Introduction to Linear Transformations

Add problems that rotate line to a direction, use $\mathbb{R}_z$ and then rotate back. Work on the x=y=z line. Show similarity works for rotations.

Recommended as prerequisites

- **Vector Space handout**
- $|V\rangle$ **represents a general vector**
- **Eigenvalue handout**

Concepts of primary interest:

- **Passive transformations vs. active operations**
- **Rotations**
- **Orthogonal transformations and real spaces**
- **Unitary transformations and complex spaces**
- **Similarity transformations - special case of congruent**
- **Diagonalizing symmetric and Hermitian matrices**
- **Transformation of scalars, vectors and matrices**
- **Scalars: Magnitudes of vectors, determinants,…**

Sample calculations:

- **Sequences of rotations**
- **Computing the form of λ that diagonalizes a matrix**

Function of a matrix – substitute matrix into

power series expansion for the function

Tools of the trade:

- **Cyclic relations:** $\mathbb{R}_{\mathbf{z}}(\phi) \rightarrow \mathbb{R}_{\mathbf{x}}(\phi) \rightarrow \mathbb{R}_{\mathbf{y}}(\phi) \rightarrow \mathbb{R}_{\mathbf{z}}(\phi) \ldots$

**ADD terms for maps injective surjective bijective

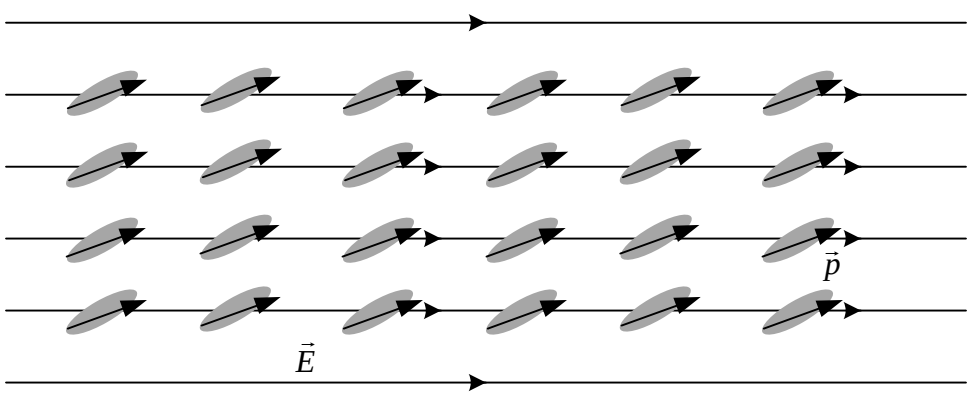
BEWARE: This handout treats the restricted case of Cartesian tensors. See references 8 and 9 for a more general treatment. It also concentrates on rotations at the expense of general transformations.

Contact: tank@alumni.rice.edu

Physical phenomena can involve relations between vector quantities that are not parallel. One such case is the relation between the electric field $\vec{E}$ applied to a molecule and the electric (dipole) polarization $\vec{p}$ of that molecule. This relation requires a matrix representation:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \qquad \text{or} \quad \mathbb{a}\ \vec{E} = \vec{p} \qquad [\text{LT.1}]$$

The symbol $\mathbb{a}$ represents the polarizability of the molecule. The matrix character of $\mathbb{a}$ arises because, for some shapes, the molecules do not polarize with equal ease in all directions. For example, if a molecule is cigar shaped, charge can move along the cigar more freely than across it. Hence if the field is applied at 45^0 relative to the long axis, the average direction for the polarization might be expected to lie between the direction of the field and that of the long axis of the cigar, a relation more complex than simple proportionality.



Prolate ellipsoidal molecules in an applied uniform electric field

This relationship is a linear mapping from the space of the $\vec{E}$'s into the space of the $\vec{p}$'s; both are three dimensional vector spaces. In the language of functions, the relation above is a function with *domain*, the space of $\vec{E}$'s, and r*ange*, the space of $\vec{p}$'s. This function could be more general involving second as well as first powers of

the E's, a case of interest for strong applied fields such as those associated with a tightly focused laser beam. For now, developments are to be restricted to the simpler cases of *linear mappings* from one vector space to another for this introduction.

The domain for the mapping is the space $\mathbb{V}$, and its target is another space $\mathbb{W}$ or perhaps a subspace of $\mathbb{W}$. That is: the function maps the space $\mathbb{V}$ into the space $\mathbb{W}$. A mapping associates every single element in the domain with a single element in the target. More than one element in the domain can be associated with (mapped into) a particular element in the target, but a single element in the domain may not be mapped into multiple elements in the target. The actual set of elements in the target that are linked to elements in the domain is the range of the mapping. For a linear mapping, the elements in $\mathbb{V}$ are mapped into a subspace of $\mathbb{W}$. (Recall that the full space $\mathbb{W}$ is itself a subset of $\mathbb{W}$ in this context. A subspace is a vector space, and every element in the subspace is also an element of the original space.)

For the space of three dimensional displacements, a rotation of 30^0 about the *z* axis maps each vector into exactly one vector. Projection into the *x-y* plane maps 'several' vectors into each vector in the target. The range of the *x-y* plane projection is itself a vector space – in this case, a space of lower dimension.

For a linear process $\mathcal{L}$, $\mathcal{L}\,[\,a|v_1\rangle + b|v_2\rangle\,] = a\,\mathcal{L}\,[\,|v_1\rangle\,] + b\,\mathcal{L}\,[\,|v_2\rangle\,]$ for $a,b \in \mathbb{K}$, the field for the vector spaces. That is: a linear transformation preserves addition and scalar multiplication. The sum of two vectors is mapped into the sum of the mappings of the two vectors. A multiple of a vector is mapped into that same multiple of the mapping of the vector. It appears that $\mathbb{W}$ might be an equivalent copy of $\mathbb{V}$. Clearly, the $|0\rangle$ element of $\mathbb{V}$ must map to the zero element of $\mathbb{W}$ as it is the 0 multiple of any vector. One might consider the mapping of a three dimensional space into a two dimensional space. Projecting three-dimensional positions onto the *x-y* plane is one example. $(x,y,z) \rightarrow (x,y,0)$. The $|0\rangle$ element of $\mathbb{V}$ maps to the $|0\rangle$ element of $\mathbb{W}$, but so

do many others. Every vector in the entire one-dimensional ***subspace*** (0, 0, z) of $\mathbb{V}$. The set of elements mapped to the zero-vector in the target space is the ***kernel*** of the mapping. The dimension of $\mathbb{V}$ is equal to the dimension of its image in $\mathbb{W}$ under the mapping plus the dimension of the kernel of the transformation.

Other complications arise if the dimension of $\mathbb{V}$ is less than the dimension of $\mathbb{W}$. In this case, $\mathbb{V}$ is mapped into a proper subspace of $\mathbb{W}$. A basis set for $\mathbb{V}$ is mapped into a spanning set for the image of $\mathbb{V}$ in $\mathbb{W}$ under the mapping indicating that the image of $\mathbb{V}$ can have at most the dimension of the original space $\mathbb{V}$. (The projection example shows that it can have a smaller dimension.)
In any case, the image of $\mathbb{V}$ in $\mathbb{W}$ under a linear mapping is a subspace of $\mathbb{W}$.

Exercise: Consider the mapping of $(x,y)\rightarrow (x,y,z)$ defined by $(x,y)\rightarrow (x,y,x\text{-}y)$. Show that the target $(x,y,x\text{-}y)$ is a two dimension subspace of (x,y,z) where x,y,z are any real numbers. Choose any basis set for (x,y). Show that its image under the linear transformation is a basis set for the image of the full space (x,y) under the transformation.
A space can never be mapped to a range of higher dimension than that of the domain by a linear mapping.

Further limiting the scope, only mappings for which the kernel contains only the zero element are to be considered. In this case each element of the initial space is mapped into a distinct element of the target space. The mapping defines a ***one-to-one*** relation between elements in the domain and range spaces and therefore has an inverse mapping matching the same element pairs in the reverse sense..

Rotations are a great example. One could grab a physical problem and rotate is by θ about the z-axis. Each vector in the space would be mapped into another vector by the rotation, and only the zero (*null*) vector would be mapped into the zero vector.

This radical an intrusion into a physical problem is not always advisable. The problem is to be studied, not disturbed. A mapping corresponding to changing the physical situation is an ***active*** mapping whereas one in which the observer changes the reference coordinate system without altering the physical situation is a ***passive*** transformation.[1] In an active mapping, one might rotate a wheel and its surroundings by +15 degrees about the z-axis. The corresponding passive transformation leaves the physical objects untouched and rotates the coordinate axes by -15 degrees about the z-axis. Adopting a new reference system with rotated axes, making a coordinate transformation, is equivalent to changing the set of basis vectors used to represent the space.

Physics, as nature presents it, comes without coordinate systems; they are added by the observer, and the particular choice is somewhat arbitrary. Coordinate transformations provide translations from the description of one observer using one coordinate system choice into the terms of another observer who chooses to use a different coordinate system. Rotations and vector behavior are particularly rich examples to illustrate transformation properties and to motivate further studies.

The nature of physical vectors: (The discussion that follows is not complete. It only introduces some concepts for initial consideration.) For the applications in physics, one crucial addition, a special transformation property is required to supplement the requirements of an abstract mathematical vector space. A collection of entities can be a **physical vector** only if its representation transforms under a rotation of coordinates (observer) in the same manner as does the representation of a

---

[1] The terms active transformation and passive transformation are more standard than the active operation or mapping and the passive transformation that are used in this handout in which a transformation is assumed to be passive. A mapping of a space into the same space is an operation (performed by an operator). The operator acts on a vector returns another vector in the same space. A transformation translates between the representations of ***a vector*** as computed using distinct basis sets for the same space. A transformation, in our terminology, does not change the vector.

spatial displacement, the prototype for a physical vector. Observer is to be synonymous with the reference frame or coordinate system used by that observer. The postulate is that the laws of physics are to be invariant under rotations of axes about the origin. Physical laws are to be expressed in terms of scalars, vectors and other things (*tensors of rank 2, …* [*scalars are tensors of rank zero and vectors are tensors of rank one, a tensors of rank two is like a matrix*]) to which undergraduates need not be seriously exposed. Scalars have the simplest transformation rule; they stay the same. All observers report the same value for a scalar entity such as mass or temperature. Next consider a vector relation, Newton's second law, $\vec{F}_{net} = m\,\vec{a}$. The mass is a scalar and so is identical according to all observers and is said to be an *invariant*. Each vector has a representation, its three components as determined by an observer. Clearly transformations such as rotations scramble the components of the vectors as reported by the different observers. Newton's law remains valid component by component only if, as the data is transformed from one observer's view to another, all the vectors have their components scrambled according to the same set of rules so that equalities are preserved. All the vectors transform together. They change in unison. Hence: A collection of entities is a *physical vector* only if its collection of components transforms under a rotation (of the coordinates axes) according to the same rules as do the components of a displacement, our reference model for a physical vector. Rotations of coordinate axes with a fixed common origin plus inversions are the classes of coordinate transformation relevant to this discussion.[1]

A collection of quantities is a physical vector <u>only</u> if they obey the transformation rule [LT.8].

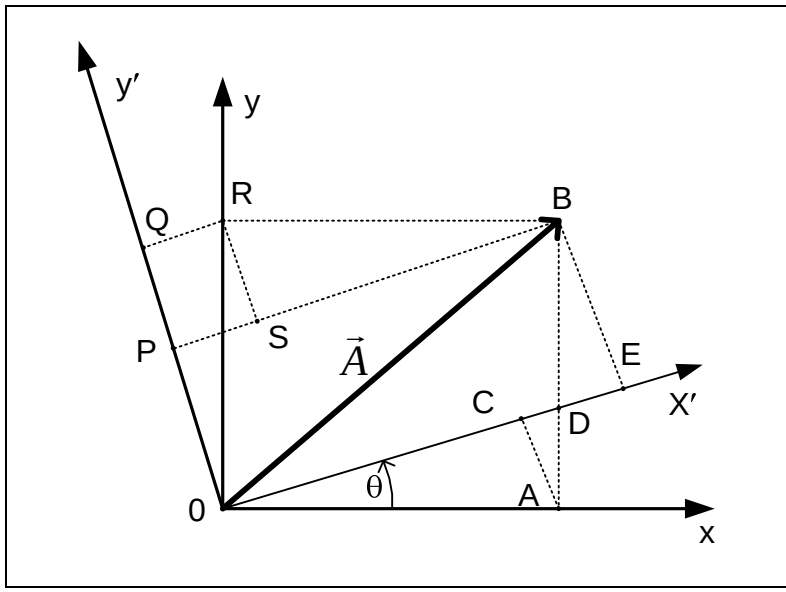
---

[1] If you wish to know more, read the first chapter of ***Classical Dynamics of Particles and Systems*** by J. B. Marion and S. T. Thornton. A shorter discussion can be found in the mechanics text by Taylor. It includes the generalization to space-time vectors.

Not all physical vectors are in the same space. Forces do not have the same dimension as accelerations. Adding a force to an acceleration makes no sense. There could be a space of accelerations and a space of forces.

A problem might involve two interacting particles with the goal being to express the force exerted on particle two by particle one. Limiting the problem to some simple classical models (one might consider the case of two particles coupled together by an ideal spring), the result could be expressed as $\vec{F}_{21}\left(\vec{r}_1, \vec{r}_2\right)$. In the trivial case in which both particles are at rest, both $\vec{r}_1$ and $\vec{r}_2$ change if the observer changes position. Clearly nothing should have changed with regard to the interaction between the particles, and it becomes clear the force should not be a function of their individual positions, $\vec{r}_1$ and $\vec{r}_2$, but rather one of $\vec{r}_2 - \vec{r}_1$, the displacement from particle one to particle two, a quantity that is independent of the position of the observer. Considerations along this line lead to the selection of displacements as the prototype for vector behavior.

Rotations about the origin are a particularly fruitful example. As a first look, the transformation from an initial reference system to a primed system with axes rotated by θ about the z-axis relative to the unprimed system is studied. The unprimed observer reports: $\vec{A} = a_x \hat{i} + a_y \hat{j} + a_z \hat{k}$. The task is to find the representation according to the primed observer: $\vec{A} = a'_x \hat{i}' + a'_y \hat{j}' + a'_z \hat{k}'$. Note: the vector $\vec{A}$ is a physical entity; it does not change as the result of a passive transformation, only its *representation* changes. In the relation: $\vec{A} = a_x \hat{i} + a_y \hat{j} + a_z \hat{k}$, the equal sign means that the right-hand side is a representation for the physical vector $\vec{A}$.



$$a_x = OA = RB$$
$$a_y = OR = AB = AD + DB$$
$$a'_x = OC + CD + DE$$
$$= OC + (AD + DB)\sin\theta$$
$$a'_x = a_x \cos\theta + a_y \sin\theta$$
$$a'_y = OP = OQ - QP$$
$$= OQ - RB \sin\theta$$
$$a'_y = a_y \cos\theta - a_x \sin\theta$$

With some effort, one finds that $a'_x = a_x \cos\theta + a_y \sin\theta,\ a'_y = a_x(-\sin\theta) + a_y \cos\theta, a'_z = a_z$. The trivial transformation for the z-component results because primed observer frame was rotated by θ about the z-axis of the original unprimed observer. The two z axes coincide.

$$\begin{bmatrix} a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \text{[LT.2]}$$

A more automated procedure to develop these rotation matrices is needed. Projection, one of our standard procedures, is the tool of choice for this task. To simplify the notation the subscripts 1, 2, 3 are to be substituted for *x, y, z*. Consider a vector $\vec{v}$ represented in an original 'fixed' system and in a primed system rotated about the origin relative to the first. It is the representation of the vector $\vec{v}$, its components and the coordinate directions that change, not the physical vector $\vec{v}$. The physical vector has a distinct representation in each distinct frame.

$$\vec{v} = v'_x\,\hat{i}' + v'_y\,\hat{j}' + v'_z\,\hat{k}' = v_x\,\hat{i} + v_y\,\hat{j} + v_z\,\hat{k} \text{ and } \vec{v} = v'_1\,\hat{i}' + v'_2\,\hat{j}' + v'_3\,\hat{k}' = v_1\,\hat{i} + v_2\,\hat{j} + v_3\,\hat{k}$$

The *x* or 1 component in the primed frame is **projected** out of the sum using the inner product and the mutual orthogonality of each set of coordinate directions.

$$\hat{i}' \cdot \vec{v} = \hat{i}' \cdot \left(v'_x \, \hat{i}' + v'_y \, \hat{j}' + v'_z \, \hat{k}' \right) = \hat{i}' \cdot \left(v_1 \, \hat{i} + v_2 \, \hat{j} + v_3 \, \hat{k} \right)$$

$$\hat{i}' \cdot \vec{v} = v'_1 \left(\hat{i}' \cdot \hat{i}' \right) + v'_2 \left(\hat{i}' \cdot \hat{j}' \right) + v'_3 \left(\hat{i}' \cdot \hat{k}' \right) = v_1 \left(\hat{i}' \cdot \hat{i} \right) + v_2 \left(\hat{i}' \cdot \hat{j} \right) + v_3 \left(\hat{i}' \cdot \hat{k} \right)$$

$$\hat{i}' \cdot \vec{v} = v'_1 = v_1 \left(\hat{i}' \cdot \hat{i} \right) + v_2 \left(\hat{i}' \cdot \hat{j} \right) + v_3 \left(\hat{i}' \cdot \hat{k} \right)$$

These relations can be collected into a compact matrix form:

$$\begin{bmatrix} v'_1 \\ v'_2 \\ v'_3 \end{bmatrix} = \begin{bmatrix} (\hat{i}' \cdot \hat{i}) & (\hat{i}' \cdot \hat{j}) & (\hat{i}' \cdot \hat{k}) \\ (\hat{j}' \cdot \hat{i}) & (\hat{j}' \cdot \hat{j}) & (\hat{j}' \cdot \hat{k}) \\ (\hat{k}' \cdot \hat{i}) & (\hat{k}' \cdot \hat{j}) & (\hat{k}' \cdot \hat{k}) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \qquad \text{[LT.3]}$$

The transformation can be represented as matrix multiplication:

$$\begin{bmatrix} v'_1 \\ v'_2 \\ v'_3 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad \text{where } \lambda_{ij} = \hat{e}'_i \cdot \hat{e}_j$$

The element λ_{mn} is $\hat{e}'_m \cdot \hat{e}_n = \cos(\hat{e}'_m, \hat{e}_n)$, the cosine of the angle between the *m* axis in the rotated (primed) system and the *n* axis in the original system. The mth row is the direction cosines of the mth primed axis relative to the unprimed axes, and the nth column is the direction cosines of the nth unprimed axis relative to the primed system axes. The mechanics text by J. B. Marion and S. T. Thornton is recommended for further reading on this point.

Consider the transformation of $\hat{i}$.

$$\begin{bmatrix} (\hat{i}' \cdot \hat{i}) \\ (\hat{j}' \cdot \hat{i}) \\ (\hat{k}' \cdot \hat{i}) \end{bmatrix} = \begin{bmatrix} (\hat{i}' \cdot \hat{i}) & (\hat{i}' \cdot \hat{j}) & (\hat{i}' \cdot \hat{k}) \\ (\hat{j}' \cdot \hat{i}) & (\hat{j}' \cdot \hat{j}) & (\hat{j}' \cdot \hat{k}) \\ (\hat{k}' \cdot \hat{i}) & (\hat{k}' \cdot \hat{j}) & (\hat{k}' \cdot \hat{k}) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \; ; \; \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The entries in a column are just the components of one of the coordinate directions (unit vectors) as measured by the primed observer. Hence, the sum $\sum_{i=1 to 3} \lambda_{i\ell} \lambda_{im}$

represents the inner product of the ℓ and m unprimed coordinate directions. Therefore $\sum_{i=1to3} \lambda_{i\ell}\lambda_{im} = \delta_{\ell m}$ as those coordinate directions are mutually orthogonal and normalized.

$$\sum_{i=1to3} \lambda_{i\ell}\lambda_{im} = \delta_{\ell m} = \sum_{i=1to3} \lambda^t_{\ell i}\lambda_{im} \quad \Rightarrow \quad \lambda^{\mathbf{t}}\ \lambda = \mathbb{1}$$

The inverse transformation λ^{-1} is found by interchanging the primed and unprimed coordinate directions.

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} (\hat{i}\cdot\hat{i}') & (\hat{i}\cdot\hat{j}') & (\hat{i}\cdot\hat{k}') \\ (\hat{j}\cdot\hat{i}') & (\hat{j}\cdot\hat{j}') & (\hat{j}\cdot\hat{k}') \\ (\hat{k}\cdot\hat{i}') & (\hat{k}\cdot\hat{j}') & (\hat{k}\cdot\hat{k}') \end{bmatrix} \begin{bmatrix} v'_1 \\ v'_2 \\ v'_3 \end{bmatrix} \qquad \text{[LT.4]}$$

It is therefore trivially obvious to the casual observer upon cursory examination that the inverse transformation matrix is the ***transpose*** of the original transformation matrix. Consider the role of each element in the matrix. Note that the inverse just transposes the roles of the coordinate directions in the primed and unprimed systems.

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \qquad \lambda^{\mathbf{t}} = \lambda^{\mathbf{-1}}$$

and as $\sum_{i=1,,,3} \lambda_{\ell i}\lambda_{mi} = \sum_{i=1,,,3} \lambda_{\ell i}\lambda^t_{im} = \delta_{\ell m} \Rightarrow \lambda\lambda^{\mathbf{t}} = \mathbb{1}.$ [LT.5]

Any matrix that has the property $\mathbb{M}^{\mathbf{t}} = \mathbb{M}^{\mathbf{-1}}$ is an ***orthogonal*** matrix. The property of determinants that $|\mathbb{M}^{\mathbf{t}}| = |\mathbb{M}|$ ensures that $|\mathbb{M}| = \pm 1$ for an orthogonal matrix. The matrices that represent rotations are real, orthogonal and have determinant 1. The matrices that represent inversions are real orthogonal and have determinant -1. For now, only rotations in real vector spaces are considered. The transformation matrices considered are real and orthogonal.

The term orthogonal matrix is appropriate because the rows of an orthogonal matrix are the components of a set of mutually orthogonal vectors. The same can be said

about the columns of an orthogonal matrix.

Exercise: Use the identification $\lambda_{ij} = \cos(\hat{e}'_i, \hat{e}_j)$ to develop the matrix equation appropriate for a rotation of the observer frame by θ about the z-axis.

$\{\hat{i}, \hat{j}, \hat{k}\} \leftrightarrow \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$

Compare with: Compare with: $a'_x = a_x \cos\theta + a_y \sin\theta,\ a'_y = a_x(-\sin\theta) + a_y \cos\theta, a'_z = a_z$.

Repeat to find $\mathbb{R}_x(\phi)$ for a rotation of the observer frame by ϕ about the *x*-axis.

Exercise: Given the matrix $\mathbb{R}_z(\theta)$ that represents a rotation of the observer frame by θ about the z-axis, the inverse should be of the same form but with θ replaced by -θ. Show explicitly that the matrix obtained by this replacement is the transpose of the original and also is its inverse.

Exercise: The transformation character of a displacement is the prototype for the transformation properties of all vectors. Consider a displacement from the origin which is common to the original coordinate system and to the rotated system.

$$\Delta\vec{r} = x'_1\ \hat{i}' + x'_2\ \hat{j}' + x'_3\ \hat{k}' = x_1\,\hat{i} + x_2\,\hat{j} + x_3\,\hat{k}$$

As displacements are vectors, $x'_i = \sum_j \lambda_{ij}\ x_j$. The inverse transformation is:

$$x_i = \sum_j \lambda^{-1}_{ij}\ x'_j = \sum_j \lambda^{t}_{ij}\ x'_j = \sum_j \lambda_{ji}\ x'_j\,.$$

Evaluate $\left(\frac{\partial x'_m}{\partial x'_n}\right)$ which has two primes. Use the result compute $\left(\frac{\partial x_i}{\partial x'_j}\right)$ with $x_i = \sum_k \lambda_{ki}\ x'_k$.

The relation between the λ's and $\left(\frac{\partial x_i}{\partial x'_j}\right)$ is important. Study it; save it!

Note that the transformation rule is based on the rule for the transformation of the components of displacements. All physical vectors must use this same coordinate based rule. A simple consequence is that

a vector must have as many components as the coordinate space has dimensions. Some of those components may be zero in a particular representation (coordinate system), but there are always three components for a physical vector in a three dimensional universe. This statement is not intended to conflict with the practice of restricting problems to two (or even one) dimensions with a corresponding reduction in the number of components for vectors such as force and velocity. Further, in special relativity, physicists work in a four-dimensional space-time vector space. Physical vectors are 4D when considered in (special) relativity and must transform according to the same Lorentz transformations that apply to displacements between events (space-time intervals).

The problem that brought us here is:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \qquad \text{or} \quad \mathbb{a}\ \vec{E} = \vec{p}$$

It is a mapping of one vector space into another. The vectors are physical vectors however so we should examine the changes in representation under a coordinate transformation. Above, $\mathbb{\lambda}$ represented a rotation applied to the reference frame. Here $\mathbb{S}$ is to represent a rotation or perhaps a more general coordinate transformation. Each vector must transform according to the rule so

$$\begin{bmatrix} \tilde{E}'_x \\ \tilde{E}'_y \\ \tilde{E}'_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} \qquad \tilde{E}'_i = \sum_j S_{ij}\, \tilde{E}_j$$

Notation Review: the symbols with twiddles above column matrices represent the array of values that would be reported by a particular observer. The unprimed observer represents the actual physical vector $\vec{E}$ as the array of values $\tilde{E} = \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}$ while the primed observer reports $\tilde{E}' = \begin{bmatrix} \tilde{E}'_x \\ \tilde{E}'_y \\ \tilde{E}'_z \end{bmatrix}$.

In the polarization problem, $\vec{E}$ and $\vec{p}$ represent physical vectors which exist independent of observers. $\vec{E}$ has the component representation $\tilde{E}$ according to the unprimed observer and $\tilde{E}^{/}$ according to the primed observer. The same relation holds for the components of $\vec{p}$ as recorded by the two observers.

$$\begin{bmatrix} \tilde{p}'_x \\ \tilde{p}'_y \\ \tilde{p}'_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \tilde{p}_x \\ \tilde{p}_y \\ \tilde{p}_z \end{bmatrix} \qquad \tilde{p}'_i = \sum_j S^{-1}_{ij}\, \tilde{p}_j \qquad \tilde{p}' = \mathbb{S}\ \tilde{p}$$

$$\begin{bmatrix} \tilde{p}_x \\ \tilde{p}_y \\ \tilde{p}_z \end{bmatrix} = \begin{bmatrix} S^{-1}_{11} & S^{-1}_{12} & S^{-1}_{13} \\ S^{-1}_{21} & S^{-1}_{22} & S^{-1}_{23} \\ S^{-1}_{31} & S^{-1}_{32} & S^{-1}_{33} \end{bmatrix} \begin{bmatrix} \tilde{p}'_x \\ \tilde{p}'_y \\ \tilde{p}'_z \end{bmatrix} \qquad \tilde{p}_i = \sum_j S_{ij}\, \tilde{p}'_j \qquad \tilde{p} = \mathbb{S}^{-1}\ \tilde{p}'$$

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \qquad \text{or} \qquad \mathbb{a}\ \vec{E} = \vec{p}$$

The components of all physical vectors must follow the same transformation rule for a change of observer. What rule do the elements in the $\tilde{\tilde{\alpha}}$ and $\tilde{\tilde{\alpha}}'$, representations of $\mathbb{a}$, follow?

Notation Preview: the symbols with twiddles above represent row, column or square matrices of the components of a vector or elements of a matrix represent the array of values that would be reported by a particular observer. The unprimed observer represents the actual physical property $\mathbb{a}$ as the array of values $\tilde{\tilde{\alpha}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$

while the primed observer reports $\tilde{\tilde{\alpha}}^{/} = \begin{bmatrix} \alpha^{/}_{11} & \alpha^{/}_{12} & \alpha^{/}_{13} \\ \alpha^{/}_{21} & \alpha^{/}_{22} & \alpha^{/}_{23} \\ \alpha^{/}_{31} & \alpha^{/}_{32} & \alpha^{/}_{33} \end{bmatrix}$.

** The tildas (~) may appear ***either above or below*** the symbol. Conclude that this notation has not been implemented consistently. The management is working on the issue. $\Rightarrow \tilde{x} \equiv \underset{\sim}{x}$, …

We want to find $\tilde{\tilde{\alpha}}'$ such that: $\begin{bmatrix} \alpha'_{11} & \alpha'_{12} & \alpha'_{13} \\ \alpha'_{21} & \alpha'_{22} & \alpha'_{23} \\ \alpha'_{31} & \alpha'_{32} & \alpha'_{33} \end{bmatrix} \begin{bmatrix} \tilde{E}'_x \\ \tilde{E}'_y \\ \tilde{E}'_z \end{bmatrix} = \begin{bmatrix} \tilde{p}'_x \\ \tilde{p}'_y \\ \tilde{p}'_z \end{bmatrix}$ First, pre-multiply by $\mathbb{S}$:

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} S^{-1}_{11} & S^{-1}_{12} & S^{-1}_{13} \\ S^{-1}_{21} & S^{-1}_{22} & S^{-1}_{23} \\ S^{-1}_{31} & S^{-1}_{32} & S^{-1}_{33} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

Using: $\begin{bmatrix} \tilde{E}'_x \\ \tilde{E}'_y \\ \tilde{E}'_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}$ and $\begin{bmatrix} \tilde{p}'_x \\ \tilde{p}'_y \\ \tilde{p}'_z \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \tilde{p}_x \\ \tilde{p}_y \\ \tilde{p}_z \end{bmatrix}$

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} S^{-1}_{11} & S^{-1}_{12} & S^{-1}_{13} \\ S^{-1}_{21} & S^{-1}_{22} & S^{-1}_{23} \\ S^{-1}_{31} & S^{-1}_{32} & S^{-1}_{33} \end{bmatrix} \begin{bmatrix} \tilde{E}'_x \\ \tilde{E}'_y \\ \tilde{E}'_z \end{bmatrix} = \begin{bmatrix} \tilde{p}'_x \\ \tilde{p}'_y \\ \tilde{p}'_z \end{bmatrix}$$

So we find:

$$\begin{bmatrix} \alpha'_{11} & \alpha'_{12} & \alpha'_{13} \\ \alpha'_{21} & \alpha'_{22} & \alpha'_{23} \\ \alpha'_{31} & \alpha'_{32} & \alpha'_{33} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} S^{-1}_{11} & S^{-1}_{12} & S^{-1}_{13} \\ S^{-1}_{21} & S^{-1}_{22} & S^{-1}_{23} \\ S^{-1}_{31} & S^{-1}_{32} & S^{-1}_{33} \end{bmatrix}$$

$$\tilde{\tilde{\alpha}}' = \mathbb{S}\ \tilde{\tilde{\alpha}}\ \ \mathbb{S}^{-1}. \qquad \text{[LT.6]}$$

Notation Review: the symbols with twiddles above represent row, column or square matrices of the components of a vector or elements of a matrix represent the array of values that would be reported by a particular observer. The symbol $\tilde{F}$ is the collection of components as measured by an observer for the vector $\vec{F}$ using his (unprimed) set of axes while another (primed) observer records the components of the same vector as $\tilde{F}'$ relative to his different (primed) set of axes.

$$F_x\,\hat{i} + F_y\,\hat{j} + F_z\,\hat{k} = \vec{F} = F_x'\,\hat{i}' + F_y'\,\hat{j}' + F_z'\,\hat{k}'$$

When to polarizabilty is considered, the unprimed observer represents the actual physical property $\mathbb{\alpha}$ as the array of values $\tilde{\tilde{\alpha}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$ while the primed observer reports $\tilde{\tilde{\alpha}}' = \begin{bmatrix} \alpha_{11}' & \alpha_{12}' & \alpha_{13}' \\ \alpha_{21}' & \alpha_{22}' & \alpha_{23}' \\ \alpha_{31}' & \alpha_{32}' & \alpha_{33}' \end{bmatrix}$.

The components of all physical vectors must follow the same transformation rule for a change of observer. A more compact derivation of the rules for $\mathbb{\alpha}$ follows?

The relation in the original coordinate system is $\tilde{p} = \tilde{\tilde{\alpha}}\ \tilde{E}$. The same rule in the new system has the form: $\tilde{p}' = \tilde{\tilde{\alpha}}'\ \tilde{E}'$. Under our transformation rule for **all** vectors, $\tilde{p}' = \mathbb{S}\ \tilde{p}$ and $\tilde{E}' = \mathbb{S}\ \tilde{E}$. Combining, $\tilde{p}' = \mathbb{S}\ \tilde{\tilde{\alpha}}\ \tilde{E}$. To relate $\tilde{p}'$ to $\tilde{E}'$, we recall that $\tilde{E}' = \mathbb{S}\ \tilde{E}$. Inserting $\mathbb{1} = \mathbb{S}^{-1}\,\mathbb{S}$, the equation becomes $\tilde{p}' = \mathbb{S}\ \tilde{\tilde{\alpha}}\ \ \mathbb{S}^{-1}\,\mathbb{S}\ \tilde{E}'$. Comparing with $\tilde{p}' = \tilde{\tilde{\alpha}}'\,\tilde{E}'$ => $\tilde{\tilde{\alpha}}' = \mathbb{S}\ \tilde{\tilde{\alpha}}\ \ \mathbb{S}^{-1}$. The components of the polarizabilty tensor transform by pre-multiplication by $\mathbb{S}$ and post-multiplication by $\mathbb{S}^{-1}$. This bizarre process is called a **similarity** transformation. Remember that term!

A general treatment of all possible transformations would be confusing; let's drop back to just rotations for which $\lambda^{-1} = \lambda^{t}$ and review the rules. The transformation $\mathbb{S}$ is now restricted to be an orthogonal transformation λ leading to the component form of the rule:

$$\tilde{\tilde{\alpha}}' = \lambda\ \tilde{\tilde{\alpha}}\ \ \lambda^{-1} \text{ or } \tilde{\tilde{\alpha}}'_{ij} = \sum_{m,\ell} \lambda_{im}\, \tilde{\tilde{\alpha}}_{m\ell}\, \lambda^{t}_{\ell j} = \sum_{m,\ell} \lambda_{im}\, \tilde{\tilde{\alpha}}_{m\ell}\, \lambda_{j\ell} = \sum_{m,\ell} \lambda_{im}\, \lambda_{j\ell}\, \tilde{\tilde{\alpha}}_{m\ell}$$

Important Observation #1: Considering $\tilde{\tilde{\alpha}}$ as a matrix, what is the relation between the determinants $|\tilde{\tilde{\alpha}}\ |$ and $|\tilde{\tilde{\alpha}}'|$? The determinant of a matrix is a scalar as is the magnitude of a vector. How do we expect scalars to transform?

Important Observation #2: In classical physics, matrices representing the properties of physical entities are expected to be real symmetric matrices (for most cases of interest). It is important to note that the real symmetric property of a matrix is preserved under a similarity transformation using a real orthogonal matrix (See problem 2). $\mathbb{M}' = \lambda\, \mathbb{M}\, \lambda^{-1}$, $\mathbb{M}'$ is real symmetric if $\mathbb{M}$ is real symmetric and λ is a transformation matrix with real elements. $\mathbb{M}' = \lambda\, \mathbb{M}\, \lambda^{-1}$, $\mathbb{M}'$ is real symmetric if $\mathbb{M}$ is real symmetric and λ is a transformation matrix with real elements.

Important Observation #3: Quantum Mechanics requires a vector space with complex scalars. Matrices representing physical entities are expected to be Hermitian matrices ($a_{ij} = a_{ji}^{*}$). It is important to note that the Hermitian property of a matrix is preserved by a similarity transformation using *unitary* matrices. One problem for this section is to prove that the Hermitian property is preserved by similarity transformations using unitary matrices. $\mathbb{O}' = \mathbb{U}\, \mathbb{O}\, \mathbb{U}^{-1}$

Important Observation #4: The trace of a matrix is defined to be the sum of its diagonal elements. Tr[$\mathbb{M}$] $= \sum_{ij} \delta_{ij}\, m_{ij}$ The trace of a matrix is invariant under similarity transformation. One problem for this section is to prove this property. The trace of a square matrix is a scalar invariant.

Terminology: A matrix transformation of the form $\mathbb{M}' = \mathbb{S}\,\mathbb{M}\,\mathbb{S}^{-1}$ is called a similarity transformation.

Scalars: Physical entities or fields that can be expressed by a single numerical value (with appropriate dimensions) at each point in space and which have the same value according to all observers. Examples are temperature and mass.

$$T(x', y', z') = T(x, y, z) \qquad \text{[LT.7]}$$

The transformation rule for scalars is 'they are invariant'. As scalars transform using zero powers λ of they are designated tensors of rank zero.

Vectors are entities or fields that have as many components as the dimension of the universe with these components transforming according to the same rule that is followed by displacements are vectors.

$$\tilde{F}' = \lambda\, \tilde{F} \qquad \tilde{F}_i' = \sum_j \lambda_{ij}\, \tilde{F}_j \qquad \text{[LT.8]}$$

All physical vectors transform using the identical rule. As vectors transform using one power of λ they are designated tensors of rank one.

Tensors (of rank *2*): Physical entities or fields that have as many components as the dimension of the universe raised to the second power with these components following the transformation rule are tensors.

$$\tilde{\tilde{\alpha}}' = \lambda \; \tilde{\tilde{\alpha}} \;\; \lambda^{\mathbf{-1}} = \lambda \; \tilde{\tilde{\alpha}} \;\; \lambda^{\mathbf{t}} \; ; \;\; \tilde{\tilde{\alpha}}'_{ij} = \sum_{m,\ell} \lambda_{im} \, \tilde{\tilde{\alpha}}_{m\ell} \, \lambda^{t}_{\ell j} = \sum_{m,\ell} \lambda_{im} \, \tilde{\tilde{\alpha}}_{m\ell} \, \lambda_{j\ell} = \sum_{m,\ell} \lambda_{im} \, \lambda_{j\ell} \, \tilde{\tilde{\alpha}}_{m\ell}$$

As these tensors transform using two factors of λ they are designated tensors of rank two. (The calculus of tensors includes generalizations of the gradient, curl, divergence and so forth; they are not to be mentioned again after the next paragraph.)

Tensors of rank *r*: Physical entities or fields that have as many components as the dimension of the universe to the *r* power **and** that have components which obey the transformation rule:

$$\tilde{\tilde{\tilde{\alpha}}}'_{i_1 i_2 \ldots i_r} = \sum_{j_1 j_2 \ldots j_r} \lambda_{i_1 j_1} \lambda_{i_1 j_2} \ldots \lambda_{i_r j_r} \, \tilde{\tilde{\tilde{\alpha}}}_{j_1 j_2 \ldots j_r} \qquad \text{[LT.9]}$$

Scalars are tensors of rank zero. Vectors are tensors of rank one. '*Tensors*' are tensors of rank two. ***Tensors of all ranks transform using factors of the same transformation λ. Hence tensors can be multiplied to yield tensors of higher rank or contracted (inner product taken) to form tensors of lower rank.*** Consider two vectors $\vec{A} = A_x \, \hat{i} + A_y \, \hat{j} + A_z \, \hat{k}$ and $\vec{B} = B_x \, \hat{i} + B_y \, \hat{j} + B_z \, \hat{k}$. The inner product of the two rank one tensors is a scalar (rank zero). $\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z$. The direct product of two vectors yields a tensor of rank two. The matrix representations of these operations are:

CONTRACTION
INNER PRODUCT

$$\begin{bmatrix} A_x & A_y & A_z \end{bmatrix} \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = \vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z$$

This (contraction) inner product forms a lower rank tensor. In this case it is a scalar.

DIRECT PRODUCT

$$\begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \begin{bmatrix} A_x & A_y & A_z \end{bmatrix} = \begin{bmatrix} B_x A_x & B_x A_y & B_x A_z \\ B_y A_x & B_y A_y & B_y A_z \\ B_z A_x & B_z A_y & B_z A_z \end{bmatrix}$$

This direct product yields a higher rank tensor called a *dyadic*.

Using the displacement from a common origin as our vector

$$\vec{r} = x\,\hat{i} + y\,\hat{j} + z\,\hat{k} = x_1\,\hat{i} + x_2\,\hat{j} + x_3\,\hat{k}$$

The inner product contraction is the scalar

$$\begin{bmatrix} x & y & z \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{r}\cdot\vec{r} = xx + yy + xx = x_1x_1 + x_2x_2 + x_3x_3$$

and the direct product is the dyadic.

$$\mathbb{d} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} x_1x_1 & x_1x_2 & x_1x_3 \\ x_2x_1 & x_2x_2 & x_2x_3 \\ x_3x_1 & x_3x_2 & x_3x_3 \end{bmatrix} = \tilde{\tilde{d}}$$ which has elements $[\,\tilde{\tilde{d}}\,]_{ij} = x_i x_j$.

It is easy to show that the contraction of two vectors transforms as a scalar. That is: it is an invariant.

Exercise: Using $x_i' = \sum_{j=1}^{3} \lambda_{ij}\, x_j$, show that $\vec{r}\cdot\vec{r} = \sum_{j=1}^{3} x_j\, x_j \equiv \sum_{k=1}^{3} x_k'\, x_k'$. Assume that the λ_{ij} are the elements of the matrix representing an real orthogonal transformation.

The direct product $\mathbb{d}$ transforms as a rank two tensor. Or more precisely, the ***representation*** of the dyadic transforms in accord with the rule for rank two tensors. $\underset{\approx}{d}' = \lambda\, \underset{\approx}{d}\, \lambda^{-1}$. Recall $x_i' = \sum_j \lambda_{ij}\, x_j$ so:

$$d_{ij}' = x_i' x_j' = \sum_{m,\ell} \lambda_{im}\, x_m \lambda_{j\ell} x_\ell = \sum_{m,\ell} \lambda_{im}\, x_m x_\ell \lambda_{j\ell}$$

$$= \sum_{m,\ell} \lambda_{im}\, x_m x_\ell \lambda^t_{\ell j} = \sum_{m,\ell} \lambda_{im}\, x_m x_\ell \lambda^{-1}_{\ell j} = \sum_{m,\ell} \lambda_{im}\, d_{m\ell}\, \lambda^{-1}_{\ell j}$$

How does one prove that a quantity transforms properly? The quantity must be defined by some prescription. The coordinate dyadic above is the direct product of the position vector with itself. That is the unprimed observer assigns the array of values $\underset{\approx}{d}_{ij} = x_i\, x_j$ as his representation of $\mathbb{d}$. Using the same prescription, but in his terms, the

primed observer assigns $\underset{\approx}{d}'_{ij} = x_i' \, x_j'$ as his representation of $\mathbb{d}$. The quantity is a proper tensor if the ***representations*** transform canonically $\Rightarrow \underset{\approx}{d}' = \lambda \, \underset{\approx}{d} \, \lambda^{-1}$.

It is crucial that physical tensors transform using the same basic transformation λ as do vectors; they just use more copies of it. This commonality means it is possible to combine vectors to form tensors and to take the inner product of a rank two tensor with a vector to yield a vector (contract a vector and a rank two tensor).

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \qquad \text{or} \quad \mathbb{a} \, \vec{E} = \vec{p} \quad \Rightarrow \quad p_i = \sum_j \alpha_{ij} \, E_j$$

Summing over the common index between the polarizability tensor and the electric field executes a contraction or a rank 2 tensor with a rank 1 tensor to yield a rank 1 tensor.

Exercise: One tensor quantity is the quadrupole moment of a charge distribution. It has elements

$$\mathbb{Q} \rightarrow \sum_{\alpha=1}^{N} q_\alpha \left[3 \, x_{\alpha i} \, x_{\alpha j} - r_\alpha^2 \, \delta_{ij} \right]$$

The charges q_α are scalar values. The first term is therefore a scalar times the coordinate dyadic $x_i x_j$ which has been shown to transform properly as a rank two tensor. Show that δ_{ij} transforms as a tensor of rank under rotations so that overall $\mathbb{Q}$ is a properly transforming rank two tensor.

When the term tensor appears without a rank specified, rank two is usually assumed. Avoid using the terms tensor of rank zero for scalars and tensor of rank one for vectors as the practice causes tremors, sweats and nightmares. Substitute the kinder, gentler term 'physical matrix' for a tensor of rank two whenever possible.

The gradient of a scalar appears to be a scalar. Does it transform properly?

Recall that $x_j = [\lambda^{-1}]_{ji}\, x'_i = \lambda_{ij}\, x'_i$. It follows that $\dfrac{\partial x_j}{\partial x'_i} = \lambda_{ij}$.

Proof: The gradient of a scalar function transforms as a vector.

$$\left(\vec{\nabla} f\right)'_i = \frac{\partial f}{\partial x'_i} = \sum_{j=1}^{3}\left(\frac{\partial f}{\partial x_j}\right)\left(\frac{\partial x_j}{\partial x'_i}\right) = \sum_{j=1}^{3}\left(\frac{\partial x_j}{\partial x'_i}\right)\left(\frac{\partial f}{\partial x_j}\right) = \sum_{j=1}^{3} \lambda_{ij}\left(\frac{\partial f}{\partial x_j}\right) = \sum_{j=1}^{3} \lambda_{ij}\left(\vec{\nabla} f\right)'_j$$

Compare with the rules that must be obeyed by vectors. $\tilde{F}'_i = \sum_j \lambda_{ij}\, \tilde{F}_j$

The gradient of a scalar function transforms as a vector $\Rightarrow$ it is a vector.

Some properties of rotations:

Rotations (all orthogonal transformations) preserve inner products and hence all lengths and angular separations. $\vec{A}\cdot\vec{B} = AB\cos\theta_{AB}$

Consider

$$\vec{A}'\cdot\vec{B}' = \sum_{i=1}^{3} A'_i\, B'_i = \sum_{i=1}^{3}\sum_{j=1}^{3}\lambda_{ij} A_j \sum_{k=1}^{3}\lambda_{ik} B_k = \sum_{j,k=1}^{3} A_j\, B_k \sum_{i=1}^{3}\lambda_{ij}\lambda_{ik}$$

$$= \sum_{j,k=1}^{3} A_j\, B_k\, \delta_{jk} = \sum_{j=1}^{3} A_j\, B_j = \vec{A}\cdot\vec{B} = AB\cos\theta$$

Used the rotation (orthogonal transformation) property:

$$\sum_{i=1,,,3}\lambda_{\ell i}\lambda_{mi} = \sum_{i=1,,,3}\lambda_{\ell i}\lambda^{t}_{im} = \sum_{i=1,,,3}\lambda_{\ell i}\lambda^{-1}_{im} = [\mathbb{1}]_{\ell m} = \delta_{\ell m}$$

Sequences of rotations are to be studies. To begin, the matrices representing a few rotations are displayed.

$^{\pi}/_{2}$ about the three axis $\quad \lambda\,(^{\pi}/_{2}, 3) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbb{R}_z(^{\pi}/_{2})$

$^{\pi}/_{2}$ about the two axis $\quad \lambda\,(^{\pi}/_{2},2) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \mathbb{R}_y(^{\pi}/_{2})$

$^{\pi}/_{2}$ about the one axis $\quad \lambda\,(^{\pi}/_{2},1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} = \mathbb{R}_x(^{\pi}/_{2})$

π about the two axis $\quad \lambda\,(\pi,2) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \mathbb{R}_y(\pi)$

Sequences of Rotations:

Consider the case of a primed frame rotated by θ about the z-axis relative to the unprimed frame and a double-primed frame rotated by ϕ about the z-axis relative to the primed frame. The matrix for rotation by θ acts on the left of the vector as does the subsequent rotation by ϕ.

$$\begin{bmatrix} a_1' \\ a_2' \\ a_3' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad \begin{bmatrix} a_1'' \\ a_2'' \\ a_3'' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1' \\ a_2' \\ a_3' \end{bmatrix}$$

Applying the transformations in sequence leads to:

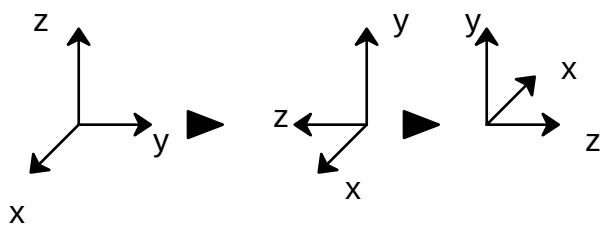
$$\begin{bmatrix} a_1'' \\ a_2'' \\ a_3'' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

The matrix representing the sequence is formed by sequentially multiplying ***on the left*** with the matrices for each operation. Operations applied later appear further to the left in the matrix product. See problem 1.

Consider the sequence $^{\pi}/_{2}$ about the one axis followed by π about the two axis.

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

'In sequence' means that each subsequent operation is represented by multiplication of the left side.



initial **after FIRST** **after SECOND**

Exercise: The matrix representing the sequence of rotations can be read as row one is the components of the new *x*-axis as represented in terms of the original axes. The final matrix may also be compared to the template form: **MULTIPLY SEQUENTIALLY ON THE LEFT.**

$$\lambda_{\mathbf{NET}} = \lambda_{\mathbf{SECOND}}\,\lambda_{\mathbf{FIRST}} = \begin{bmatrix} (\hat{i}''\cdot\hat{i}) & (\hat{i}''\cdot\hat{j}) & (\hat{i}''\cdot\hat{k}) \\ (\hat{j}''\cdot\hat{i}) & (\hat{j}''\cdot\hat{j}) & (\hat{j}''\cdot\hat{k}) \\ (\hat{k}''\cdot\hat{i}) & (\hat{k}''\cdot\hat{j}) & (\hat{k}''\cdot\hat{k}) \end{bmatrix}$$

where the final axes are the double primed directions and the original coordinate directions are unprimed. Check to see that the result above agrees with these forms.

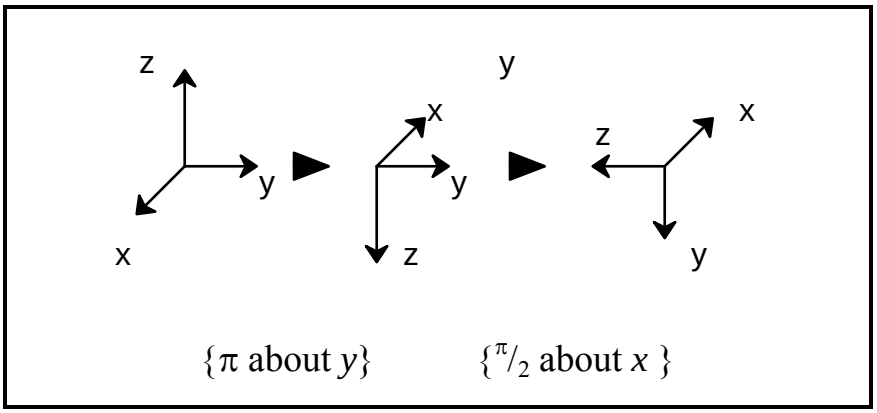
Contrast the first sequence with the sequence π about the two axis followed by $^{\pi}/_{2}$ about the one axis.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$\{^{\pi}/_{2}$ about $x\}$ $\{\pi$ about $y\}$ = *net rotation*

$$\lambda_{\mathbf{SECOND}}\,\lambda_{\mathbf{FIRST}} = \lambda_{\mathbf{NET}}$$

initial **after FIRST** **after SECOND**



Demonstrate this behavior by imagining the *x, y* and *z*-axis to be rigidly fixed to your textbook and executing the rotation sequence. Repeat imagining that the axis directions are fixed in space and only the text is rotated. In the passive view, the text remains fixed in space and the rotations are applied in sequence to the reference axes. This view is illustrated above. Viewed in anyway you wish, the actions of rotations do not satisfy a commuting addition and hence cannot be properly represented as vectors, but rather require a matrix representation.

Angular velocities and infinitesimal rotations can be represented as vectors for most purposes. The violation of the commutative property is proportional to the square of the rotation angle and is therefore negligible for infinitesimal rotations. More is to be presented later.

Multiplicity: There are an infinite number of rotation sequences that yield any specified reorientation of a set of coordinate axes in spite of the fact that only three angles are needed to set the final orientation. One could take two angles to set the direction of the final z-axis and one more to set the orientation about this axis. The accepted convention is somewhat different. Reorient the axes in three steps. First: rotate by ϕ about the *z*-axis. Second: rotate by θ about the new (or current) *x*-axis. Three: rotate by ψ about the new (or current) *z*-axis. These angles are the Euler (or Eulerian) angles. Be alert to the introduction of Eulerian angles in your study of

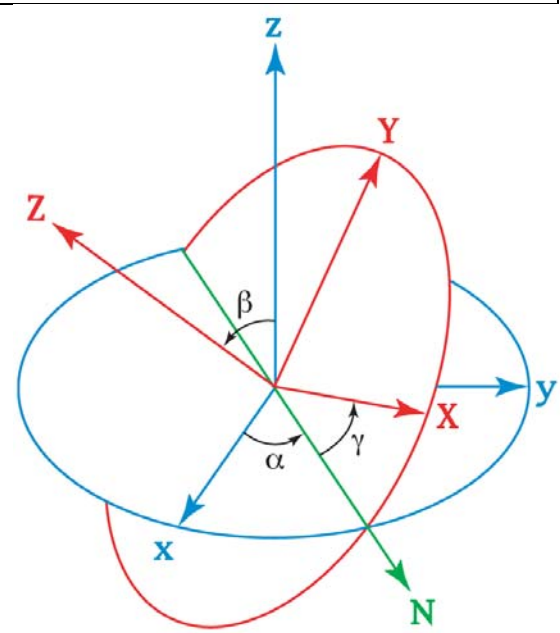
mechanics. Memorize every detail of the figures representing the rotations. The time spent will be well spent.

$$\lambda(\phi,\theta,\psi)=\begin{bmatrix}\cos\psi & \sin\psi & 0\\ -\sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix}$$

Euler's Prescription: Call initial direction of the *z*-axis $\hat{z}_{initial}$ and the final direction of the *z*-axis $\hat{z}_{final}$. The vector $\hat{z}_{initial}\times\hat{z}_{final}$ is perpendicular to both and so lies somewhere in the original *x*-*y* plane, perpendicular to $\hat{z}_{initial}$. The first rotation is CCW by ϕ about $\hat{z}_{initial}$ to rotate the *x* -axis $\hat{x}$ from its initial direction $\hat{x}_{initial}$ to the direction of $\hat{z}_{initial}\times\hat{z}_{final}=\hat{x}_{intermediate}$. This $\hat{x}_{intermediate}$ direction is perpendicular to the plane containing $\hat{z}_{initial}$ and $\hat{z}_{final}$ so a CCW rotation by θ can be made that sweeps the *z*-axis from $\hat{z}_{initial}$ to $\hat{z}_{final}$. As the *z*-axis is in its final orientation, all that remains is a CCW rotation by ψ to sweep the *x* and *y* axes from their intermediate to final orientations.

Conventional Order Alternative: Rotation about *z* then *y′* and finally *z″*. STAY ALERT.

The Euler angles: Consider a general rotation of a rigid body about an axis point. *Paint on* a set of Cartesian axes centered on that fixed point. Imagine a circle in the horizontal *x*-*y* plane that is concentric with the origin. It defines a plane and so will have a line of intersection with the horizontal plane after a general rotation. Begin by rotating the axes by α about the current *z* axis to swing the *x* axis into the line of intersection which is called the line of nodes. Rotate by β about that new *x* axis to swing the *z* axis into its final orientation. Finally, rotate by γ about the final *z* direction to swing the *x* and *y* axes into their final



http://upload.wikimedia.org/wikipedia/commons/8/82/Euler.png

orientations. Map: $\{\alpha,\beta,\gamma\} \rightarrow \{\phi,\theta,\psi\}$

Determinant: Note that the determinant of each matrix above is +1 and that $\lambda(\phi,\theta,\psi)$ is the most general rotation matrix. The determinant of any rotation matrix is 1.

Why would we ever use a rotation?

Consider an ellipse:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

The major and minor axes are parallel to the *x* and *y*-axes and the semi-major and semi-minor axis lengths are the larger and smaller of *a* and *b*. A more general ellipse has the representation:

$$A x^2 + 2 C x y + B y^2 = 1 \qquad \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} A & C \\ C & B \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=1$$

The axes of this ellipse are not parallel to the *x* and y axes, but rather they lie at some angle with respect to them. Transform to a rotated reference system rotated by θ about the *z*-axis using the prescription: $\tilde{\tilde{\alpha}}' = \lambda\ \tilde{\tilde{\alpha}}\ \ \lambda^{-1} = \lambda\ \tilde{\tilde{\alpha}}\ \ \lambda^{t}$

Note that the *z*-parts are suppressed to save ink.

$$\begin{bmatrix} x' & y' \end{bmatrix}\begin{bmatrix} A' & C' \\ C' & B' \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}=1 \quad \text{where} \quad \begin{bmatrix} A' & C' \\ C' & B' \end{bmatrix}=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} A & C \\ C & B \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\begin{bmatrix} A' & C' \\ C' & B' \end{bmatrix}=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} A\cos\theta + C\sin\theta & -A\sin\theta + C\cos\theta \\ C\cos\theta + B\sin\theta & -C\sin\theta + B\cos\theta \end{bmatrix}$$

$$\begin{bmatrix} A' & C' \\ C' & B' \end{bmatrix}=\begin{bmatrix} A\cos^2\theta + B\sin^2\theta + C\sin 2\theta & ({}^1/_2)(B-A)\sin 2\theta + C\cos 2\theta \\ ({}^1/_2)(B-A)\sin 2\theta + C\cos 2\theta & A\sin^2\theta + B\cos^2\theta - C\sin 2\theta \end{bmatrix}$$

The final form follows after the application of a few trigonometric identities.

The goal is to zero-out the off-diagonal terms by a choice of θ.

$$(\tfrac{1}{2})(B-A)\sin 2\theta + C\cos 2\theta = 0 \Rightarrow \tan(2\theta) = \frac{2C}{A-B}$$

That is it. The axes of the ellipse are rotated by θ relative to the original Cartesian axes. With the choice of angle above, the off-diagonal elements are zero and the diagonal elements are the eigenvalues of the matrix (*for our case, the squares of the semi-minor and semi-major axes*). Inverse tangent yields two angle separated by π. This multiplicity corresponds to choosing one pair of eigen-directions or another equivalent pair that follows a rotation by $^{\pi}/_{2}$. A proof follows that the three eigen-directions of a real symmetric (3D) matrix are mutually orthogonal.

German: *Eigenshaft* => nature, virtue, feature, quality, property, attribute character, characteristic, characteristic quality

Exercise: If tan(2θ) = 0, what is sin(2θ)? Use these values to find diagonal form of the matrix of coefficients.

Choosing $\tan(2\theta) = \dfrac{2C}{A-B}$, $\begin{bmatrix} A' & C' \\ C' & B' \end{bmatrix} = \begin{bmatrix} A\cos^2\theta + B\sin^2\theta + C\sin 2\theta & 0 \\ 0 & A\sin^2\theta + B\cos^2\theta - C\sin 2\theta \end{bmatrix}$.

The values on the diagonal are the eigenvalues of the matrix.

Application to moments of inertia:

Moments of inertia for a collection of point masses: (See the Introduction to Eigenvalues handout.)

$$I_{ij} = \sum_{masses\ \alpha} m_\alpha \left[r_\alpha^2\ \delta_{ij} - x_{\alpha i} x_{\alpha j} \right]$$

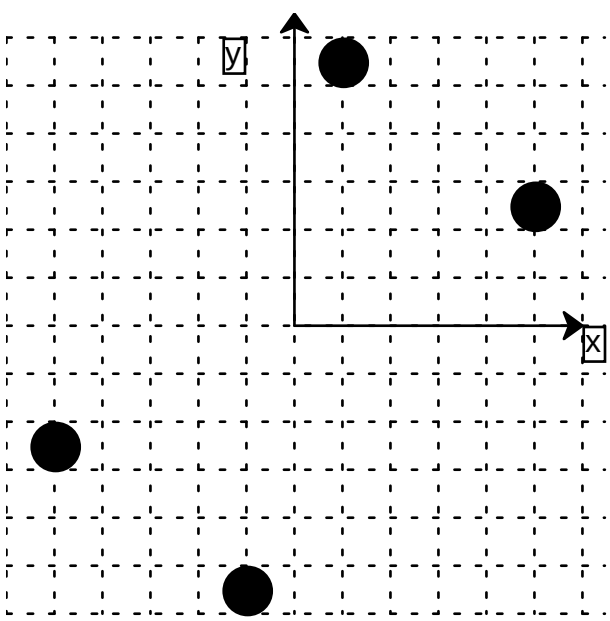
Using the mapping: $x_1 \leftrightarrow x; x_2 \leftrightarrow y; x_3 \leftrightarrow z$, for example x_{72} is the y coordinate of particle 7.

$$\mathbb{I} = \begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} = \sum_{masses\ \alpha} m_\alpha \begin{bmatrix} y_\alpha{}^2 + z_\alpha{}^2 & -x_\alpha\, y_\alpha & -x_\alpha\, z_\alpha \\ -y_\alpha\, x_\alpha & x_\alpha{}^2 + z_\alpha{}^2 & -y_\alpha\, z_\alpha \\ -z_\alpha\, x_\alpha & -z_\alpha\, y_\alpha & x_\alpha{}^2 + y_\alpha{}^2 \end{bmatrix}$$

The sum over the particles uses the particle label (index) α as illustrated. The diagonal elements of the inertia tensor are the familiar moments of inertia about the *x*, *y* and z-axes, the sum of each mass multiplied by its perpendicular distance from the axis squared. For example:

$$I_{xx} = \sum_\alpha m_\alpha\, r^2_{perpendicular} = \sum_\alpha m_\alpha \left(y_\alpha^2 + z_\alpha^2 \right)$$

The off-diagonal elements are called skew moments.



m_1: 1 kg @ (10m, 5m, 0m)

m_2: 1 kg @ (2m, 11m, 0m)

m_3: 1 kg @ (-10m, -5m, 0m)

m_4: 1 kg @ (-2m, -11m, 0m)

For this example,

$$\mathbb{I} = \begin{bmatrix} 292 & -144 & 0 \\ -144 & 208 & 0 \\ 0 & 0 & 500 \end{bmatrix} \quad \text{Desired} \quad \vec{L} = \begin{bmatrix} 292 & -144 & 0 \\ -144 & 208 & 0 \\ 0 & 0 & 500 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \lambda \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$$

As you will learn in your intermediate mechanics class, $\mathbb{I}\vec{\omega} = \vec{L}$, the angular momentum of the rigid assembly of masses rotating at $\vec{\omega}$. The assembly can only rotate smoothly (torque-free) if $\vec{L}$ is parallel to $\vec{\omega}$. That is: A classic eigenvalue problem is to be

solved. ($\mathbb{I}\vec{\omega}_i = \vec{L} = \lambda_i \vec{\omega}_i$) As a simplifying step, the equation is to be divided by 4 *kg-m*^{2}. As a result the values found are one-fourth the eigenvalues of the original problem.

$$\begin{bmatrix} 73 & -36 & 0 \\ -36 & 52 & 0 \\ 0 & 0 & 125 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mu \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad where\ \mu = \lambda/4$$

The results of the ellipse example are to be exercised before the classical solution method.

$$\tan(2\theta) = \frac{2C}{A-B} = \frac{2(-36)}{(73-52)} = -\frac{72}{21} = -\frac{24}{7}$$

This equation can be 'solved' to find the possible angles.

$$\tan(2\theta) = -\frac{72}{21} = \frac{\sin(2\theta)}{\cos(2\theta)} = \frac{2\sin(\theta)\cos(\theta)}{\cos^2(\theta)-\sin^2(\theta)} = \tau \qquad \text{(defines } \tau \text{ as } -{}^{72}/_{21} = -{}^{24}/_{7}\text{)}$$

$$2\sin(\theta)\cos(\theta) = \tau\left(\cos^2(\theta)-\sin^2(\theta)\right)$$

$$4\sin^2(\theta)\cos^2(\theta) = \tau^2\left(\cos^4(\theta)-2\sin^2(\theta)\cos^2(\theta)+\sin^4(\theta)\right) \qquad \text{(square)}$$

$$4\sin^2(\theta)\left[1-\sin^2(\theta)\right] = \tau^2\left(\left[1-\sin^2(\theta)\right]^2 - 2\sin^2(\theta)\left[1-\sin^2(\theta)\right]+\sin^4(\theta)\right)$$

$$4\sin^2(\theta)-4\sin^4(\theta) = \tau^2\left(1-4\sin^2(\theta)+4\sin^4(\theta)\right)$$

$$0 = 4\left(\tau^2+1\right)\sin^4(\theta)-4\left(\tau^2+1\right)\sin^2(\theta)+\tau^2$$

$$\sin^2(\theta) = \frac{4\left(\tau^2+1\right)\pm\sqrt{16\left(\tau^4+2\tau^2+1\right)-16\left(\tau^4+\tau^2\right)}}{8\left(\tau^2+1\right)}$$

$$\sin^2(\theta) = \frac{1}{2}\pm\frac{\sqrt{\left(\tau^2+1\right)}}{2\left(\tau^2+1\right)} = \frac{1}{2}\pm\frac{1}{2\sqrt{\left(\tau^2+1\right)}}$$

Substituting $\tau = -{}^{24}/_{7}$, $\sin^2(\theta) = \frac{1}{2}\pm\frac{21}{150} = \frac{1}{2}\pm\frac{7}{50} = \frac{16}{25}\ ;\ \frac{9}{25}$ or sinθ is to be chosen from the list $\{\pm\frac{4}{5};\pm\frac{3}{5}\}$. But wait, the equation was squared during the solution

process. Expect that half the values listed are bogus. Checking the condition $\tan(2\theta) = -\frac{24}{7}$, $\sin\theta = +\frac{4}{5}$ and $-\frac{3}{5}$ are found to be the valid pair of values. These values identify angles separated by $^{\pi}/_{2}$ which correspond to orthogonal eigen-directions.

The inertia tensor is diagonal in a reference frame rotated by an angle $\theta_1 = +\sin^{-1}(^4/_5)$ or

$\theta_2 = -\sin^{-1}(^3/_5)$ about the z-axis. The $\theta_2 = -\sin^{-1}(^3/_5)$ case is to be examined.

The eigenvalues for μ are the diagonal elements calculated for $\theta_2 = -\sin^{-1}(^3/_5)$.

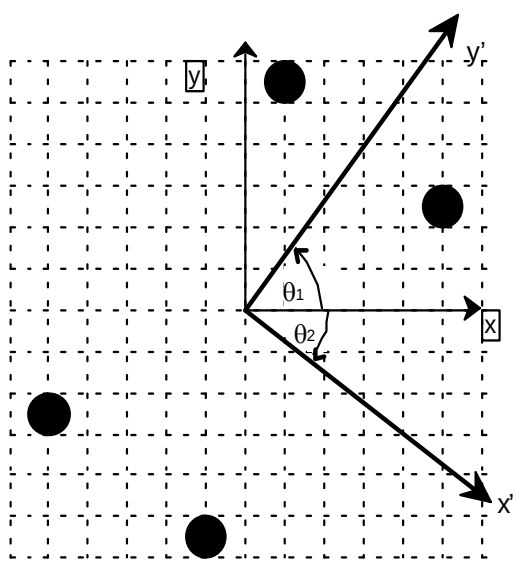
$$\mu_1 = A\cos^2\theta + B\sin^2\theta + C\sin 2\theta = 73(.64) + 52(.36) + (-36)(-.96) = 100$$

$$\mu_2 = A\sin^2\theta + B\cos^2\theta - C\sin 2\theta = 73(.36) + 52(.64) - (-36)(-.96) = 25$$

$$\mu_3 = 125$$

Armed with the values of μ, recall that the λ's are 4 *kg-m*^{2} times these values.

$\lambda = \{400, 100, 500\}$ *kg-m*^{2}.



The two angles are shown. They locate the two eigen-directions for the array of masses. Note that eigen-directions reflect the symmetry properties of the body. In the new coordinate system, the masses are 5 meters from one axis and ten meters from the other.

m_1: 1 kg @ (10m, 5m) →

(10 cosθ + 5 sinθ, -10 sinθ + 5

cosθ)

= (8 - 3, -10 (-.6)+ 5(.8)) = (5, 10)

Note: each square is 2 x 2.

Exercise: Use the transformation equations: (**remember** $\theta = -\sin^{-1}(^3/_5)$)

$$x' = x\cos\theta + y\sin\theta; \quad y' = x(-\sin\theta) + y\cos\theta; \quad z' = z$$

to compute the coordinates of the masses in the rotated frame. Compute the moment of inertia tensor as observed in the rotated frame using the transformed coordinates. Masses are neither vectors nor (rank two) tensors. How do the values of the masses of each object as determined by the original observer compare with the values found by the observer in the rotated frame? What designation is given to quantities with values that are invariant, the same as reported by observers in the primed and unprimed frames?

The classical solution of the matrix eigenvalue problem leads to the characteristic equation for the scaled eigenvalues, the μ's:

$$\left[(73-\mu)(52-\mu)-36^2\right]\left[125-\mu\right]=0$$

$$\left[\mu^2-125\mu+2500\right]\left[125-\mu\right]=0$$

$$\left[(100-\mu)(25-\mu)\right]\left[125-\mu\right]=0$$

Hence μ = {25, 100, 125} and λ = {100, 400, 500} *kg-m*^{2}.

The final task is to determine the directions of the eigenvectors.

$$\begin{bmatrix} (73-\mu) & -36 & 0 \\ -36 & (52-\mu) & 0 \\ 0 & 0 & (125-\mu) \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For μ= 100

$$\begin{bmatrix} (-27) & -36 & 0 \\ -36 & (-48) & 0 \\ 0 & 0 & (100) \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow\ 3a + 4b = 0 \ \ or\ \hat{e}_{\mu=100} = \frac{4\,\hat{i} - 3\,\hat{j}}{5}$$

For μ= 25

$$\begin{bmatrix} (48) & -36 & 0 \\ -36 & (27) & 0 \\ 0 & 0 & (100) \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow\ 4a - 3b = 0 \ \ or\ \hat{e}_{\mu=25} = \frac{3\,\hat{i} + 4\,\hat{j}}{5}$$

For μ= 125, the z-direction works.

This first direction corresponds to the *x' axis* about which the mass distribution has a moment of inertia of 4μ *kg-m*^{2} or 400 *kg-m*^{2}. What angle does this axis make relative to the original *x*-axis?

The collection of eigen-directions: {($^4/_5$, -$^3/_5$.0) ; ($^3/_5$, $^4/_5$.0); (0, 0,1) }

Observations: A simple rotation diagonalized the matrix of coefficients for the equation of an ellipse. The procedure also worked for a simple moment of inertia matrix. Both initial matrices were real symmetric.

The unsupported conclusion is that a real symmetric matrix can be diagonalized by an orthogonal transformation of which rotations are the most interesting. To diagonalize a general 3 x 3 real symmetric matrix, an arbitrary rotation involving the three independent Euler angles is needed.

Orthogonal transformations include rotations and inversions. One need only multiply the collection of all rotations by [(*x,y,z*)->(-*x,y,z*)] reflection across the *y-z* plane or [(*x,y,z*)->(-*x*,-*y*-,*z*)] inversion through the origin to fill out all

possible orthogonal transformations in 3D. Rotations do not distort physical objects as they are transformed; inversions begin and end in configurations in which all separations are the same, but no continuous transformation can be imagined that preserves separations throughout the transformation. The set of all 3 by 3 orthogonal matrices forms the group O(3). If the inversions are excluded, the remaining transformations form the *special* subgroup SO(3) representing rotations in three dimensions. The matrices representing transformations of the O(3) type have a determinant of ±1 while all those for SO(3) have determinant +1.

Exercise: Consider the eigenvalue problem above.

$$\begin{bmatrix} (73-\mu) & -36 & 0 \\ -36 & (52-\mu) & 0 \\ 0 & 0 & (125-\mu) \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{vmatrix} (73-\mu) & -36 & 0 \\ -36 & (52-\mu) & 0 \\ 0 & 0 & (125-\mu) \end{vmatrix} = 0$$

Show that the determinant condition leads to the equation $(125 - \mu)(2500 - 125\,\mu + \mu^2)$ which has roots {25, 100, 125}.

The physical importance is that all (rank two) tensor properties that are naturally real symmetric appear as diagonal matrices in some rotated frame. Consider the inertia tensor $\mathbb{I}$ for a football. Common sense and symmetry principles suggest that one axis be along the long axis of the ball and that the other two axes be any mutually perpendicular directions perpendicular to that axis.

Examples of physical (rank two) tensors: Quadrupole moments: $\mathbb{Q}$, Moment of inertia: $\mathbb{I}$, Polarizability: $\mathbb{\alpha}$. Note that these matrices represent physical properties of physical entities, and that they are real symmetric.

The usual case: Note that matrices representing physical properties of physical entities (at rest) are real symmetric ($\mathbb{M} = \mathbb{M}^t$). The matrices λ that represent rotations are to transform from one coordinate system (not a part of the physics) to another coordinate system (not part of the physics). These matrices are orthogonal ($\mathbb{O}^{-1} = \mathbb{O}^t$) rather than real symmetric. Other physical concepts are represented by anti-symmetric matrices ($\mathbb{A} = -\mathbb{A}^t$). I am unable to make a clear statement, but these anti-symmetric matrices are tied to our attempt to involve motion and position at the same time and more specifically, they involve a cross product or curl with the arbitrary sign convention, the right-hand rule.

What can be learned from rotation matrices?:

Suppose that an inertia tensor has been diagonalized,

$$\mathbb{I} = \begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix}$$

and that the body has a four-fold rotation symmetry about the z-axis. The last statement means, when the body is rotated by 90^0 about the z-axis, an identical mass distribution results. So: $\mathbb{I} \equiv \mathbb{I}^{/} = \lambda\, \mathbb{I}\, \lambda^{\mathbf{t}}$.

$$\lambda = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -I_x & 0 \\ I_y & 0 & 0 \\ 0 & 0 & I_z \end{bmatrix} = \begin{bmatrix} I_y & 0 & 0 \\ 0 & I_x & 0 \\ 0 & 0 & I_z \end{bmatrix}$$

Not surprisingly, the 90^0 rotation interchanges the *x* and the *y* elements. As the body is invariant under the rotation, the inertia tensor must be identical: $I_x = I_y$. The conclusion is that, when diagonalized, the *x* and *y* diagonal elements are equal for a body with four-fold rotation symmetry about the *z*-axis.

Suppose the body only has three-fold rotation symmetry about the *z*-axis. That is a rotation be 120^0 returns the body to an identical configuration.

$$\begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} I_x + \frac{3}{4} I_y & 0 & 0 \\ 0 & \frac{3}{4} I_x + \frac{1}{4} I_y & 0 \\ 0 & 0 & I_z \end{bmatrix}$$

Again the result requires that $I_x = I_y$. In fact, the result holds for the z-axis being an *n-fold* rotational symmetry axis for any integer $n > 2$.

How does one find λ to diagonalize a general real symmetric matrix?:

A prescription to build λ can be developed from results that have been established. Consider the generic representation of the eigenvectors and eigenvalue problem for a real symmetric matrix where the left-hand equation is in the original coordinate representation while the right-hand is the form after rotating the reference frame to yield the diagonal form of $\mathbb{M}$.

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} = \mu_i \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} \quad where\ i = 1,2,3 \quad \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{bmatrix} \begin{bmatrix} a_i^{/} \\ b_i^{/} \\ c_i^{/} \end{bmatrix} = \mu_i \begin{bmatrix} a_i^{/} \\ b_i^{/} \\ c_i^{/} \end{bmatrix}$$

form before rotation form in preferred frame

The preferred frame has its axes along the eigen-directions. **Watch the Notation**: The symbol μ represents an eigenvalue; he symbol λ already represents the rotation matrix with elements λ_{ij}.

Just as the original observer represents *x* direction as $\hat{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, the rotated observer would could to represent the *one* axis as $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. The axis is in the physical direction of the first eigenvector.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \tilde{e}_1' = \lambda\ \tilde{e}_1 = = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} e_{1x} \\ e_{1y} \\ e_{1z} \end{bmatrix}$$

Hence, after rotation, the *normalized* eigenvectors in (*a, b, c*) form must transform into the canonical forms {{1, 0, 0}, … , …}:

$$\begin{bmatrix} a_1^{/} \\ b_1^{/} \\ c_1^{/} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}; \qquad \begin{bmatrix} a_2^{/} \\ b_2^{/} \\ c_2^{/} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad \begin{bmatrix} a_3^{/} \\ b_3^{/} \\ c_3^{/} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The eigenvectors must be listed in right-hand rule order. If not, the transformation is a rotation plus an inversion.

Exercise: What directions are to be chosen as the natural coordinate directions in the eigen-frame? Give the coordinate triplets for those coordinate directions as observed in the eigen-frame.

Recalling that the inverse of λ is its transpose:

$$\begin{bmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} \\ \lambda_{12} & \lambda_{22} & \lambda_{32} \\ \lambda_{13} & \lambda_{23} & \lambda_{33} \end{bmatrix} \begin{bmatrix} a_i^{/} \\ b_i^{/} \\ c_i^{/} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{21} & \lambda_{31} \\ \lambda_{12} & \lambda_{22} & \lambda_{32} \\ \lambda_{13} & \lambda_{23} & \lambda_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \text{ or } \lambda_{11} = a_1;\ \lambda_{12} = b_1;\ \lambda_{13} = c_1$$

Each row of the rotation matrix is the components of the corresponding eigen-direction.

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

Recall: Our conventions for 3D require that the coordinate vector triad be listed in *right-hand* order. As the directions are orthogonal, this convention requires at most that the sign of the third eigenvector be changed. **If the vectors are not listed in *right-hand* order, the transformation represented is a rotation plus an inversion.** Inversions are discussed later.

For the moment of inertia problem above, the final *x-y-z* sequence of eigenvalues as computed in the rotated frame is: 400, 100, 500 *kg-m*^{2}. Listing the eigenvectors as rows in the corresponding sequence:

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} = \begin{bmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Applying these values to the reduced problem:

$$\lambda\, \mathbb{I}\, \lambda^{\mathbf{t}} = \begin{bmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 73 & -36 & 0 \\ -36 & 52 & 0 \\ 0 & 0 & 125 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 & 0 \\ -3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 80 & 15 & 0 \\ -60 & 20 & 0 \\ 0 & 0 & 125 \end{bmatrix} = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Fantastic, it works in a simple case. More difficult examples are reserved for the homework problem section.

Exercise: Verify that the eigenvectors are list in right-hand order in the example above. Interchange the eigenvector rows to form any other order. Check that the order is right-handed. If not multiply row three by negative one. Verify that the new eigenvector set is right-handed. Apply the new rotation to diagonalize the original moment of inertia. Note that the result is still diagonal. What has changed?

Second Sample Calculation of λ from the Eigenvectors:

The matrix $\mathbb{A} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ has eigenvalues $\{1, 3, 3\}$.

a.) Pick mutually orthogonal eigenvectors for these eigenvalues and use them to develop an orthogonal transformation matrix λ that diagonalizes $\mathbb{A}$. Verify that your eigenvector sequence is in right-hand order. If not, multiply the third vector by negative one.

b.) Apply the resulting similarity transformation to demonstrate that $\mathbb{A}$ is transformed to a diagonal form with the eigenvalues as the diagonal elements. $\mathbb{A}^{/} = \lambda\, \mathbb{A}\, \lambda^{-1}$; $\mathbb{A}^{/}$ diagonal.

Work with the distinct eigenvalue first. $\lambda = 1$.

$$\begin{bmatrix} 2-\lambda & 0 & 1 \\ 0 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 2-1 & 0 & 1 \\ 0 & 3-1 & 0 \\ 1 & 0 & 2-1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$a + c = 0 \quad \text{and } 2\,b = 0 \;\Rightarrow\; c = -\,a \;\text{ and } b = 0. \qquad \hat{e}_{\lambda 1} = \frac{\hat{i}-\hat{k}}{\sqrt{2}}$$

Work with $\lambda = 3$ to find the condition satisfied by the components of the other two eigen-directions.

$$\begin{bmatrix} 2-3 & 0 & 1 \\ 0 & 3-3 & 0 \\ 1 & 0 & 2-3 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \;\Rightarrow -\,a + c = 0 \;\text{ and } 0\,b = 0.\;\; c = a \;\text{ and } b \;\text{ unrestricted.}$$

Choose and values that satisfy the requirements. Example: $c = a$ and $b = 0$.

$\hat{e}_{\lambda 3A} = \frac{\hat{i}+\hat{k}}{\sqrt{2}}$.

The third direction must be orthogonal, and it must satisfy the RHR so use the cross product.

$$\hat{e}_{\lambda 3B} = \hat{e}_{\lambda 1} \times \hat{e}_{\lambda 3A} = \left(\frac{\hat{i}-\hat{k}}{\sqrt{2}}\right) \times \left(\frac{\hat{i}+\hat{k}}{\sqrt{2}}\right) = \frac{-\hat{j}-\hat{j}}{2} = -\hat{j}\,.$$

As rows: $\hat{e}_{\lambda 1} = \frac{\hat{i}-\hat{k}}{\sqrt{2}} \rightarrow \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$; $\hat{e}_{\lambda 3A} = \frac{\hat{i}+\hat{k}}{\sqrt{2}} \rightarrow \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$; $\hat{e}_{\lambda 3B} = -\hat{j} = (0,-1,0))$

$$\lambda = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1 & 0 \end{bmatrix}$$

$$\lambda\, \mathbb{A}\, \lambda^{-1} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1 & 0 \end{bmatrix} \left(\begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & -1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \right)$$

$$\lambda\,\mathbb{A}\,\lambda^{-1} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 3/\sqrt{2} & 0 \\ 0 & 0 & -3 \\ -1/\sqrt{2} & 3/\sqrt{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \underset{\approx}{A}' \quad \text{(check)}$$

Eigenvectors for distinct eigenvalues are orthogonal.

Orthogonal transformations preserve the original Euclidean inner product for a real space.

$$\begin{bmatrix} a_i & b_i & c_i \end{bmatrix} \begin{bmatrix} a_j \\ b_j \\ c_j \end{bmatrix} = a_i\, a_j + b_i\, b_j + c_i\, c_j = \tilde{v}_i^t\, \tilde{v}_j \quad \text{where} \quad \tilde{v}_j = \begin{bmatrix} a_j \\ b_j \\ c_j \end{bmatrix} \quad \textit{and} \quad \tilde{v}_i^t = \begin{bmatrix} a_i & b_i & c_i \end{bmatrix}$$

Orthogonal transformations preserve all inner products so that this standard method can be used before and after a transformation.

The eigenvalue problem for real symmetric matrices (and its transpose):

$$\mathbb{M}\ \tilde{v}_j = \lambda_j\, \tilde{v}_j \qquad (\mathbb{M}\ \tilde{v}_j)^t = (\lambda_j\, \tilde{v}_j)^t \rightarrow \tilde{v}_j^t\ \mathbb{M}^{\mathrm{t}} = \lambda_j\ \tilde{v}_j^t \text{ and finally } \tilde{v}_j^t\ \mathbb{M} = \lambda_j\ \tilde{v}_j^t$$

The last form follows because real symmetric matrices are their own transposes.

NOTE: The real-symmetric nature is required to complete this proof.

eigenvalue equation for kets: $\mathbb{M}\ \tilde{v}_j = \lambda_j\, \tilde{v}_j$

eigenvalue equation for bras: $\tilde{v}_j^t\ \mathbb{M} = \lambda_j\ \tilde{v}_j^t$

Theorem: The eigenvectors of a real symmetric matrix that correspond to distinct eigenvalues are orthogonal.

$$(\tilde{v}_i^t\, \mathbb{M}\ )\ \tilde{v}_j\ =\ \tilde{v}_i^t\ (\mathbb{M}\ \tilde{v}_j) \qquad \text{associative property}$$

$$(\lambda_i\, \tilde{v}_i^t)\ \tilde{v}_j = \tilde{v}_i^t\, (\lambda_j\, \tilde{v}_j) \qquad \text{eigenvalue property}$$

$$(\lambda_i\ - \lambda_j)\ \tilde{v}_i^t\ \tilde{v}_j = 0 \qquad \text{the } \lambda\text{'s are just numbers}$$

The conclusion is that if the eigenvalues λ_i and λ_j are distinct (not equal) then their corresponding eigenvectors must be orthogonal ($\tilde{v}_i^t \; \tilde{v}_j = 0$).

It is common that real symmetric matrices represent concepts (material properties) in classical physics. The eigenvectors of these matrices are orthogonal if they are associated with distinct eigenvalues. The eigenvectors that correspond to a degenerate eigenvalue may be chosen to be orthogonal, and processes such as Gram-Schmidt can produce a suitable orthogonal set from an initial set.

The Cartesian form of the inner product is our usual starting point. Orthogonal transformations preserve inner products so the standard definition as the sum of the products of the corresponding components remains valid. $\Rightarrow$ The eigenvectors can be unity normalized.

$$\begin{bmatrix} a_i & b_i & c_i \end{bmatrix} \begin{bmatrix} a_j \\ b_j \\ c_j \end{bmatrix} = a_i\, a_j + b_i\, b_j + c_i\, c_j = \tilde{v}_i^t\, \tilde{v}_j = \delta_{i\,j}$$

A symmetric matrix can be diagonalized by an orthogonal transformation. The form of the transformation matrix is the collection of the eigenvectors as the rows of the matrix. The orthonormal form of the eigenvectors ensures that the matrix is orthogonal.

Other transformations proper and otherwise:

Rotations are but one class of orthogonal transformations. As demonstrated, the determinant of the coefficients representing a rotation is one. Orthogonal transformations with determinant one are designated as proper transformations. Orthogonal transformations are severely restricted as they must preserve all inner

products, but a reflection across a coordinate plane does qualify. Consider the mapping by a single reflection across the *y*-*z* coordinate plane:

$$x' = -x; \quad y' = y; \quad z' = z$$

Vector inner products are preserved as:

$$\tilde{a}'_x \tilde{b}'_x + \tilde{a}'_y \tilde{b}'_y + \tilde{a}'_z \tilde{b}'_z = (-\tilde{a}_x)(-\tilde{b}_x) + \tilde{a}_y \tilde{b}_y + \tilde{a}_z \tilde{b}_z = \tilde{a}_x \tilde{b}_x + \tilde{a}_y \tilde{b}_y + \tilde{a}_z \tilde{b}_z$$

The determinant of this orthogonal transformation is negative one, and it is an *improper* transformation. If this transformation is followed by a reflection across the *x*-*z* plane, the net mapping is:

$$x' = -x; \quad y' = -y; \quad z' = z$$

and the determinant is one, and therefore is a proper transformation. This result is not surprising as the particular sequence of two reflections is identical to a rotation of π about the *z* axis. Viewed otherwise, the product of two transformations that each have determinant negative one is a transformation with determinant one. Hence the product of two improper transformations is a proper one, a rotation. Continuing, one has the inversion mapping

$$x' = -x; \quad y' = -y; \quad z' = -z$$

that may be achieved as a sequence of three reflections across the three coordinate planes. It is improper as it has determinant $(-1)^3$.

A reflection across any plane through the origin is the combination of a rotation to orient that plane as a coordinate plane, the reflection and a rotation back to the original plane. The determinant of the entire operation is (+1)(-1)(+1)=-1. The mapping is improper as are all that contain an odd number of reflections.

Orthogonal transformations can represent gentle passive changes of reference frame. Improper transformation seem to be a little ruder, but at least distances, shapes, angles, …. are preserved although inversion itself is painful. Physical vectors behave well

under all orthogonal transformations. The class of pretenders, pseudo-vectors, co-varies with vectors under proper transformations but fails under improper ones. The components of a vector change sign under an inversion, but those of pseudo-vectors do not. Pseudo-vectors include cross products and curls with common examples being torque, magnetic field, angular momentum, and angular velocity. Although pseudo-vectors are not quite vectors, it is permissible to pretend that they are as long as one restricts reference frame changes to proper ones such as rotations. Do not ponder the preceding paragraph; just steel yourself against the day in graduate school when the cross product is declared to be an anti-symmetric tensor rather than a vector. In three dimensions, a tensor has nine components. If it is anti-symmetric, the diagonal entries must be zero leaving three independent off-diagonal values, the three components of the pseudo-vector. Above all remember that for us, it just does not matter.

More general transformations fail to preserve inner products. Recall that preserving inner products leaves distances and angles (size and shapes) unchanged. An isotropic scaling changes the inner product. Suppose the basis vectors $\left\{2\hat{i}, 2\hat{j}, 2\hat{k}\right\}$ are adopted. Then the new vector components relative to this basis are the original values divided by 2.

$$\tilde{a}_x^{/}\left(2\hat{i}\right)+\tilde{a}_y^{/}\left(2\hat{j}\right)+\tilde{a}_z^{/}\left(2\hat{k}\right)=\tfrac{1}{2}\tilde{a}_x\left(2\hat{i}\right)+\tfrac{1}{2}\tilde{a}_y\left(2\hat{j}\right)+\tfrac{1}{2}\tilde{a}_z\left(2\hat{k}\right)=\tilde{a}_x\,\hat{i}+\tilde{a}_y\,\hat{j}+\tilde{a}_z\,\hat{k}$$

The inner product changes under a scaling!

$$\begin{bmatrix}\tilde{a}_x & \tilde{a}_y & \tilde{a}_z\end{bmatrix}\begin{bmatrix}\tilde{b}_x\\ \tilde{b}_y\\ \tilde{b}_z\end{bmatrix}=\begin{bmatrix}\tilde{a}_x & \tilde{a}_y & \tilde{a}_z\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\tilde{b}_x\\ \tilde{b}_y\\ \tilde{b}_z\end{bmatrix}\rightarrow\begin{bmatrix}\tilde{a}_x^{/} & \tilde{a}_y^{/} & \tilde{a}_z^{/}\end{bmatrix}\begin{bmatrix}2^2 & 0 & 0\\ 0 & 2^2 & 0\\ 0 & 0 & 2^2\end{bmatrix}\begin{bmatrix}\tilde{b}_x^{/}\\ \tilde{b}_y^{/}\\ \tilde{b}_z^{/}\end{bmatrix}$$

The inner product takes the form

$$\langle a|b\rangle=\sum_{i,j=1}^{3}\tilde{a}_i^{/}\,g_{ij}\,\tilde{b}_j^{/}$$

where g_{ij} is the metric. The metric was ignored in Cartesian problems because it is just the identity. It has been lurking around other places such as in spherical coordinates.

$$|d\vec{r}|^2 = \begin{bmatrix} dr & d\theta & d\phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & (r\sin\theta)^2 \end{bmatrix} \begin{bmatrix} dr \\ d\theta \\ d\phi \end{bmatrix}$$

Projection is a useful mapping which alters inner product. Projection collapses the space in some manner rather than just changing the way we view as in the case of a reference frame rotation. The mapping for projection into the *x*-*y* plane is:

$x' = x; \quad y' = y; \quad z' = NULL$

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Projections are operations that yield the same result if repeated. $P\left[P\left[|v\rangle\right]\right] = P\left[|v\rangle\right]$

Our example, in a sense, maps (*x*,*y*,*z*) to (*x*,*y*,*0*). If the projection is repeated, clearly (*x*,*y*,0) is mapped to (*x*,*y*,0) so $P\left[P\left[|(x,y,z)\rangle\right]\right] = P\left[|(x,y,0)\rangle\right] = (x,y,0)$.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The operator identity is $P[P[\ \ldots\ ]] = P[\ \ldots\ ]$ or $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

In prose, the projection finds and the part parallel to the $z = 0$ plane and removes that part perpendicular to that plane. Projecting a second time yields the same answer as the previous projection removed all non-parallel contributions.

Angles and more can be distorted by a shear.

$$x' = x + y; \quad y' = y; \quad z' = z$$

Exercise: Provide the matrix representation of the shear mapping above. Find the angle between the mappings of the *x* and *y* directions. Compute the magnitude of the mappings of the *x* and *y* directions.

We love rotations. We detest transformations that fail to preserve inner products.

JUMP TO PAGE 45.

The angular velocity as a vector:

Rotations failed to qualify as vectors early in the game as the net action of rotations is different if they are applied in a different order. Rotations lack a commutative addition.

A GENERAL PRINCIPLE OF EXPANSIONS is that if an expansion retains any second (n^{th}) order terms, it must retain all terms of that order if valid conclusions are to be drawn. The treatment below ignores the second order terms in the expansion of the cosines. The flavor of the conclusion is nonetheless correct. This section is not to be corrected before next year. More on this point in the Perturbations Handout that has not been written.

The development below succeeds because the terms of interest must involve interaction between the two rotations. That means that all the terms of interest are products that are first order in the first rotation and first order in the second rotation. As a consequence one finds all the terms of interest with only first order expansions for the actions of the individual rotations.

The expansion to second order in the cross product form is:

$$\delta \vec{r}_1 = \delta\vec{\theta}_1 \times \vec{r} + \left(\tfrac{1}{2}\right) \delta\vec{\theta}_1 \times \left(\delta\vec{\theta}_1 \times \vec{r}\right)$$

If one compares the action of infinitesimal rotation sequences, the crux of the issue becomes apparent. For small angles $\cos\Delta\theta \to 1$ and $\sin\Delta\theta \to \Delta\theta$. A small rotation

about the three axis followed by a small rotation about the one axis has the matrix representation:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \Delta\theta_1 \\ 0 & -\Delta\theta_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \Delta\theta_3 & 0 \\ -\Delta\theta_3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \Delta\theta_3 & 0 \\ -\Delta\theta_3 & 1 & \Delta\theta_1 \\ \Delta\theta_1\,\Delta\theta_3 & -\Delta\theta_1 & 1 \end{bmatrix}$$

so the change is:

$$\begin{bmatrix} 1 & \Delta\theta_3 & 0 \\ -\Delta\theta_3 & 1 & \Delta\theta_1 \\ \Delta\theta_1\,\Delta\theta_3 & -\Delta\theta_1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \Delta\theta_3 & 0 \\ -\Delta\theta_3 & 0 & \Delta\theta_1 \\ \Delta\theta_1\,\Delta\theta_3 & -\Delta\theta_1 & 0 \end{bmatrix}$$

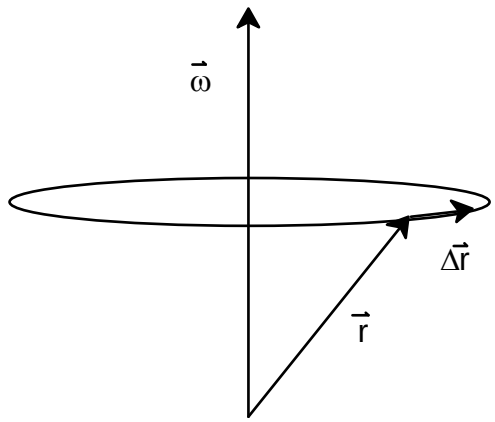
Reversing the order of the rotations:

$$\begin{bmatrix} 1 & \Delta\theta_3 & 0 \\ -\Delta\theta_3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \Delta\theta_1 \\ 0 & -\Delta\theta_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \Delta\theta_3 & \Delta\theta_1\,\Delta\theta_3 \\ -\Delta\theta_3 & 1 & \Delta\theta_1 \\ 0 & -\Delta\theta_1 & 1 \end{bmatrix}$$

so the change is:

$$\begin{bmatrix} 1 & \Delta\theta_3 & \Delta\theta_1\,\Delta\theta_3 \\ -\Delta\theta_3 & 1 & \Delta\theta_1 \\ 0 & -\Delta\theta_1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \Delta\theta_3 & \Delta\theta_1\,\Delta\theta_3 \\ -\Delta\theta_3 & 0 & \Delta\theta_1 \\ 0 & -\Delta\theta_1 & 0 \end{bmatrix}$$

The results differ in second order in the small angles. If the second order pieces were negligible, the actions would commute. Angular velocities afford just this release. Represent $\Delta\theta_i$ *as* $\omega_i\ \Delta t$ and divide by Δt. In the limit $\Delta t \to 0$, the offending second order terms disappear. Angular velocities and infinitesimal rotations can be considered as vectors under proper transformations of reference frames.

	This result is so important that it is to be repeated. Consider rotating a body with $\vec{r}(t)$ representing the position vector of a point on the body. The body rotates about the origin at $\vec{\omega}$ for a short time Δt. The new position of the point is: $\vec{r}(t+\Delta t)=\vec{r}(t)+\vec{\omega}\Delta t\times\vec{r}(t)$. (NOTE: This transformation is *active*.)

Now consider the change due to two rotation steps $\delta\vec{\theta}_1$ and $\delta\vec{\theta}_2$. (***Note that the notation chosen begs the 'Is it a vector ?' question. The vector-like notation does not qualify the rotations as vectors. They must satisfy the properties of vector addition and scalar multiplication.***

$$\delta\vec{r}_1=\delta\vec{\theta}_1\times\vec{r}(t)$$

$$\delta\vec{r}_2=\delta\vec{\theta}_2\times\vec{r}(t)+\delta\vec{\theta}_2\times\left[\delta\vec{\theta}_1\times\vec{r}(t)\right]$$

or: $$\delta\vec{r}=\delta\vec{\theta}_1\times\vec{r}(t)+\delta\vec{\theta}_2\times\vec{r}(t)+\delta\vec{\theta}_2\times\left[\delta\vec{\theta}_1\times\vec{r}(t)\right]$$

If the rotations are applied in the reverse order, one would find:

$$\delta\vec{r}=\delta\vec{\theta}_1\times\vec{r}(t)+\delta\vec{\theta}_2\times\vec{r}(t)+\delta\vec{\theta}_1\times\left[\delta\vec{\theta}_2\times\vec{r}(t)\right]$$

Look at the velocities:

$$\vec{v}=\frac{\delta\vec{r}}{\delta t}=\frac{1}{\delta t}\left[\delta\vec{\theta}_1\times\vec{r}(t)+\delta\vec{\theta}_2\times\vec{r}(t)+\delta\vec{\theta}_1\times\left[\delta\vec{\theta}_2\times\vec{r}(t)\right]\right] \text{ or}$$

$$\vec{v}=\frac{\delta\vec{r}}{\delta t}=\frac{1}{\delta t}\left[\delta\vec{\theta}_1\times\vec{r}(t)+\delta\vec{\theta}_2\times\vec{r}(t)+\delta\vec{\theta}_2\times\left[\delta\vec{\theta}_1\times\vec{r}(t)\right]\right]$$

In the limit that δt goes to zero, the δθs go to zero, but the ratios $\delta\vec{\theta}_i/\delta t\rightarrow\vec{\omega}_i$, finite angular velocities.

$$\vec{v}=\frac{\delta\vec{r}}{\delta t}=\left[\vec{\omega}_1\times\vec{r}(t)+\vec{\omega}_2\times\vec{r}(t)+\vec{\omega}_1\times\left[\delta\vec{\theta}_2\times\vec{r}(t)\right]\right]\rightarrow\left(\vec{\omega}_1+\vec{\omega}_2\right)\times\vec{r}(t)$$

$$\vec{v} = \frac{\delta \vec{r}}{\delta t} = \left[\vec{\omega}_2 \times \vec{r}(t) + \vec{\omega}_1 \times \vec{r}(t) + \vec{\omega}_2 \times \left[\delta\vec{\theta}_1 \times \vec{r}(t) \right] \right] \rightarrow \left(\vec{\omega}_2 + \vec{\omega}_1 \right) \times \vec{r}(t)$$

The angular velocities obey a commutative addition. Angular velocities (and infinitesimal rotations) behave like vectors in most cases of interest. Finite rotations do not. One application is the precessing bicycle wheel. The total angular velocity is the vector is represented as the vector sum of the rapid rotation about the axis of the wheel $\omega_{wheel}\ (a\hat{x}is)$ plus the vertical component of rotational velocity $\Omega_{prec}\ \hat{k}$.

***** Add a section defining ω and demonstrating that its members transform like those of a vector under proper orthogonal transformations, but not under the improper ones.

Matrix Representations of Linear Operators

Great effort has been taken to be vague about whatever it is that is being discussed. Mathematics abstracts common behaviors and rules from systems and hides all the particulars that are not essential to express those rules. A fantastic game ! At crunch time, a concrete representation is necessary, and, for things linear, matrices are a common choice. Matrices are to be proposed to represent vectors and linear mappings. The net effect of applying one mapping and then applying a second mapping to the result of the first is to be represented by multiplying by the matrix representation of the each mapping successively on the left. Settle in. The next few paragraphs are to meander over a range of topics - some of them advanced. Think about what is said, but take solace from the promise that the discussion is to end soon and is probably not critical to your immediate success.

The general topic is linear mappings of a vector space $\mathbb{V}$ into a vector space $\mathbb{W}$. Assume that a basis set $\{|e_1^v\rangle, |e_2^v\rangle, \ldots, |e_n^v\rangle\}$ has been chosen for $\mathbb{V}$ and that $\{|e_1^w\rangle, |e_2^w\rangle, \ldots, |e_m^w\rangle\}$ has been chosen for $\mathbb{W}$. A vector in either space is to be represented by the column vector of its coefficients in the chosen basis.

$$|v_i\rangle = a_{i1}|e_1^v\rangle + a_{i2}|e_2^v\rangle + a_{in}|e_n^v\rangle \quad \text{and} \quad |w_j\rangle = c_{j1}|e_1^w\rangle + c_{j2}|e_2^w\rangle + c_{jm}|e_m^w\rangle$$

Hence:

$$|v_i\rangle \to \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix} \quad \text{and} \quad |w_j\rangle \to \begin{bmatrix} c_{j1} \\ c_{j2} \\ \vdots \\ c_{jm} \end{bmatrix}$$

Next consider a linear mapping $\mathcal{L}$ of a vector space $\mathbb{V}$ into a vector space $\mathbb{W}$. The mapping is fully characterized by the collection of its actions on each basis vector for $\mathbb{V}$. The mapping of each $|e_i^v\rangle$ is $\mathcal{L}\,[|e_i^v\rangle]$, an element of $\mathbb{W}$ that therefore is a linear combination of the basis vectors for $\mathbb{W}$.

$$\tau_{1i}^{std.}|e_1^w\rangle + \tau_{2i}^{std.}|e_2^w\rangle + \ldots + \tau_{mi}^{std.}|e_m^w\rangle = \mathcal{L}\,[|e_i^v\rangle]$$

(The symbol τ is to be used for the matrix representation of the mapping $\mathcal{L}$.) The superscript *std.* specifies that the values $\tau_{j\,i}^{std.}$ are for the representation in the *standard* basis sets. This notation is usually omitted, and use of the standard basis sets is assumed. When the mapping acts on a general member of $\mathbb{V}$,

$$|V\rangle = a_1^V|e_1^v\rangle + a_2^V|e_2^v\rangle + \ldots + a_n^V|e_n^v\rangle$$

it is mapped to $\mathcal{L}\,[|V\rangle] = b_1^{L[V]}|e_1^w\rangle + b_2^{L[V]}|e_2^w\rangle + \ldots + b_m^{L[V]}|e_m^w\rangle$ where

$$\begin{bmatrix} c_1^{L[V]} \\ c_2^{L[V]} \\ \ldots \\ c_m^{L[V]} \end{bmatrix} = \begin{bmatrix} \tau_{11}^{std} & \tau_{12}^{std} & \ldots & \tau_{1n}^{std} \\ \tau_{21}^{std} & \tau_{22}^{std} & \ldots & \tau_{2n}^{std} \\ \ldots & \ldots & \ldots & \ldots \\ \tau_{m1}^{std} & \tau_{m2}^{std} & \ldots & \tau_{mn}^{std} \end{bmatrix} \begin{bmatrix} a_1^V \\ a_2^V \\ \ldots \\ a_n^V \end{bmatrix} \qquad \tilde{c} = \tilde{\tau}\,\tilde{a}\,.$$

The action of the linear mapping is represented by a matrix. Each column of the mapping matrix is the coefficients for the mapping of the corresponding basis vector. The symbols $\tilde{c}$ and $\tilde{a}$ represent the components of the vectors and $\tilde{\tilde{\tau}}$ represents the linear mapping. A change of basis set for either space requires re-computing $\tilde{a}$ or $\tilde{c}$ and $\tilde{\tilde{\tau}}$. In terms of the new basis sets, the vectors and the linear mapping are represented by arrays of values $\tilde{a}'$ or $\tilde{c}'$ and $\tilde{\tilde{\tau}}'$.

Exercise: Assume that both basis sets have elements that are mutually orthogonal and that are unity normalized. ($\langle e_i | e_j \rangle = \delta_{ij}$) Show that $\tau_{mn} = \langle e_m^w | \{ \hat{\tau} | e_n^v \rangle \}$ where $\mathcal{L} \, | e_i^v \rangle$ = $\hat{\tau} \, | e_i^v \rangle$. $\tau_{mn} = \langle e_m^w | \{ \hat{\tau} | e_n^v \rangle \} = \langle e_m^w | \hat{\tau} \, | e_n^v \rangle$ is one element of the matrix representing $\hat{\tau}$.

Operators: Next, some non-standard notation to separate cases of interest. A mapping is to be anointed an ***operator*** if it maps a space into that same space.

For an operator, element $\tau_{\ell m}$ is the ℓ-character amplitude in the output for unity (1) amplitude character *m* as the input for the operator. The mapping of a space into itself is to be liberally interpreted. A vector in $\mathbb{V}$ may be mapped into a dimensioned scalar value times a vector in $\mathbb{V}$.

Schrödinger quantum mechanics has the space of wave functions with differential operators such as the hamiltonian, the energy operator. The vectors in Schrödinger space are the wave functions $\psi(x)$ and the energy operator is:

$$\hat{H} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}\right) + V(x)$$

leading to the infamous and profoundly productive Schrödinger equation.

$$\hat{H}\psi(x) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2}\right) + V(x)\psi(x) = i\hbar\frac{\partial\psi}{\partial t}$$

The *std.* basis is the energy eigenfunction basis found by solving the equation:

$$\hat{H}\varphi_n(x) = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\varphi_n}{\partial x^2}\right) + V(x)\varphi_n(x) = E_n\,\varphi_n(x)$$

The E_n are the characteristic energy levels for the system, and the matrix representing the operator $\hat{H}$ is particularly simple.

$$\hat{H} \rightarrow \begin{bmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 \\ \vdots & & & \\ 0 & 0 & \dots & E_n \end{bmatrix}$$

A small complication is that *n* is usually infinite.

If a state of the quantum system is expanded in terms of the energy eigenfunction basis, $\psi(x) = \sum_n a_n\,\varphi_n(x)$, then the hamiltonian operating on the function yields $\hat{H}\psi(x) = \sum_n a_n\,\hat{H}\varphi_n(x) = \sum_n a_n\,E_n\,\varphi_n(x)$. As the $\varphi_n(x)$ are eigenfunctions, the hamiltonian, an operator for energy, does not mix the states of one characteristic energy with those of a distinct characteristic energy as indicated by the characteristic diagonal form of its matrix.

Momentum, a second physical observable, is represent by the operator: $\hat{p} = -i\hbar\frac{\partial}{\partial x}$.

The energy eigenfunctions are rarely also momentum eigenfunctions so momentum is represented by a complicated matrix:

$$\hat{p} \rightarrow \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \\ p_{n1} & p_{11} & \cdots & p_{nn} \end{bmatrix} \quad \text{where } p_{mn} = \langle \varphi_m | i\hbar \frac{\partial}{\partial x} | \varphi_n \rangle$$

As the matrix for $\hat{p}$ is not diagonal, making a measurement of momentum on an energy eigenstate mixes in other energy eigenstates. A year of quantum mechanics identifies the eigenstates of momentum (the momentum basis set) as: $\frac{1}{\sqrt{2\pi}} e^{i\left(p/\hbar\right)x}$.

Using the inner product to project out the p-momentum character part of $\varphi_n(x)$.

$$\int_{lower}^{upper} \left(\frac{1}{\sqrt{2\pi}} e^{i\left(p/\hbar\right)x} \right)^* \varphi_n(x)\, dx = \tilde{\varphi}_n(p/\hbar)$$

The Fourier transform of the spatial wave function is the amplitude of the momentum character components of the wavefunction. That is: $\tilde{\varphi}_n(p)$ is the Fourier transform of $\varphi_n(x)$. Note that the twiddle is used differently here. (Again: $\tilde{\varphi}_n(p)$ is the Fourier transform of $\varphi_n(x)$.)

! TIME OUT!: Leave this for graduate school. For now, realize the matrices can represent operators. If the representation is in terms of the eigenstates for that operator, then its matrix is diagonal. If the matrix has off-diagonal elements, then the operator mixes the states or causes transitions between states. Certain inner products are to be designated as matrix elements o_{mn}.

$$o_{mn} = \langle \varphi_m | \hat{o} | \varphi_n \rangle$$

Operators operate to the right so: $o_{mn} = \langle \varphi_m | \hat{o} | \varphi_n \rangle \Rightarrow \langle \varphi_m | \hat{o}\, \varphi_n \rangle \equiv \langle \hat{o}^\dagger\, \varphi_m | \varphi_n \rangle$. Dirac notation is a little ambiguous, so the action-to-the-right rule must be added. The action of the operator $\hat{o}$ to the right in the KET is identical to the action of its Hermitian conjugate $\hat{o}^\dagger$ in the BRA to the left.

Let's look ahead just a little further before returning to more concrete tasks.

Transformations: A (*basis*) transformation is to be analogous to a change of coordinates such as a rotation of a set of *x*-*y*-*z* axes. A physical vector does not change when the axes are rotated, but its component representation is different after the rotation. Transformations translate the components of a vector and the elements of an operator's matrix from those for one basis set to those for another basis set.

Suppose that the basis set for $\mathbb{V}$ is transformed by so the components of a vector are mapped to those in the new basis by $\mathbb{T}_V$ and the basis for $\mathbb{W}$ is transformed such that the components of a vector are mapped to those in the new basis by $\mathbb{T}_W$.

$$\tilde{b}' = \mathbb{T}_W\, \tilde{b} = \mathbb{T}_W\, \tilde{\tilde{\tau}}\, \tilde{a} = (\mathbb{T}_W\, \tilde{\tilde{\tau}}\, \mathbb{T}_V^{-1})\,(\mathbb{T}_V\, \tilde{a}) = (\mathbb{T}_W\, \tilde{\tilde{\tau}}\, \mathbb{T}_V^{-1})\,(\tilde{a}')$$

The result is that the transformation rule for the matrix representing a transformation is a distressing mess: $\tilde{\tilde{\tau}}' = (\mathbb{T}_W\, \tilde{\tilde{\tau}}\, \mathbb{T}_V^{-1})$. Fortunately, life is rarely this difficult. If the τ is an operator, then $\mathbb{V}$ and $\mathbb{W}$ are the same vector space, so $\mathbb{T}_V = \mathbb{T}_W$. For operators:

$$\tilde{a}' = \mathbb{T}\, \tilde{a} \qquad \tilde{b}' = \mathbb{T}\, \tilde{b} = \qquad \tilde{\tilde{\tau}}' = (\mathbb{T}\, \tilde{\tilde{\tau}}\, \mathbb{T}^{-1})$$

Example: Consider the operation $\vec{d} = (3\hat{i}\cdot\vec{r})\,\hat{i} + (\hat{i}\cdot\vec{r})\,\hat{j} + (2\,\hat{j}\cdot\vec{r})\,\hat{j}$ and the basis $\{\hat{i}, \hat{j}\}$ for both the initial and target two dimension spaces. The operator is applied to the two basis vectors for the space of two dimensional displacements.

$$\vec{d}(\hat{i}) = (3)\,\hat{i} + (1)\,\hat{j} + (0)\,\hat{j} \quad \text{and} \quad \vec{d}(\hat{j}) = (0)\,\hat{i} + (0)\,\hat{j} + (2)\,\hat{j}$$

Following the rule: *Each column of the transformation matrix is the coefficients from the mapping of the corresponding basis vector*,

$$\tau = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \text{ or } \begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad \begin{aligned} d_x &= 3\,x \\ d_y &= x + 2\,y \end{aligned}$$

Next consider this transformation as described by an observer rotated by + $^{\pi}/_{4}$ about the z-axis. The new basis directions for the rotated frame $\left\{ \hat{i}^{/}, \hat{j}^{/} \right\}$ are represented in terms of the basis vector used to define the mapping as $\left(\frac{1}{\sqrt{2}}\right)\left[\,\hat{i} + \hat{j}\,\right]$ and $\left(\frac{1}{\sqrt{2}}\right)\left[-\hat{i} + \hat{j}\,\right]$. Inverting the equations, $\hat{i} = \left(\frac{1}{\sqrt{2}}\right)\left[\,\hat{i}^{/} - \hat{j}^{/}\,\right]$ and $\hat{j} = \left(\frac{1}{\sqrt{2}}\right)\left[\,\hat{i}^{/} + \hat{j}^{/}\,\right]$.

A vector $\vec{v} = a_1\,\hat{i} + a_2\,\hat{j} = a_1\left(\frac{1}{\sqrt{2}}\right)\left[\,\hat{i}^{/} - \hat{j}^{/}\,\right] + a_2\left(\frac{1}{\sqrt{2}}\right)\left[\,\hat{i}^{/} + \hat{j}^{/}\,\right] = a_1^{/}\,\hat{i}^{/} + a_2^{/}\,\hat{j}^{/}$

That is the component transformation rule is:

$$\begin{bmatrix} a_1^{/} \\ a_2^{/} \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \left(\frac{1}{\sqrt{2}}\right)\begin{bmatrix} a_1 + a_2 \\ -a_1 + a_2 \end{bmatrix}$$

which identifies

$$\mathbb{R} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Compute the action of the operator on the new basis vectors.

$\vec{d}(\hat{i}^{/}) = \vec{d}(\left(\frac{1}{\sqrt{2}}\right)[\hat{i} + \hat{j}\,]) = (\frac{3}{\sqrt{2}})\,\hat{i} + (\frac{1}{\sqrt{2}})\,\hat{j} + (\sqrt{2})\,\hat{j} =$ $3\left(\frac{1}{\sqrt{2}}\right)[\hat{i} + \hat{j}\,] + (0)\left(\frac{1}{\sqrt{2}}\right)[-\hat{i} + \hat{j}\,] = 3\,\hat{i}^{/}$	$\hat{\tau}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$
$\vec{d}(\hat{j}^{/}) = \vec{d}(\left(\frac{1}{\sqrt{2}}\right)[-\hat{i} + \hat{j}\,]) = (-\frac{3}{\sqrt{2}})\,\hat{i} + (-\frac{1}{\sqrt{2}})\,\hat{j} + (\sqrt{2})\,\hat{j} =$ $(-1)\left(\frac{1}{\sqrt{2}}\right)[\hat{i} + \hat{j}\,] + (2)\left(\frac{1}{\sqrt{2}}\right)[-\hat{i} + \hat{j}\,] = -1\,\hat{i}^{/} + 2\,\hat{j}^{/}$	$\hat{\tau}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

Therefore the transformed representation for the operator matrix in the new basis is:

$$\tau^{/} = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}$$

This result is to be compared with that found by transforming the original matrix with the rotation matrix for + $^{\pi}/_{4}$ about the *z*-axis.

$$\tau' = \mathbb{R}\,\tau\,\mathbb{R}^{-1} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}$$

This stuff just works! There is another complication. While all classical physics is real, quantum mechanics requires complex values. The next topic is the generalization from vector spaces over real scalars to vector spaces over the scalar field of complex numbers.

Complex spaces and abstract basis (coordinate) transformations using Hermitian matrices and unitary transformations:

The focus now shifts from three dimensional real spaces of *physical* vectors to abstract complex spaces of arbitrary dimension. It is important to note the parallels between the current development for complex spaces with Hermitian operators and unitary transformations and the one for real spaces with symmetric operators and orthogonal transformations. The changes are just the generalizations necessary to accommodate complex numbers.

Begin with a complex vector space with an orthonormal basis: $\langle e_i | e_j \rangle = \delta_{ij}$

$$\{....|V_i\rangle,...\} = a_{i1}|e_1\rangle + a_{i2}|e_2\rangle + ... + a_{in}|e_n\rangle \;\; ; a_{ij} \in \mathbb{C}$$

As discussed in the vector space handout, the inner product must be modified for complex vectors to ensure that a vector's inner product with itself is positive (inner Product property I4).

$$|v_j\rangle \rightarrow \begin{bmatrix} a_{j1} \\ a_{j2} \\ \cdots \\ a_{jn} \end{bmatrix} = \tilde{v}_j \quad \text{and} \quad \langle v_i | \rightarrow \begin{bmatrix} a_{i1}^* & a_{i2}^* & \cdots & a_{in}^* \end{bmatrix} = \tilde{v}_i^\dagger$$

The BRA is represented as the *complex conjugate* of the transpose of the BRA column vector.

$$\begin{bmatrix} a_{i1}^* & a_{i2}^* & \dots & a_{in}^* \end{bmatrix} \begin{bmatrix} a_{j1} \\ a_{j2} \\ \dots \\ a_{jn} \end{bmatrix} = a_{i1}^* a_{j1} + a_{i2}^* a_{j2} + \dots + a_{in}^* a_{jn} = \tilde{v}_i^\dagger \tilde{v}_j = \langle v_i | v_j \rangle \qquad \text{[LT.10]}$$

Recall that the dagger **†** represents the Hermitian conjugate, the complex conjugate of the transpose.

Exercise: $|V_i\rangle = a_{i1}|e_1\rangle + a_{i2}|e_2\rangle + \dots + a_{in}|e_n\rangle$; give the corresponding representation of $\langle V_j|$. Compute $\langle V_i|V_j\rangle$ using $\langle e_i|\, e_j\rangle = \delta_{ij}$ and $\langle V_j| = a_{1j}^* \langle e_1| + a_{2j}^* \langle e_2| + \dots + a_{nj}^* \langle e_n|$. Compare with [LT.10].

For a change of basis (reference system), the components of the vectors obey a transformation rule. The *e-basis* is our standard basis and the *b-basis* in some other orthonormal basis for the same space.

$$|V_i\rangle = \tilde{a}_{i1}|e_1\rangle + \tilde{a}_{i2}|e_2\rangle + \dots + \tilde{a}_{in}|e_n\rangle = \tilde{a}_{i1}^b|b_1\rangle + \tilde{a}_{i2}^b|b_2\rangle + \dots + \tilde{a}_{in}^b|b_n\rangle$$

$$\tilde{a}_{im}^b = \langle b_m|V_i\rangle = \sum_\ell \langle b_m|e_\ell\rangle \tilde{a}_{i\ell} = \sum_\ell \tilde{\tilde{U}}_{m\ell}\, \tilde{a}_{i\ell}$$

$$\begin{bmatrix} \tilde{a}_{i1}^b \\ \tilde{a}_{i2}^b \\ \dots \\ \tilde{a}_{in}^b \end{bmatrix} = \begin{bmatrix} \langle b_1|e_1\rangle & \langle b_1|e_2\rangle & \dots & \langle b_1|e_n\rangle \\ \langle b_2|e_1\rangle & \langle b_2|e_2\rangle & \dots & \langle b_2|e_n\rangle \\ \dots & \dots & \dots & \dots \\ \langle b_n|e_1\rangle & \langle b_n|e_2\rangle & \dots & \langle b_n|e_n\rangle \end{bmatrix} \begin{bmatrix} \tilde{a}_{i1} \\ \tilde{a}_{i2} \\ \dots \\ \tilde{a}_{in} \end{bmatrix}; \quad \tilde{\tilde{U}} = \begin{bmatrix} \langle b_1|e_1\rangle & \langle b_1|e_2\rangle & \dots & \langle b_1|e_n\rangle \\ \langle b_2|e_1\rangle & \langle b_2|e_2\rangle & \dots & \langle b_2|e_n\rangle \\ \dots & \dots & \dots & \dots \\ \langle b_n|e_1\rangle & \langle b_n|e_2\rangle & \dots & \langle b_n|e_n\rangle \end{bmatrix}$$

The inverse transformation just interchanges the role of the *e's* and *b's*.

$$\left[U^{-1}\right]_{ij} = \langle e_i|b_j\rangle; \qquad \left[U\ \right]_{ij} = \langle b_i|e_j\rangle$$

(See inner product space property I3.) As the space is complex, $\langle e_i | b_j \rangle = \langle b_j | e_i \rangle^*$, the inverse transformation is the complex conjugate of the transpose – that is: the Hermitian conjugate of the forward transformation matrix. $\tilde{\tilde{U}}^{-1} = (\tilde{\tilde{U}}^t)^* = \tilde{\tilde{U}}^{\mathbf{t}}$

A matrix is a ***unitary*** matrix if its Hermitian conjugate is its inverse.

Properties of an inner product:	
I1. $\langle I \vert (\vert M \rangle + \vert N \rangle) = \langle I \vert M \rangle + \langle I \vert N \rangle = c_{im} + c_{in}$	*Distributive*
I2. $\langle I \vert (d \vert M \rangle) = d \langle I \vert M \rangle$ where d is any scalar.	*Scaling*
I3. $\langle I \vert M \rangle = \langle M \vert I \rangle^*$ asterisk means complex conjugate	*'Hermitian'*
I4. $\langle I \vert I \rangle \geq 0$ and $\langle I \vert I \rangle = 0$ if and only if $\vert I \rangle = 0$.	*Can be a Norm*

A matrix for which the Hermitian conjugate is the inverse is designated as a ***unitary*** matrix. The basis transformations between orthonormal basis sets for complex vector spaces are unitary. Unitary matrices are to Hermitian matrices what orthogonal matrices are to real symmetric matrices. (Real symmetric matrices are there own transposes and can be diagonalized by an orthogonal transformation. If a transformation is orthogonal, its inverse transformation is the transpose of the forward transformation. Hermitian matrices can be diagonalized using a unitary transformation.)

$$\mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix} \text{ and } \mathbb{U}^\dagger = \mathbb{U}^{-1} \text{ as } \sum_{i=1,2,3} U_{\ell i} U^*_{mi} = \sum_{i=1,2,3} U_{\ell i} U^\dagger_{im} = \delta_{\ell m} \Rightarrow \mathbb{U}\,\mathbb{U}^\dagger = \mathbb{1}.$$

Equivalently, the rows of a unitary matrix are the coefficients of normalized mutually orthogonal vectors under our standard inner product for complex spaces.

$$\begin{bmatrix} U_{i1}^* & U_{i2}^* & \dots & U_{in}^* \end{bmatrix} \begin{bmatrix} U_{j1} \\ U_{j2} \\ \dots \\ U_{jn} \end{bmatrix} = U_{i1}^* U_{j1} + U_{i2}^* U_{j2} + \dots + U_{in}^* U_{jn} = \tilde{u}_i^{\mathrm{t}} \tilde{u}_j = \delta_{ij}$$

The same statement applies to the columns considered as components of vectors.

What happens to inner products?

$$\sum_m (\tilde{a}_{jm}^b)^* \tilde{a}_{im}^b = \sum_m (\sum_s \tilde{\tilde{U}}_{ms} \tilde{a}_{js})^* (\sum_t \tilde{\tilde{U}}_{mt} \tilde{a}_{jt}) = \sum_{s,t} (\tilde{a}_{js})^* \tilde{a}_{jt} \sum_m (\tilde{\tilde{U}}_{ms})^* \tilde{\tilde{U}}_{mt}$$

use: $$\sum_m (\tilde{\tilde{U}}_{ms})^* \tilde{\tilde{U}}_{mt} = \sum_m (\tilde{\tilde{U}}_{sm}^t)^* \tilde{\tilde{U}}_{mt} = \sum_m (\tilde{\tilde{U}}_{sm}^{\mathrm{t}}) \tilde{\tilde{U}}_{mt} = \sum_m (\tilde{\tilde{U}}_{sm}^{-1}) \tilde{\tilde{U}}_{mt} = \delta_{st}$$

$$\sum_m (\tilde{a}_{jm}^b)^* \tilde{a}_{im}^b = \sum_s (\tilde{a}_{js})^* \tilde{a}_{js}$$

Unitary transformations preserve inner products of vectors in complex spaces. In particular, each vector transforms into a vector of the same magnitude.

Exercise: Use matrix properties to avoid the detailed component sums in the proof that unitary transformations preserve inner products $\tilde{v}^\dagger \tilde{w} = \tilde{v}'^\dagger \tilde{w}'$ if $\tilde{v}' = \mathbb{U}\ \tilde{v}$, $\tilde{w}' = \mathbb{U}\ \tilde{w}$ and $\mathbb{U}^{-1} = \mathbb{U}^\dagger$.

Hermitian Conjugate of an Operator:

The plan is to drop back from the comfortable world of matrices and to use abstract vector space notation to review the action on operators and to define the Hermitian conjugate of an operator and from that to characterize Hermitian operators. To close the loop, the matrix representation of a Hermitian operator is necessarily a Hermitian matrix.

What does an operator do? An operator maps one vector in a space to another vector in that same space.

$$\hat{o}\left|V_i\right\rangle = \left|V_j\right\rangle \quad for \ \left|V_i\right\rangle, \left|V_j\right\rangle \in \mathbb{V}$$

Given a basis set for $\mathbb{V}$, the (linear) operator is completely characterized by its action on that basis.

$$\left|V_i\right\rangle = \sum_m a_{im}\left|e_m\right\rangle \ \text{ and } \ \left|V_j\right\rangle = \sum_k a_{jk}\left|e_k\right\rangle$$

$$a_{jk} = \sum_m \left\langle e_k\right|\hat{o}\left|e_m\right\rangle a_{im} = \sum_m o_{km}\ a_{im}$$

The matrix representing the operator has elements $o_{km} = \left\langle e_k\right|\hat{o}\left|e_m\right\rangle$.

Exercise: The matrix that represents an operator has elements O_{mn} = $\left\langle m\right|\hat{O}\left|n\right\rangle = \left\langle m\right|\hat{O}n\left\rangle$. Show that Hermitian operators are represented by Hermitian matrices.

Exercise: Show that if a Hermitian matrix is transformed using a similarity transformation with a unitary matrix that the result is also a Hermitian matrix. Given $\tilde{\tilde{H}} = \tilde{\tilde{H}}^{\dagger}$ and $\mathbb{U}^{-1} = \mathbb{U}^{\dagger}$ show that $\tilde{\tilde{H}}' = (\tilde{\tilde{H}}')^{\dagger}$ where $\tilde{\tilde{H}}' = \mathbb{U}\tilde{\tilde{H}}\ \mathbb{U}^{-1}$. Use matrix properties rather than detailed matrix element based proofs.

Which way do operators operate? Operators operate to their right. A few examples have included differential operators which are known to operate to the right.

$$(g(x))^* \left(d/dx\right) f(x) = (g(x))^* \ df/dx$$

Given that $\hat{o}\left|V_i\right\rangle = \left|V_j\right\rangle$, the rule is:

$$\left\langle V_m\right|\hat{o}\left|V_i\right\rangle \rightarrow \left\langle V_m\right|\left\{\hat{o}\left|V_i\right\rangle\right\} = \left\langle V_m\right|\left\{\left|V_j\right\rangle\right\} = \left\langle V_m\right|V_j\left\rangle$$

This prescription directs that the operator act to the right mapping $|V_i\rangle$ to $|V_j\rangle$ and that then the inner product of V_m and V_j be computed. Well, not exactly. An example follows in which some flesh is to be wrapped around this abstract skeleton.

Hermitian conjugate takes the operator to the left. To define action to the left, a partner operator $\hat{o}^\dagger$, the Hermitian conjugate is introduced. The Hermitian conjugate lives to act on the left (on the *BRA* part of the *BRA-KET*). It is defined by its action:

$$\left\langle \hat{o}^\dagger V_m \middle| V_j \right\rangle = \left\langle V_m \middle| \hat{o}\, V_j \right\rangle \quad \forall \left|V_j\right\rangle, \left|V_m\right\rangle \in \mathbb{V}$$

That's it. The Hermitian conjugate of an operator ô is whatever it takes operating on the *BRA* vector to yield the same inner product as if the operator ô had acted on the *KET*.

Example: Finding a Hermitian conjugate. The abstract formulation above applies universally to all vectors spaces real and imagined providing information proportional to the inverse of its scope. Consider the space of all functions of *x* defined on the interval (*a, b*) that are continuously differentiable to all orders. Further all functions in the space have the property that *f(a) = f(b)*. The operator is $\hat{D} = {}^{d}\!/_{dx}$ and the inner product is the function space standard.

$$\langle g | h \rangle = \frac{1}{N} \int_a^b \left(g(x)\right)^* h(x)\, dx$$

$$\langle g | \hat{D} | h \rangle = \left\langle g \middle| \hat{D}\, h \right\rangle = \frac{1}{N} \int_a^b \left(g(x)\right)^* {}^{dh}\!/_{dx}\; dx$$

Finally! A use for Calculus II. Integrate by parts.

$$\frac{1}{N} \int_a^b \left(g(x)\right)^* {}^{dh}\!/_{dx}\; dx = \frac{1}{N} \left\{ \left[g(x)^*\, h(x) \right]_a^b + \int_a^b \left(-{}^{dg}\!/_{dx} \right)^* h(x)\; dx \right\} = 0 + \left\langle \left(-{}^{d}\!/_{dx} \right) g \middle| h \right\rangle$$

$$\left\langle g \middle| \hat{D}\, h \right\rangle = \left\langle (-\hat{D})\, g \middle| h \right\rangle \; or \; \left(-{}^{d}\!/_{dx} \right) = \left({}^{d}\!/_{dx} \right)^\dagger$$

Note that the boundary conditions $g(a) = g(b)$ and $h(a) = h(b)$ are assumed. These conditions are often stated as the functions are periodic with period $b-a$ or that the functions vanish at a and b.

The plan is to leave the comfortable world of matrices and to use abstract vector space notation to review the action on operators and to define the Hermitian conjugate of an operator and from that to characterize Hermitian operators. To close the loop, the matrix representation of a Hermitian operator is necessarily a Hermitian matrix.

Hermitian operator: An operator is Hermitian if it is its own Hermitian conjugate. The example above $^{d}/_{dx}$ did not meet this standard, but all operators that represent physical observables do so let us press onward. What can be said about the matrix that represents a Hermitian operator? The matrix elements have the form: $\langle e_k | \hat{H} | e_m \rangle = \langle e_k | \hat{H}\, e_m \rangle = \langle \hat{H}^{\dagger} e_k | e_m \rangle$. The last equality follows from the definition of the Hermitian conjugate operator. **IMPORTANT:** There is another process that moves things from the *KET* to the *BRA*. The third property of the inner product interchanges the *BRA* and *KET* at the cost of one complex conjugation.

I3. $\langle I | M \rangle = \langle M | I \rangle^{*}$ asterisk means complex conjugate *'Hermitian'*

$$\langle e_k | \hat{H} | e_m \rangle = \langle e_k | \hat{H}\, e_m \rangle = \langle \hat{H}^{\dagger} e_k | e_m \rangle = \langle e_m | \hat{H}^{\dagger} e_k \rangle^{*}$$

If the operator is Hermitian, then $\hat{H} = \hat{H}^{\dagger}$.

$$\langle e_k | \hat{H} | e_m \rangle = \langle e_k | \hat{H}\, e_m \rangle = \langle \hat{H}^{\dagger} e_k | e_m \rangle = \langle e_m | \hat{H}^{\dagger} e_k \rangle^{*} = \langle e_m | \hat{H}^{\dagger} | e_k \rangle^{*} = \langle e_m | \hat{H} | e_k \rangle^{*}$$

$$H_{km} = (H_{mk})^{*} \quad \text{or} \quad \mathbb{H} = (\mathbb{H}^{t})^{*} = \mathbb{H}^{t}$$

The matrix representing the operator is equal to the complex conjugate of its transpose. That is: it's Hermitian. Hermitian is used to describes a bundle of features because all those features are related.

Exercise: Consider the space of all functions of x defined on the interval (a, b) that are continuously differentiable to all orders. Further all functions in the space have the property that $f(a) = f(b)$. The operator is $\hat{d} = -i\left({}^{d}/_{dx}\right)$ and the inner product is the function space standard. The factor $i = \sqrt{-1}$.

$$\langle g | h \rangle = \frac{1}{N} \int_a^b \left(g(x)\right)^{*} h(x)\, dx$$

Find the differential operator representing $\hat{d}^{\dagger}$. Is $\hat{d}$ Hermitian? Explain.

Exercise: Consider the space of all functions of *x*, a single real variable, defined on the interval (*a, b*) that are continuously differentiable to all orders. Further all functions in the space have the property that *f(a) = f(b)*. The operator is $\hat{\ell} = x$ and the inner product is the function space standard.

$$\langle g | h \rangle = \frac{1}{N} \int_a^b \left(g(x) \right)^* h(x)\, dx$$

Find the differential operator representing $\hat{\ell}^\dagger$. Is $\hat{\ell}$ Hermitian? Explain.

Operations that equal the identity when applied a second time their own inverses!

**** Let's find a term for this distinction.

Unary minus:	- (- A) = A
Multiplicative inverse:	1/(1/A) = A
Complex conjugate:	$(A^*)^* = A$
Matrix transpose:	$(\mathbb{M}^t)^t = \mathbb{M}$
Matrix Hermitian Conjugate:	$(\mathbb{M}^t)^t = \mathbb{M}$

Exercise: Use the definition of Hermitian conjugate and property I3 of the inner product to show that $\left(\hat{o}^\dagger\right)^\dagger = \hat{o}$. Use $\left\langle \hat{o}^\dagger V_m \middle| V_j \right\rangle = \left\langle V_m \middle| \hat{o}\, V_j \right\rangle \quad \forall \left| V_j \right\rangle, \left| V_m \right\rangle \in \mathbb{V}$.

It is understood that operators act to the right so $\langle V_m | \hat{o} | V_j \rangle \Rightarrow \left\langle V_m \middle| \hat{o}\, V_j \right\rangle$. Adding the Hermitian conjugate operator: $\langle V_m | \hat{o} | V_j \rangle \Rightarrow \left\langle V_m \middle| \hat{o}\, V_j \right\rangle = \left\langle \hat{o}^\dagger V_m \middle| V_j \right\rangle$. Beware of a careless interpretations that arise from too much familiarity with Hermitian operators for which $\langle V_m | \hat{h} | V_j \rangle \Rightarrow \left\langle V_m \middle| \hat{h}\, V_j \right\rangle = \left\langle \hat{h} V_m \middle| V_j \right\rangle$ as $\hat{h}^\dagger = \hat{h}$. *In general, the action direction and ordering of operators matters*.

The eigenvalue problem:

Hermitian operators operate in a complex space and have eigenvectors with their associated eigenvalues.

$$\hat{H} \left| V_i \right\rangle = \lambda_i \left| V_i \right\rangle \qquad \text{Eigenvalue equation}$$

An eigenvalue might be found by solving a differential equation and applying boundary conditions. Only the matrix method is to be reviewed.

Considered in its matrix representation

$$\langle V_i| \rightarrow \begin{bmatrix} a_{i1}^* & a_{i2}^* & \dots & a_{in}^* \end{bmatrix} = \tilde{v}_i^\dagger \quad ; \quad |V_j\rangle \rightarrow \begin{bmatrix} a_{j1} \\ a_{j2} \\ \dots \\ a_{jn} \end{bmatrix} = \tilde{v}_j \quad ; \quad \langle V_i| = \left((|V_i\rangle)^t \right)^* = (|V_i\rangle)^\dagger$$

$$\hat{H}|V_i\rangle = \lambda_i |V_i\rangle \rightarrow \mathbb{H}\ \tilde{v}_j = \lambda_j\ \mathbb{1}\ \tilde{v}_j \quad \text{eigenvalue equation}$$

set $\quad |\mathbb{H} - \lambda\ \mathbb{1}| = 0 \quad$ characteristic equation

Find the set of n (for $n \times n$ $\mathbb{H}$) eigenvalues. Substitute each eigenvalue in turn into the eigenvalue equation and solve for the eigen-directions. The scale of each eigenvector being set by normalizing once the eigenvalue equation provides the ratios of the various components.

where the dagger (**t**) represents the Hermitian conjugate, the complex conjugate of the transpose. Compute the Hermitian conjugate of the eigenvalue equation.

$$(\mathbb{H}\ \tilde{v}_j = \lambda_j\ \tilde{v}_j)^{\mathbf{t}} = (\tilde{v}_j^\dagger\ \mathbb{H}^{\mathbf{t}} = \tilde{v}_j^\dagger\ \lambda_j^*) \quad (\textit{in general})$$

$$= (\tilde{v}_j^\dagger\ \mathbb{H} = \tilde{v}_j^\dagger\ \lambda_j^*) \quad (\textit{for Hermitian})$$

Theorem: The eigenvalues of a Hermitian matrix are real.

$$\tilde{v}_j^\dagger\ \mathbb{H}\ \tilde{v}_j = (\tilde{v}_j^\dagger \mathbb{H})\ \tilde{v}_j = \tilde{v}_j^\dagger\ (\mathbb{H}\ \tilde{v}_j) \quad \text{associate the multiplication}$$

$$= \tilde{v}_j^\dagger\ \lambda_j\ \tilde{v}_j = \tilde{v}_j^\dagger\ \lambda_j^*\ \tilde{v}_j \quad \text{use the eigenvalue equations}$$

$$= \lambda_j\ \tilde{v}_j^\dagger\ \tilde{v}_j = \lambda_j^*\ \tilde{v}_j^\dagger\ \tilde{v}_j \quad \text{or} \quad \lambda_j^* = \lambda_j \quad \text{real, no imaginary part}$$

Recall

that the inner product of a vector with itself is $\tilde{v}_j^\dagger\ \tilde{v}_j$ is non-zero as long as $\tilde{v}_j^\dagger\ \tilde{v}_j$ is not the zero vector

Theorem: The eigenvectors of distinct eigenvalues are orthogonal.

$\tilde{v}_j^\dagger \mathbb{H}\ \tilde{v}_i\ -\ \tilde{v}_j^\dagger\ \mathbb{H}\ \tilde{v}_i = 0$ a simple truth

$(\tilde{v}_j^\dagger \mathbb{H})\tilde{v}_i\ -\ \tilde{v}_j^\dagger\ (\mathbb{H}\ \tilde{v}_i) = 0$ associate the multiplication

$(\tilde{v}_j^\dagger\ \lambda_j^*\ )\tilde{v}_i\ -\ \tilde{v}_j^\dagger\ (\lambda_i \tilde{v}_i) = 0$ use the eigenvalue equations

$(\lambda_j^* \tilde{v}_j^\dagger\ \tilde{v}_i)\ -\ \lambda_i\ \tilde{v}_j^\dagger\ \tilde{v}_i\ = 0$ factor out constants

$\left(\lambda_j^* - \lambda_i\right) \tilde{v}_j^\dagger\ \tilde{v}_i = 0$ eigenvalues assumed not equal

$\left(\lambda_j - \lambda_i\right) \tilde{v}_j^\dagger\ \tilde{v}_i = 0$ eigenvalues are real and distinct

$\tilde{v}_j^\dagger\ \tilde{v}_i = 0$ orthogonal eigenvector for distinct values

---

Abstract Vector Space Notation.

The previous theorems are to be repeated to demonstrate the abstract vector space notation and to exercise the definition of Hermitian conjugate and inner product property I3. The abstract notation proofs establish the results for all complex (and real) vector spaces. More abstract developments have a broader scope of validity.

Begin with the eigenvalue equation for a Hermitian operator and a complex space.

$$\hat{H}\left|\varphi_i\right\rangle = \lambda_i\ \left|\varphi_i\right\rangle$$

Consider a matrix element of a Hermitian operator taken between two eigenvectors/states.

Start with $\left\langle\varphi_i\middle|\hat{H}\middle|\varphi_j\right\rangle$ which involves two distinct *non-zero* eigenstates.

$\left\langle\varphi_i\middle|\hat{H}\middle|\varphi_j\right\rangle = \left\langle\varphi_i\middle|\hat{H}\,\varphi_j\right\rangle = \left\langle\hat{H}^\dagger\varphi_i\middle|\,\varphi_j\right\rangle$ *acts right; move left by dagger-op*

$\left\langle\varphi_i\middle|\hat{H}\,\varphi_j\right\rangle = \left\langle\varphi_i\middle|\lambda_j\,\varphi_j\right\rangle = \left\langle\hat{H}^\dagger\varphi_i\middle|\,\varphi_j\right\rangle = \left\langle\varphi_j\middle|\,\hat{H}^\dagger\varphi_i\right\rangle^*$ *eigenvalue equation plus I3*

$\lambda_j\left\langle\varphi_i\middle|\,\varphi_j\right\rangle = \left\langle\varphi_j\middle|\,\hat{H}^\dagger\varphi_i\right\rangle^* = \left\langle\varphi_j\middle|\,\hat{H}\ \varphi_i\right\rangle^* = \left\langle\varphi_j\middle|\,\lambda_i\varphi_i\right\rangle^*$ *λ is a scalar and* $\hat{H}^\dagger = \hat{H}$

$\lambda_j \langle \varphi_i | \varphi_j \rangle = \langle \varphi_j | \lambda_i \varphi_i \rangle^* = \lambda_i^* \langle \varphi_j | \varphi_i \rangle^* = \lambda_i^* \langle \varphi_i | \varphi_j \rangle$ *just a scalar and I3*

$(\lambda_j - \lambda_i^*) \langle \varphi_i | \varphi_j \rangle = 0$

Theorem: The eigenvalues of a Hermitian operator are real.

Choose $i = j$. $(\lambda_j - \lambda_j^*) \langle \varphi_j | \varphi_j \rangle = 0 \quad \Rightarrow (\lambda_j - \lambda_j^*) = 0$

The eigenstates are nonzero so the self-inner product is real positive. It follows that λ_j is real.

All the eigenvalues of Hermitian operators are real.

Theorem: Eigenvectors with distinct eigenvalues are orthogonal.

$(\lambda_j - \lambda_i^*) \langle \varphi_i | \varphi_j \rangle = (\lambda_j - \lambda_i) \langle \varphi_i | \varphi_j \rangle = 0$ used: *eigenvalues are real*

$(\lambda_j - \lambda_i) \neq 0 \Rightarrow \langle \varphi_i | \varphi_j \rangle = 0$ used: *eigenvalues are distinct*

Eigenstates with distinct eigenvalue have zero mutual inner product; they are orthogonal.

In cases in which the eigenvalues are degenerate, the independent eigenstates can be chosen to be orthogonal. Gram-Schmidt provides is a procedure suited to developing a mutually orthogonal set from an original set.

The previous two proofs can be condensed if the abstract notation is pushed a little further.

$$\langle V_i | \hat{O} | V_j \rangle = \langle V_i | \hat{O}\, V_j \rangle = \langle \hat{O}^\dagger\, V_i | V_j \rangle \quad \forall |V_i\rangle, |V_j\rangle \in \mathbb{V}$$

Taken with property I3 and the eigenvalue relation

$$\hat{O} | V_j \rangle = | \hat{O} V_j \rangle = o_j | V_j \rangle \Rightarrow \langle \hat{O}^\dagger\, V_j | = o_j^* \langle V_j |$$

You should attempt more economical proofs based on this relation.

How do you find the unitary matrix $\mathbb{U}$ that diagonalizes a Hermitian matrix $\mathbb{H}$?

This section is represented as 3D, but the development should be understood to be N dimensional. The solution follows from result previously derived. Consider the generic representation of the eigenvectors and eigenvalue problem for a Hermitian matrix for a complex space where the left-hand equation is in the original coordinate representation while the right-hand is the form after transforming to the preferred basis set (the eigenvector basis) to yield the diagonal form of $\mathbb{H}$.

$$\begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{12} & H_{22} & H_{23} \\ H_{13} & H_{23} & H_{33} \end{bmatrix}\begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} = E_i \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} \quad where\ i = 1,2,3 \quad \begin{bmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_3 \end{bmatrix}\begin{bmatrix} a_i^{/} \\ b_i^{/} \\ c_i^{/} \end{bmatrix} = E_i \begin{bmatrix} a_i^{/} \\ b_i^{/} \\ c_i^{/} \end{bmatrix}$$

original basis form in preferred basis

The preferred frame has its axes along the eigen-directions (functions). The symbol E_i represents the eigenvalue of the Hermitian matrix (operator) $\mathbb{H}$ which has elements H_{mn}.

Hence, after transformation, the *normalized* eigenvectors must be in the canonical form (up to a multiplicative phase which may be set to one):

$$\begin{bmatrix} a_1^{/} \\ b_1^{/} \\ c_1^{/} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix}\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}; \qquad \begin{bmatrix} a_2^{/} \\ b_2^{/} \\ c_2^{/} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad \begin{bmatrix} a_3^{/} \\ b_3^{/} \\ c_3^{/} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Recalling that the inverse of $\mathbb{U}$ is its Hermitian conjugate:

$$\begin{bmatrix} U_{11}^* & U_{21}^* & U_{31}^* \\ U_{12}^* & U_{22}^* & U_{32}^* \\ U_{13}^* & U_{23}^* & U_{33}^* \end{bmatrix}\begin{bmatrix} a_1^{/} \\ b_1^{/} \\ c_1^{/} \end{bmatrix} = \begin{bmatrix} U_{11}^* & U_{21}^* & U_{31}^* \\ U_{12}^* & U_{22}^* & U_{32}^* \\ U_{13}^* & U_{23}^* & U_{33}^* \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$$

or $U_{11} = a_1^*;\ U_{12} = a_2^*;\ U_{13} = a_3^*$

Each row of the unitary transformation matrix is the complex conjugates of the components of the corresponding *N component* eigenvector.

$$\mathbb{U} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix} = \begin{bmatrix} a_1^* & b_1^* & c_1^* \\ a_2^* & b_2^* & c_2^* \\ a_3^* & b_3^* & c_3^* \end{bmatrix}$$

Exercise: Each row of the transformation matrix $\mathbb{U}$ is the complex conjugate of an eigenvector. What are the columns of $\mathbb{U}^{-1}$? $\mathbb{U}^{-1}$ is the same as what other operation on $\mathbb{U}$?

Hermitian matrices (operators) represent measurable values in quantum physics. The eigenvectors (functions) of these matrices (operators) are orthogonal if they are associated with distinct eigenvalues. The eigenvectors that correspond to a degenerate eigenvalue may be chosen to be orthogonal, and processes such as Gram-Schmidt can produce a suitable orthogonal set from an initial set.

Unitary transformations preserve all inner products so the standard definition as the sum of the products of the KET components with the complex conjugate of the BRA component coefficients remains valid. ⇒ The eigenvectors can be unity normalized.

$$\begin{bmatrix} a_{i1}^* & a_{i2}^* & \dots & a_{in}^* \end{bmatrix} \begin{bmatrix} a_{j1} \\ a_{j2} \\ \dots \\ a_{jn} \end{bmatrix} = a_{i1}^* a_{j1} + a_{i2}^* a_{j2} + \dots + a_{in}^* a_{jn} = \tilde{v}_i^\dagger \, \tilde{v}_j$$

A Hermitian matrix can be diagonalized by a unitary transformation. The form of the transformation matrix is the collection of the complex conjugates of the eigenvectors as the rows of the matrix. The orthonormal form of the eigenvectors for the complex space ensures that the matrix is unitary.

Oscillator space example: This example is so simple that it can only illustrate basic features, but here it goes.

Our oscillator space is the space of all solutions for the harmonic oscillator with unit frequency.

$$x(t) = c\left[\sqrt{2}\,\cos(t)\right] + d\left[\sqrt{2}\,\sin(t)\right] = c^{/}\left[e^{it}\right] + d^{/}\left[e^{-it}\right]$$

Proposed is a standard basis set $\left\{\ \hat{e}_1(t) = \sqrt{2}\,\cos(t)\ ;\ \hat{e}_2(t) = \sqrt{2}\,\sin(t)\ \right\}$ plus an alternative basis set for the space $\left\{\ \hat{b}_1(t) = e^{it}\ ;\ \hat{b}_2(t) = e^{-it}\ \right\}$. The normalizations have been chosen so that both sets are orthonormal using the inner product.

$$\langle f\,|\,g\rangle = \tfrac{1}{2\pi}\int_0^{2\pi} f^*(t)\ g(t)\ dt$$

Using our canonical recipe, the unitary transformation from the standard basis to the alternative basis has elements

$$U_{ij} = \left\langle \hat{b}_i \middle| \hat{e}_j \right\rangle \ \text{ for example } \ U_{11} = \left\langle \hat{b}_1 \middle| \hat{e}_1 \right\rangle = \tfrac{1}{2\pi}\int_0^{2\pi}\left(e^{it}\right)^*\ \sqrt{2}\,\cos(t)\ dt = \tfrac{1}{\sqrt{2}}$$

Continuing:

$$\mathbb{U} = \begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix} \text{ and hence } \mathbb{U}^{\dagger} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{bmatrix}$$

Let's check this by computing $\mathbb{U}\,\mathbb{U}^{\dagger}$

$$\begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{1} \quad (\textbf{\textit{It works !}})$$

Suppose that the system has a solution vector:

$$x(t) = \cos(t) = \left(1/\sqrt{2}\right)\left[\sqrt{2}\,\cos(t)\right] + (0)\left[\sqrt{2}\,\sin(t)\right]$$

Transforming to the alternative basis,

$$\begin{bmatrix} c' \\ d' \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

Hence

$$x(t) = \left(1/2\right)\left[e^{it}\right] + \left(1/2\right)\left[e^{-it}\right] = \cos(t) \quad \text{(It checks !)}$$

Now, let's move on to operators. There are few choices available as the problem is simple and as we need operators that map oscillator space into oscillator space.

The time shift operator $\hat{\tau}$: $t \to t + \pi/2$

In the standard basis $\left\{\sqrt{2}\ \cos(t) \to \sqrt{2}\ \cos(t+\pi/2) = -\sqrt{2}\ \sin(t)\,;\sqrt{2}\ \sin(t) \to \sqrt{2}\ \sin(t+\pi/2) = \sqrt{2}\ \cos(t)\right\}$

so

$$\hat{\tau} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

The element $\tau_{\ell m}$ is the ℓ-character amplitude in the output for unity (1) amplitude character *m* as the input for the operator.

Develop $\hat{\tau}$ using the prescribed method:

$$x(t) = c\left[\sqrt{2}\ \cos(t)\right] + d\left[\sqrt{2}\ \sin(t)\right]$$

$$\begin{aligned} \hat{\tau}\, x(t) &= c\left[\sqrt{2}\ \cos(t+\pi/2)\right] + d\left[\sqrt{2}\ \sin(t+\pi/2)\right] \\ &= d\left[\sqrt{2}\ \cos(t)\right] - c\left[\sqrt{2}\ \sin(t)\right] = a\left[\sqrt{2}\ \cos(t)\right] + b\left[\sqrt{2}\ \sin(t)\right] \end{aligned}$$

The action of an operator on a function (vector) is to return a function (vector). That returned function has an expansion in terms of the basis set. A matrix relation summarizes the action of the operator:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} \quad \Rightarrow \quad [\tau]_{\ell m} = \langle e_\ell \,|\, \{\tau | e_m\rangle\}\,.$$

The element $[\tau]_{\ell m}$ is the amount of ℓ character in the result for a unit of *m* character in the input. To find $[\tau]_{11}$, one assumes a pure one input (*c=1; d=0)*, and looks for one character in the result.

$$\hat{\tau}\, x(t) = 1\left[\sqrt{2}\,\cos(t+\pi/2)\right] = (-1)\left[\sqrt{2}\,\sin(t)\right]$$

There is none! $[\tau]_{11} = 0.$ For $[\tau]_{21}$, assume a pure one input (*c=1; d=0)*, and looks for character of type two ($\sqrt{2}\,\sin(t)$) in the result. The result is (-1) times type-two behavior so $[\tau]_{21} = -1.$

Exercise: Extend this reasoning to find $[\tau]_{12}$ and $[\tau]_{22}$.

Alternatively, one can compute the elements using $[\tau]_{\ell m} = \langle e_\ell | \{\tau | e_m\rangle\}$.

$$[\tau]_{11} = \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[\hat{\tau}\,\sqrt{2}\,\cos(t)\right]\,dt$$

$$= \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[\sqrt{2}\,\cos(t+\pi/2)\right]\,dt = \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[-\sqrt{2}\,\sin(t)\right]\,dt = 0$$

$$[\tau]_{12} = \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[\hat{\tau}\,\sqrt{2}\,\sin(t)\right]\,dt$$

$$= \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[\sqrt{2}\,\sin(t+\pi/2)\right]\,dt = \tfrac{1}{2\pi}\int_0^{2\pi} \sqrt{2}\,\cos(t)\left[\sqrt{2}\,\cos(t)\right]\,dt = 1$$

Exercise: Continue the $[\tau]_{\ell m} = \langle e_\ell | \{\tau | e_m\rangle\}$ method to find $[\tau]_{21}$ and $[\tau]_{22}$.

Applying the time shift to the alternative basis, $\left\{\, e^{it} \to i\, e^{it};\; e^{-it} \to -i\, e^{-it} \,\right\}$. In the alternative basis each basis vector is mapped to a constant times itself. Clearly, $\hat{\tau}$ is diagonal in this basis.

$$\hat{\tau}^{/} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

Transforming $\hat{\tau}$ to the alternative basis, ($\hat{\tau}^{/} = \mathbb{U}\ \hat{\tau}\ \mathbb{U}^{-1} = \mathbb{U}\ \hat{\tau}\ \mathbb{U}^{\mathbf{t}}$)

$$\hat{\tau}^{/} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad (amazing\ !)$$

The only other operator to be considered is time (reversal) inversion $\hat{i}$: $t \to -t$

Exercise: Show that in the standard basis, the time inversion operator matrix is:

$$\hat{i} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Exercise: Show that in the alternative basis, the time inversion operator matrix is:

$$\hat{i}^{/} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Exercise: Show that time inversion transforms in the canonical fashion.

$$\hat{i}^{/} = \mathbb{U}\ \hat{i}\ \mathbb{U}^{\mathbf{t}}$$

Quantum mechanics is formulated in a complex vector space with Hermitian operators representing physical observables.

Advanced Example: ***DO NOT READ THIS SEGMENT !!***

The topic is still oscillator space:

Standard basis: $\left\{ \hat{e}_1(t) = \sqrt{2}\ \cos(t)\ ;\ \hat{e}_2(t) = \sqrt{2}\ \sin(t) \right\}$

Alternative basis: $\left\{ \hat{b}_1(t) = e^{it}\ ;\ \hat{b}_2(t) = e^{-it} \right\}$

The operator $^{d}/_{dt}$ can be represented for each basis

$$\frac{d}{dt} \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{std basis} \qquad \frac{d}{dt} \to \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad \text{alternative}$$

It is interesting to note that in the standard basis:

$$\frac{d^2}{dt^2} \to \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{Compare: } \frac{d}{dt}\left(\frac{d}{dt}\right) \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

The point of these representations of operators is that ***they represent the operators, actions and algebra; the representations work and can be adopted as a faithful substitute for the original problem.***

Continuing in the standard basis,

$$\frac{d^3}{dt^3} \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \frac{d^4}{dt^4} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \frac{d^{4m}}{dt^{4m}} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\frac{d^{4m+1}}{dt^{4m+1}} \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \frac{d^{4m+2}}{dt^{4m+2}} \to \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \frac{d^{4m+3}}{dt^{4m+3}} \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

As all the time derivatives are known, the time development operator can be formed. Recalling the Taylor's series expansion:

$$f(t+\Delta) = f(t) + \frac{df}{dt}\Delta + \frac{1}{2!}\frac{d^2 f}{dt^2}\Delta^2 + \ldots + \frac{1}{n!}\frac{d^n f}{dt^n}\Delta^n + \ldots$$

FUNCTIONS OF MATRICES $\mathbb{M}$ ARE DEFINED BY SUBSTITUTING THE MATRIX INTO THE POWER SERIES EXPANSION OF THE FUNCTION.

Use: $\mathbb{M}^0 \equiv \mathbb{1}$; $\mathbb{M}^1 \equiv \mathbb{M}$; $\mathbb{M}^2 \equiv \mathbb{M}\,\mathbb{M}$; . . .

Now just substitute the corresponding matrices for each derivative and sum the infinite series to find …. Baloney ! *Actually, the sum can be computed in this case as the derivative matrices separate into odd and even order sequences each with alternating signs. Completing the sums yields matrices with elements that are cosine and sine. The standard method however is to transform to a basis in which the operator matrix is diagonal. This approach makes life 'easy'.*

Eigen-basis and diagonal forms to the rescue! The alternative basis is the basis in which the time derivative is diagonal.

Alternative basis: $\left\{ \hat{b}_1(t) = e^{it} \ ; \ \hat{b}_2(t) = e^{-it} \right\}$

$$\frac{d}{dt} \rightarrow \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad \Rightarrow \quad \frac{d^n}{dt^n} \rightarrow \begin{bmatrix} i^n & 0 \\ 0 & (-i)^n \end{bmatrix}$$

The nth power of a diagonal matrix is just a matrix with the nth powers of the original diagonal elements on its diagonal. In our alternative basis, the time development operator becomes

$$\hat{T}_b(\Delta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Delta^0 + \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \Delta^1 + \tfrac{1}{2!} \begin{bmatrix} i^2 & 0 \\ 0 & (-i)^2 \end{bmatrix} \Delta^2 + \ldots + \tfrac{1}{n!} \begin{bmatrix} i^n & 0 \\ 0 & (-i)^n \end{bmatrix} \Delta^n + \ldots$$

$$\hat{T}_b(\Delta) = \begin{bmatrix} (i\Delta)^0 & 0 \\ 0 & (-i\Delta)^0 \end{bmatrix} + \begin{bmatrix} (i\Delta)^1 & 0 \\ 0 & (-i\Delta)^1 \end{bmatrix} + \tfrac{1}{2!} \begin{bmatrix} (i\Delta)^2 & 0 \\ 0 & (-i\Delta)^2 \end{bmatrix} + \ldots + \tfrac{1}{n!} \begin{bmatrix} (i\Delta)^n & 0 \\ 0 & (-i\Delta)^n \end{bmatrix} + \ldots$$

$$\hat{T}_b(\Delta) = \begin{bmatrix} e^{i\Delta} & 0 \\ 0 & e^{-i\Delta} \end{bmatrix}$$

How does the time-development operator work?

Suppose that one has the state: $x(t) = a \left[e^{it} \right] + b \left[e^{-it} \right] \rightarrow \begin{bmatrix} a \\ b \end{bmatrix}$

The time operator maps this state into the state at time $t + \Delta$.

$$\hat{T}_b(\Delta) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} e^{i\Delta} & 0 \\ 0 & e^{-i\Delta} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a\, e^{i\Delta} \\ b\, e^{i\Delta} \end{bmatrix}$$

$$\hat{T}\left(x(t)\right) = a\ e^{i\Delta} \left[e^{it} \right] + b\, e^{-i\Delta} \left[e^{-it} \right] = a \left[e^{i(t+\Delta)} \right] + b\, e^{-i\Delta} \left[e^{-i(t+\Delta)} \right]$$

The time development by Δ operator transforms the state into the state that it will be at a time Δ later.

What about the standard basis?

The time development operator can be transformed into the standard basis. U is the transformation from standard to alternative so going the reverse direction:

$\widehat{T}_e(\Delta) = \mathbb{U}^{-1}\widehat{T}_b(\Delta)\ (\mathbb{U}^{-1})^{-1} = \mathbb{U}\ \mathbf{t}\widehat{T}_b(\Delta)\ \mathbb{U}$ inverse of a unitary matrix

$$= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} \begin{bmatrix} e^{i\Delta} & 0 \\ 0 & e^{-i\Delta} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \cos\Delta & \sin\Delta \\ -\sin\Delta & \cos\Delta \end{bmatrix}$$

Starting with $x(t) = c\left[\sqrt{2}\ \cos(t)\right] + d\left[\sqrt{2}\ \sin(t)\right] \rightarrow \begin{bmatrix} c \\ d \end{bmatrix}$

$$\widehat{T}_e(\Delta)\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \cos\Delta & \sin\Delta \\ -\sin\Delta & \cos\Delta \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} c\ \cos\Delta + d\ \sin\Delta \\ d\ \cos\Delta - c\ \sin\Delta \end{bmatrix}$$

$$\begin{bmatrix} c\ \cos\Delta - d\ \sin\Delta \\ d\ \cos\Delta + c\ \sin\Delta \end{bmatrix} \rightarrow x(t+\Delta) = \left(c\ \cos\Delta + d\ \sin\Delta\right)\left[\sqrt{2}\ \cos(t)\right] + \left(d\ \cos\Delta - c\ \sin\Delta\right)\left[\sqrt{2}\ \sin(t)\right]$$

$$x(t+\Delta) = c\left[\sqrt{2}\ \left(\cos\Delta\ \cos(t) - \sin\Delta\ \sin(t)\right)\right] + d\left[\sqrt{2}\ \left(\cos\Delta\ \sin(t) + \sin\Delta\ \cos(t)\right)\right]$$

$$x(t+\Delta) = c\left[\sqrt{2}\ \cos(t+\Delta)\right] + d\left[\sqrt{2}\ \sin(t+\Delta)\right]$$

It works in the standard basis as well.

What could have been learned?

A function of a matrix is computed by substituting the matrix into the power series expansion of the function. Powers of diagonal matrices are easy to compute so it is easiest to transform to a basis in which the operator matrix is diagonal. After the function of a matrix is computed, the result can be transformed to another basis using the canonical method, a similarity transformation.

Additional transformations and their consequences:

Observers should be able to use the same form of the laws of physics even after more general transformations than rotations. The Principle of Galilean Invariance states that any inertial observer can make valid observations. That is: any observer traveling at a uniform velocity. It is also assumed that information can be transmitted at infinite speed. [This is our view for this section – a view that is modified by Special Relativity.] With signal propagation at infinite speed, one can synchronize clocks instantaneously over all space. In Galilean relative all observers' clocks agree keeping exactly the same time.

Coordinate Transformation Equations for Galilean Relativity

Assume that the origins of the primed and unprimed inertial systems coincided at t=0 and that the primed observer moves at velocity $\vec{u}$ relative to the unprimed observer. As both systems are moving at uniform velocities (as they are inertial), their relative velocity $\vec{u}$ is also uniform.

$$t' = t; \quad \vec{r}' = \vec{r} - \vec{u}\ t$$

Derived transformations: (take time derivatives of $\vec{r}' = \vec{r} - \vec{u}\ t$)

velocity: $\vec{v}' = \vec{v} - \vec{u}$

acceleration: $\vec{a}' = \vec{a}$

relative velocity: $\vec{v}_2' - \vec{v}_1' = (\vec{v}_2 - \vec{u}) - (\vec{v}_1 - \vec{u}) = \vec{v}_2 - \vec{v}_1$

Velocities transform according to a rule ($\vec{v}' = \vec{v} - \vec{u}$) while accelerations, relative velocities, masses, time and forces are Galilean invariants.

What does Galilean relativity say about possible physical laws? One popular misstatement is that:

$$\frac{d\vec{p}}{dt} = \frac{d}{dt}\left(m\vec{v}\right) = \frac{dm}{dt}\,\vec{v} + m\,\frac{d\vec{v}}{dt} = \vec{F}_{net}$$

The force $\vec{F}_{net}$, the mass m and the acceleration $\frac{d\vec{v}}{dt}$ are Galilean invariants. The factor $\frac{dm}{dt}$ (combination of mass and time) would be a Galilean invariant if it meant anything. All this invariance would require that the only other factor $\vec{v}$ also be invariant if the equation were to be meaningful. The transformation rule for velocities states that velocities are not invariant under Galilean transformations.

In classical (pre-relativistic) physics, the mass of something is a constant. One can have a rail car loaded with coal and throw the coal out of the car. The mass of the car remains the same as does the mass of the coal. Mass moves around, it does not change. The factor $\frac{dm}{dt}$ does not mean that the mass of something is changing, rather it is used to represent the rate at which mass is transferred into (or out of) an open system. Consider the rocket equation.

$$m\,\frac{d\vec{v}}{dt} + \left|\frac{dm}{dt}\right|\vec{v}_{exhaust} = \vec{F}_{net} \quad \text{or} \quad m\,\frac{d\vec{v}}{dt} = -\left|\frac{dm}{dt}\right|\vec{v}_{exhaust} + \vec{F}_{net}$$

Note that $\vec{v}_{exhaust}$ appears rather than $\vec{v}$, the velocity of the rocket. The velocity $\vec{v}_{exhaust}$ is the velocity of the exhausted gases relative to the rocket; $\vec{v}_{exhaust}$ is a relative velocity and hence a Galilean invariant. Therefore it can appear in a valid equation in which all terms and factors are invariant. Note that $\left|\frac{dm}{dt}\right|$ is the rate at which fuel plus oxidizer is expelled from the rocket. Is the mass of any ***thing*** actually changing?

Some of the statements and conclusions of the previous paragraph must be modified when special relativity is considered. The point was to show that transformation invariance could provide restrictions on the form of terms that can appear in physical laws. In general the equations representing laws need not be invariant under transformation. The must be covariant. That is: the two sides of the equation must vary in the same fashion such that the equation remains valid.

Special relativity declares that there is a maximum speed at which information can be transferred leading to different final conclusions about hr transformation properties of physical quantities. Nonetheless, the conclusions above are useful as approximations in cases in which all speeds are small compared to the speed of light.

Exercise: In the Galilean context of the discussion just above, compare and contrast the situations in which an open topped coal car of mass 10,000 *kg* is rolling down a straight, level frictionless track with initial speed $v = 10\ {}^{m}/_{s}$ at $t = 0\ s$. In the first situation rain falls directly downward with no horizontal component of velocity relative to the earth and fills the car at the rate of $10\ {}^{kg}/_{s}$. What happens to the car's speed? Develop an expression for $v(t)$. In the second situation, the 10,000 *kg* car is filled with 10,000 *kg* of water and rolling down the track at $v = 10\ {}^{m}/_{s}$ when a vertical drain pipe is opened on the underside of the car allowing the water to flow out at $10\ {}^{kg}/_{s}$. How does the car's speed vary with time? What is the horizontal component of velocity of the water that has just exited the pipe? Describe the location and orientation of a drain pipe that could be used to jet (rocket) the car down the track.

The sad story of coupled oscillators and diagonalization

The coupled oscillator problem requires that two real symmetric matrices, $\mathbb{M}$ and $\mathbb{K}$, be diagonalized. The transformation to ***simultaneously*** diagonalize the two matrices is more complicated that a rotation. It may be decomposed into three steps: a rotation to diagonalize $\mathbb{M}$; a set of scalings to each coordinate to render a diagonalized $\mathbb{M}$ as a multiple of the identity and a final rotation to diagonalize whatever $\mathbb{K}$ has become after the first two steps. As $\mathbb{M}$ is a multiple of the identity after the first two steps, the final similarity transformation does not disturb it. $\mathbb{K}$ was initially real symmetric, and it maintains its real symmetric form after the first two steps so it is guaranteed that a

final rotation can render it diagonal. One of the problems is to prove that a similarity transformation of a real symmetric matrix that uses an orthogonal matrix preserves the symmetry of the matrix. The proof that a scaling preserves it as well is not difficult.

The sad part of the story is that the scaling alters the inner product such that the components of the eigenvectors must be variably weighted. This issue is discussed in the coupled oscillator handout where the mass matrix is proposed as the metric for the space.

$$\langle I | J \rangle = \begin{bmatrix} a_i & b_i & c_i \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} a_j \\ b_j \\ c_j \end{bmatrix} \text{ where } \mathbb{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

shown as an example for three degrees of freedom. This inner product has the dimensions of *kg-m*^{2}.

Any inner product that meets the defining behaviors may be used. A proposed inner product of two normal modes might be proposed to be:

$$\langle I | J \rangle = \left(\frac{\omega_i \ \omega_j}{2} \right) * \left(\begin{bmatrix} a_i & b_i & c_i \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} a_j \\ b_j \\ cj \end{bmatrix} \right)$$

which has dimensions *Joules*. (Note that the angular frequencies can be chosen positive by convention as it is their squares that arise as eigenvalues.) The extension of this proposal to a inner product of general (rather than just eigenvectors) vectors would need to be developed. That prescription would then need to be test for compliance with the four inner product axioms. If such a procedure could be found, it might be more attractive just because *Joules* are more familiar dimensions than *kg-m*^{2}.

P. K. Aravind, ***Geometrical interpretation of the simultaneous diagonalization of two quadratic forms***, American Journal of Physics (April 1989) **57**, 4, pp. 309-311.

Could it get any worse?!

In the coupled oscillator problem, the kinetic energy and the potential energy are represented by homogeneous quadratic expressions.

$$KE = \tfrac{1}{2}\sum m_{ij}\,\dot{q}_i\,\dot{q}_j \;\; and \;\; PE = \tfrac{1}{2}\sum k_{ij}\,q_i\,q_j$$

leading to the defining of two matrices:

$$\mathbb{M} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & & m_{2n} \\ \vdots & & \ddots & \vdots \\ m_{n1} & \cdots & & m_{nn} \end{bmatrix} \quad \text{and } \mathbb{K} = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} & & k_{2n} \\ \vdots & & \ddots & \vdots \\ k_{n1} & \cdots & & k_{nn} \end{bmatrix}$$

These matrices can be chosen to be real symmetric. Are they physical tensors?

What do we know about rank two tensors? The components transform between frames according to a similarity transformation $\tilde{\tilde{\alpha}}' = \mathbb{S}\;\tilde{\tilde{\alpha}}\;\;\mathbb{S}^{-1}$ where $\mathbb{S}$ is the transformation for the components of a *(vector)* displacement. A rank two tensor has [*dimension*]^2 components. All components share the same units.

Form the generalized coordinate vectors

$$\tilde{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} \quad ; \quad \tilde{q}^{\,t} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$$

When transforming coordinates, the obey the general linear form

$$\tilde{q}' = \mathbb{T}\;\tilde{q} \;\; \text{and } (\tilde{q}')^t = (\mathbb{T}\;\tilde{q})^t = \tilde{q}^{\,t}\,\mathbb{T}^t$$

where $\mathbb{T}$ is and the *n x n* matrix representing the transformation. The vector $\dot{\tilde{q}}$ follows the same transformation rule for time independent transformations. Next, recall that *PE* and *KE* are physical scalars.

$$PE = {}^1/_2\;\tilde{q}^{\,t}\,\mathbb{K}\,\tilde{q} = {}^1/_2\;(\tilde{q}')^t\,\mathbb{K}'\,\tilde{q}' \;\; \text{scalar values are invariant}$$

Use the transformation rules.

$${}^1/_2\;\tilde{q}^{\,t}\mathbb{K}\,\tilde{q} = {}^1/_2\;(\tilde{q}')^t\,\mathbb{K}'\,\tilde{q}' = {}^1/_2\,\tilde{q}^{\,t}\,\mathbb{T}^t\,\mathbb{K}'\,\mathbb{T}\,\tilde{q}$$

Hence

$\mathbb{K} = \mathbb{T}^t \mathbb{K}^{/} \mathbb{T}$ (the $\mathbb{T}^t$ -object- $\mathbb{T}$ form is called a *congruent* transformation)

This rule only represents a similarity transformation if $\mathbb{T}^t = \mathbb{T}^{-1}$ which is not true for many coupled oscillator problems. Yikes! It appears that $\mathbb{K}$ and $\mathbb{M}$ fail the '*am I a tensor?'* test.

Sure they do. A tensor has [*dimension*] ^{2} components. A tensor is a collection of components with the same dimension describing one thing. The generalized coordinate position vector does not transform like *displacement vector* so all other bets are off. The matrices $\mathbb{K}$ and $\mathbb{M}$ are just that: *matrices of physical interest, not physical tensors with standard, good transformation properties.* (Review the arguments that established that the inertia tensor obeyed the correct transformation rules. Count the elements and check that their dimensions agree.)

The congruent transformations for the coupled-oscillator problem have the form: $\mathbb{C} = \lambda_2 \ \mathbb{S} \ \lambda_1$ where λ_1 is a rotation that diagonalizes $\mathbb{M}$, $\mathbb{S}$ is a diagonal scaling ($S_{ij} = s_i \ \delta_{ij} = c \ m^{/}_{jj}$) that maps the diagonalized $\mathbb{M}=\lambda_1 \ \mathbb{M} \ (\lambda_1)^{\mathbf{t}}$ to a multiple of the identity $c \ \mathbb{1}$, and λ_2 is a rotation to diagonalize $\lambda_2 \ \mathbb{S} \ \lambda_1 \mathbb{K} \ (\lambda_2 \ \mathbb{S} \ \lambda_1)^{\mathbf{t}}$ (which is still real symmetric). In this transformed space, the eigenvectors have the standard metric, a multiple of $\mathbb{1}$. The transformed $\mathbb{M} = \lambda_2 \ \mathbb{S} \ \lambda_1 \mathbb{M} \ (\lambda_2 \ \mathbb{S} \ \lambda_1)^{\mathbf{t}}$ is also a multiple of $\mathbb{1}$, so, after the inverse transformation, $\mathbb{M}$ can be used as the metric in the original basis of generalized coordinates.

Thoughts under devlopment !!!!! DO NOT READ. TO ADD:
transformations in abstract spaces
s-p hybrid orbitals
sp2 Use the picture of a triad and require equal s character in each => all have 1/roo(3) for S. One with S and 2/3 in one P 1/root6 [root2,2,0]
other 1/root6[root2,-1, +- root3]
sp3 equal parts of S => amplitude 1/2 easy answers
*(1/2) * { [1,1,1,1]; [1,1,-1,-1]; [1,-1,-1,1]; [1,-1,1,-1] }*

Unitary and Hermitian

Recap mapping vs change of basis: An operator mapping changes or associates each vector in the space into another vector in the space or perhaps in another space. Under a change of basis transformation, nothing physics changes only the representations change.