Astronomy – General Relativity and Gravitation I
Part 2: Vectors, tensors and functions

Norman Gray

Public version 1.1, June 2013

At this point we take a holiday from the physics, in favour of mathematical preliminaries. This part is concerned with carefully defining vectors, tensors and functions, and showing how they are linked with the notion of coordinate systems. This will take us to the point where, in the next part, we can talk about doing calculus with these objects.

Aims You should

1. understand the links and distinctions between the concepts of *vector space*, *tensor*, *vector*, *one-form* and *function*;

2. be able to use the component-based notation for vectors, one-forms and tensors.

Objectives You should be able to demonstrate that you can

1. quote the definition of a $\binom{M}{N}$ tensor;

2. identify the rank $\binom{M}{N}$ of a tensor (or vector or one-form) from expressions such as $\mathsf{T}(\widetilde{p}, \widetilde{q};\ \cdot\ )$ for the partially-contracted tensor, or expressions such as $T^{ij}{}_{k}$ for its components;

3. perform calculations using component notation and the einstein summation convention; for example…

4. correctly use a given transformation matrix $\Lambda^{\bar{\imath}}{}_{j}$ or $\Lambda^{i}{}_{\bar{\jmath}}$ to transform components from one basis to another.

1 Linear algebra

See Schutz, appendix A.

The material in this section will probably be, if not familiar, at least recognisable to you, though possibly with new notation.

Here, and elsewhere in this course, the idea of *linearity* is of crucial importance; it is a very simple idea, however. Consider a function (or operator or other object) f, objects $\mathbf{x}$ and $\mathbf{y}$ in the domain of f, and numbers $\{a, b\} \in \mathbb{R}$: if $f(a\mathbf{x} + b\mathbf{y}) = af(\mathbf{x}) + bf(\mathbf{y})$, then the function f is said to be *linear*. Thus the function $f = ax$ is linear in x, but $f = ax + b$, $f = ax^2$ and $f = \sin x$ are not; matrix multiplication is linear in the (matrix) arguments, but the rotation of a solid sphere (say) is not linear in the Euler angles (note that although you might refer to $f(x) = ax + b$ as a 'straight line graph', or might refer to it as linear in other contexts, in this formal sense it is *not* a linear function, because $f(2x) \neq 2f(x)$).

1.1 Vector spaces

A set of objects V is called a *vector space* if it satisfies the following axioms (for $A, B \in V$ and $a \in \mathbb{R}$):

vecto

1. Closure: there is a binary operator '+', such that $A + B = B + A \in V$.
2. Identity: there exists an element $0 \in V$, such that $A + 0 = A$.
3. Inverse: for every $A \in V$, there exists an element $B \in V$ such that $A + B = 0$ (incidentally, these first three properties together mean that V is classified as an *abelian group*).
4. Multiplication by reals: for all a and all A, $aA \in V$ and $1A = A$.
5. Distributive: $a(A + B) = aA + aB$.

The obvious example of a vector space is the set of vectors that you learned about in school, but crucially, *anything* which satisfies these axioms is also a vector space.

Ex.2

Vectors $A_1, \ldots, A_n$ are *linearly independent* (LI) if $a_1 A_1 + a_2 A_2 + \cdots + a_n A_n = 0$ implies $a_i = 0, \forall i$. The *dimension* of a vector space n is the largest number of LI vectors which can be found. A set of n LI vectors A_i in an n-dimensional space is said to *span* the space, and is termed a *basis* for the space. Then is it a theorem that, for every vector $B \in V$, there exists a set of numbers $\{b_i\}$ such that $B = \sum b_i A_i$; these numbers $\{b_i\}$ are the *components* of the vector B with respect to the basis $\{A_i\}$.

linea
pend
dime
span
basis
comp

One can (but need not) define an *inner product* on a vector space: the inner product between two vectors A and B is written $A \cdot B$ (yes, the dot-product that you know about is indeed an example of an inner product; also note that the inner product is sometimes written $\langle A, B \rangle$, but we will reserve that notation, here, to the contraction between a vector and a one-form, defined below). This is a symmetric, linear, operator which maps pairs of vectors to the real line. That is (i) $A \cdot B = B \cdot A$, and (ii) $(aA + bB) \cdot C = aA \cdot C + bB \cdot C$. Two vectors,

inner

A and B, are *orthogonal* if $A \cdot B = 0$. An inner-product is *positive-definite* if $A \cdot A > 0$ for all $A \neq 0$, or indefinite otherwise. The *norm* of a vector A is $|A| = |A \cdot A|^{1/2}$.The symbol δ_{ij} is the *Kronecker delta symbol*, defined as

$$\delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

(below, we will use variants of this symbol with indexes raised or lowered – they mean the same: $\delta_{ij} = \delta^i{}_j = \delta_i{}^j$). A set of vectors $\{e_i\}$ such that $e_i \cdot e_j = \delta_{ij}$ (that is, all orthogonal and with unit norm) is an *orthonormal* basis. It is a theorem that, if $\{b_i\}$ are the components of an arbitrary vector B in this basis, then $b_i = B \cdot e_i$.

ortho
norm
posit
Kron
defin
symb
o

1.2 Matrix algebra

An $m \times n$ matrix A is a purely mathematical object which can can be represented by a set of *elements* denoted A_{ij}, via

$$\mathsf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}.$$

You know how to define addition of two $m \times n$ matrices, and multiplication of a matrix by a scalar, and that the result in both cases is another matrix, so the set of $m \times n$ matrices is another example of a vector space. You also know how to define matrix multiplication: a vector space with multiplication defined is an *algebra*, so what we are now discussing is matrix algebra.

A square matrix (that is, $n \times n$) may have an inverse, written A^{-1}, such that $\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = 1$ (one can define left- and right-inverses of non-square matrices, but they will not concern us). The unit matrix $\mathbb{1}$ has elements δ_{ij}. You can define the *determinant* and *trace* of a square matrix.

Make sure that you are in fact familiar with the matrix concepts in this section.

2 Tensors, vectors and one-forms

Most of the rest of this course is going to be talking about tensors one way or another, so we had better grow to love them now. See Schutz, chapter 3.

We are going to introduce tensors in a rather abstract way here; this is because, to me, this emphasises that they are in fact rather simple objects. They will become more specific when we introduce tensor components shortly, and in the rest of the course we will use these components extensively, but introducing tensor components first of all can hide the geometrical primitiveness of tensors.

2.1 Definition of tensors

For each $M, N = 0, 1, 2, \ldots$, the set of $\binom{M}{N}$ tensors is a set which obeys the axioms of a vector space in Sect. 1.1. Specifically, we refer to a $\binom{1}{0}$ tensor as a *vector* and write it as $\overline{A}$, and a $\binom{0}{1}$ tensor as a *one-form*, written $\widetilde{A}$. A $\binom{0}{0}$ tensor is just a function which maps $\mathbb{R} \to \mathbb{R}$. The clash with the terminology of Sect. 1.1 is unfortunate (because *all* of these objects are 'vectors' in the terminology of that section), but from now on when we refer to a vector space, we are referring to Sect. 1.1, and when we refer to vectors, we are referring specifically to $\binom{1}{0}$ tensors.

vecto
one-f

For the moment, you can perfectly correctly think of vectors as exactly the type of vectors you are used to – a direction in space. In part 3, we will introduce a new definition of vectors which is of crucial importance in our development of general relativity.

> *Definition:* A $\binom{M}{N}$ *tensor* is a function, linear in each argument, which takes M one-forms and N vectors as arguments, and maps them to a real number.

tenso

Because we said, above, that an $\binom{M}{N}$ tensor was an element of a vector space, we already know that if we add two $\binom{M}{N}$ tensors, or if we multiply an $\binom{M}{N}$ tensor by a scalar, then we get another $\binom{M}{N}$ tensor. This definition does seem very abstract, but most of the properties we are about to deduce follow directly from it.

For example, we can write the $\binom{2}{1}$ tensor T as

$$\mathsf{T}(\widetilde{\cdot}, \widetilde{\cdot}; \overline{\cdot}),$$

to emphasise that the function has two 'slots' for one-forms and one 'slot' for a vector. When we insert one-forms $\widetilde{p}$ and $\widetilde{q}$, and vector $\overline{A}$, we get $\mathsf{T}(\widetilde{p}, \widetilde{q}; \overline{A})$ which, by our definition of a tensor, we see must be a pure number, in $\mathbb{R}$. [Note that this 'dots' notation is an informal one, and though I have chosen to write this, below, with one-form arguments all to the left of vector ones, this is just for the sake of clarity: in general, the $\binom{1}{1}$ tensor $\mathsf{T}(\overline{\cdot}, \widetilde{\cdot})$ is a perfectly good tensor, and distinct from the $\binom{1}{1}$ tensor $\mathsf{T}(\widetilde{\cdot}, \overline{\cdot})$.]

Note firstly that there is nothing in the definition of a tensor which states that the arguments are interchangeable, thus, in the case of a $\binom{0}{2}$ tensor $\mathsf{U}(\overline{\cdot}, \overline{\cdot})$, $\mathsf{U}(\overline{A}, \overline{B}) \neq \mathsf{U}(\overline{B}, \overline{A})$ in general: if in fact $\mathsf{U}(\overline{A}, \overline{B}) = \mathsf{U}(\overline{B}, \overline{A}), \forall \overline{A}, \overline{B}$, then U is said to be *symmetric*; and if $\mathsf{U}(\overline{A}, \overline{B}) = -\mathsf{U}(\overline{B}, \overline{A}), \forall \overline{A}, \overline{B}$, it is *antisymmetric*.

Note also, that if we insert only some of the arguments into the tensor T,

$$\mathsf{T}(\widetilde{\omega}, \widetilde{\cdot}; \overline{\cdot}),$$

then we obtain an object which can take a single one-form and a single vector, and map them into a number; in other words, we have a $\binom{1}{1}$ tensor.

Note finally, that although we said that a vector was just the thing that you are used to thinking of as a vector, since a vector is 'really' a $\binom{1}{0}$ tensor, it's 'really' just a function which can map a one-form into a number; similarly, a one-form, as a $\binom{0}{1}$ tensor, is just a thing which can map a vector into a number. For our purposes in GR, we will restrict these functions so that, for arbitrary vector $\overline{A}$ and one-form $\widetilde{p}$,

$$\overline{A}(\widetilde{p}) = \widetilde{p}(\overline{A}) \equiv \left\langle \widetilde{p}, \overline{A} \right\rangle, \quad \forall \widetilde{p}, \overline{A} \qquad \text{[in GR]} \tag{2.1}$$

where the notation $\langle \cdot, \cdot \rangle$ emphasises the symmetry of this operation.

This combination of the two objects is known as the *contraction* of $\widetilde{p}$ with $\overline{A}$.

2.2 Examples of tensors

This description of tensors is very abstract, so we need some examples fairly promptly.

The most immediate example of a vector is the column vector you are familiar with, and the one forms which correspond to it are simply row-vectors.

$$\widetilde{p} = (p_1, p_2), \qquad \overline{A} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \qquad \langle \widetilde{p}, \overline{A} \rangle = (p_1, p_2) \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = p_1 A_1 + p_2 A_2.$$

Here we see the one-form $\widetilde{p}$ and vector $\overline{A}$ contracting to form a number. Or we can see $\widetilde{p}$ as a real-valued function over vectors, mapping them to numbers, and similarly $\overline{A}$, a real-valued function over one-forms.

How about tensors of higher rank? Easy: a square matrix

$$\mathsf{T} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is a function which takes one one-form and one vector, and maps them to a number, which it to say it is a $\binom{1}{1}$ tensor. If we supply only one of the arguments, to get $\mathsf{T}\overline{A}$, we get an object which has a single one-form argument, which is to say, another vector.

Can we form tensors of other ranks? Yes, though with a little more difficulty. If we have vectors $\overline{V}$ and $\overline{W}$, then we can form a $\binom{2}{0}$ object $\overline{V} \otimes \overline{W}$, whose value on the one-forms $\widetilde{p}$ and $\widetilde{q}$ is defined to be

$$(\overline{V} \otimes \overline{W})(\widetilde{p}, \widetilde{q}) \equiv \overline{V}(\widetilde{p}) \times \overline{W}(\widetilde{q}).$$

This object $\overline{V} \otimes \overline{W}$ is known as the *outer product* of the two vectors (or sometimes the direct product or tensor product); see Schutz, section 3.4. For example the object outer

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \otimes \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

is a $\binom{2}{0}$ object whose value on the two one-forms $\widetilde{p}$ and $\widetilde{q}$ is $(p_1 A_1 + p_2 A_2) \times (q_1 B_1 + q_2 B_2)$. In a similar way, we can use the outer product to form objects of other ranks from suitable combinations of vectors and one-forms. Not all tensors are necessarily outer products, though all tensors can be represented as a sum of outer products.

2.3 Fields

We will often want to refer to a *field* of objects. A field is just a function, in the sense that it maps one space to another, but in this course we restrict the term 'field' to the case where the domain is a physical space, or spacetime. That is, a field is a rule which associates a number, or some higher-rank tensor, with each point in space, or in spacetime. Air-pressure is an example of a scalar field (each point in 3-d space has a number associated with it), and the electric and magnetic fields, **E** and **B**, are vector fields (associating a vector with each point in 3-d space). field

2.4 Visualisation of vectors and one-forms

We can visualise vectors straightforwardly as arrows, having both a magnitude and a direction. In order to combine one vector with another, however, we need to add further rules, defining something like the dot product and thus – as we will soon learn – introducing concepts such as the metric (Sect. 2.6 below).

How do we visualise one-forms in such a way that we distinguish them from vectors, and in such a way that we can visualise (metric-free) operations such as the contraction of a vector and a one-form?

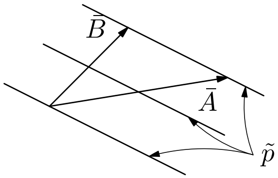

Figure 1

The most common way is to visualise a one-form as a set of planes in the appropriate space. Such a structure picks out a direction – the direction perpendicular to the planes – and a magnitude which increases as the separation between the planes *decreases*. The contraction between a vector and a one-form thus visualised is the number of the one-form planes which the vector crosses.

In Fig. 1, we see two different vectors and one one-form, $\widetilde{p}$. Although the two vectors are of different lengths (though we don't 'know' this yet, since we haven't yet talked about a metric and thus have no notion of 'length'), their contraction with the one-form is the same, 2.

You are already familiar with this picture, since you are familiar with the notion of contours on a map. These show the gradient of the surface they are mapping, with the property that the closer the coutours are together, the larger is the gradient. The three vectors shown in

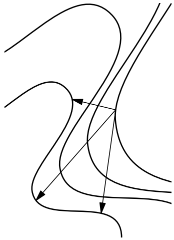

Figure 2

Fig. 2, which might be different paths up the hillside, have the same contraction – the path climbs three units – even though the three vectors have rather different lengths. When we look at the contours, we are seeing a one-form *field*, with the one-form having different values, both magnitude and direction, at different points in the space.

We will see below, in Sect. 1.3 of part 3, that the natural defintion of the gradient of a function does indeed turn out to be a one-form.

Armed with this visualisation, it is natural to talk of $\overline{A}$ and $\widetilde{p}$ as *geometrical objects*. When we do so, we are stressing the difference between these abstract objects and their numerical components with respect to a basis. This is what we meant when we talked, in Sect. 2.3 of part 1, about physical laws depending only on geometrical objects, and not on their components with respect to a set of basis vectors which we introduce only for our mensural convenience.

2.5 Components

We said, above, that the set of $\binom{M}{N}$ tensors formed a vector space. Specifically, that includes the sets of vectors and one-forms. From Sect. 1.1, this means that we can find a set of *n basis vectors* $\{\overline{e}_i\}$ and basis one-forms $\{\widetilde{\omega}^i\}$ (this is supposing that the domains of the arguments to our tensors all have the same dimensionality, n; this is not a fundamental property of tensors, but it is true in all the use we make of them, and so this avoids unnecessary complication). Note that we have written the basis vectors with lowered indexes and the basis one-forms with raised ones; this is related to the Einstein summation convention described below. ba

Armed with a set of basis vectors and one-forms, we can write a vector $\overline{A}$ and one-form $\widetilde{p}$ in components as

$$\overline{A} = \sum_i A^i \overline{e}_i; \qquad \widetilde{p} = \sum_i p_i \widetilde{\omega}^i.$$

Crucially, these components are *not* intrinsic to the geometrical objects which $\overline{A}$ and $\widetilde{p}$ represent, but instead depend on the vector or one-form basis which we select. *It is absolutely vital* that you fully appreciate that if you change the basis, you change the components of a vector or one-form (or any tensor) with respect to that basis, *but* the underlying geometrical object, $\overline{A}$ or $\widetilde{p}$ (or T), *does not change*. Though this remark seems obvious now, dealing with it in general is what much of the complication of differential geometry is all about.

Note the (purely conventional) positions of the indexes for these basis vectors and one-forms, and for the components: the components of vectors have raised indexes, and the components of one-forms have lowered indexes. This convention allows us to define an extremely useful notational shortcut, which allows us in turn to avoid writing hundreds of summation signs:

> *Einstein summation convention:* whenever we see an index repeated in an expression, once raised and once lowered, we are to understand a summation over that index.

Thus:

$$A^i \overline{e}_i \equiv \sum_i A^i \overline{e}_i; \qquad p_i \widetilde{\omega}^i \equiv \sum_i p_i \widetilde{\omega}^i.$$

We have illustrated this for components and vectors here, but it will apply quite generally below.

Rules for working with components:

1. In any expression, there must be at most two of each index, one raised and one lowered. If you have more than two, or have both raised or lowered, you've made a mistake. Any indexes 'left over' after this contraction tell you the rank of the object which this is the component of.

2. The components *are just numbers*, and so, as you learned in primary school, it doesn't matter what order you multiply them in (don't swap them past differential signs, though!).

3. The indexes are arbitrary – you can always replace an index letter with another one, as long as you do it consistently. That is, $p_i A^i = A^j p_j$, and $p_i q_j T^{ij} = p_j q_i T^{ji} = p_k q_i T^{ki}$ (though $p_k q_i T^{ki} \neq p_k q_i T^{ik}$ in general, unless the tensor T is symmetric).

What happens if we apply $\widetilde{p}$, say, to one of the basis vectors? We have

$$\widetilde{p}(\overline{e}_j) = p_i \widetilde{\omega}^i(\overline{e}_j). \tag{2.2}$$

In principle, we know nothing about the *number* $\widetilde{\omega}^i(\overline{e}_j)$, since we are at liberty to make completely independent choices of the vector and one-form bases. However, we can save ourselves ridiculous amounts of trouble by making a wise choice, and we will always choose one-form bases to have the property

$$\langle \widetilde{\omega}^i, \overline{e}_j \rangle = \widetilde{\omega}^i(\overline{e}_j) = \delta^i{}_j \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.3a}$$

$$\overline{e}_j(\widetilde{\omega}^i) = \delta_j{}^i. \tag{2.3b}$$

A one-form basis with this property is said to be *dual* to the vector basis. Returning to Eq. (2.2), therefore, we find

$$\widetilde{p}(\overline{e}_j) = p_i \widetilde{\omega}^i(\overline{e}_j) = p_i \delta^i{}_j = p_j. \tag{2.4}$$

Thus in the one-form basis which is dual to the vector basis $\{\overline{e}_j\}$, the arbitrary one-form $\widetilde{p}$ has components $p_j = \widetilde{p}(\overline{e}_j)$.

Similarly, we can apply the vector $\overline{A}$ to the one-form basis $\{\widetilde{\omega}^i\}$, and obtain

$$\overline{A}(\widetilde{\omega}^i) = A^j \overline{e}_j(\widetilde{\omega}^i) = A^j \delta_j{}^i = A^i.$$

In exactly the same way, we can apply the tensor T to the basis vectors and one-forms, and obtain

$$\mathsf{T}(\widetilde{\omega}^i, \widetilde{\omega}^j; \overline{e}_k) = T^{ij}{}_k. \tag{2.5}$$

The set of $n \times n \times n$ *numbers* $\{T^{ij}{}_k\}$ are the *components* of the tensor T in the basis $\{\overline{e}_i\}$ and its dual $\{\widetilde{\omega}^j\}$. Below, we will often denote the vector $\overline{A}$ by simply writing 'A^i', denote $\widetilde{p}$ by 'p_i', and the $\binom{2}{1}$ tensor T by '$T^{ij}{}_k$'. Because of the index convention, we will always know what sort of object we are referring to by whether the indexes are raised or lowered: (the components of) vectors always have their indexes raised, and (the components of) one-forms always have their indexes lowered.

We can, obviously, find the components of the basis vectors and one-forms by exactly this method, and find

$$\begin{aligned}
\overline{e}_1 &\to (1, 0, \ldots, 0) \\
\overline{e}_2 &\to (0, 1, \ldots, 0) \\
&\vdots \\
\overline{e}_n &\to (0, 0, \ldots, 1)
\end{aligned} \tag{2.6}$$

where the numbers on the right-hand-side are the components in the vector basis, and

$$\begin{aligned}
\widetilde{\omega}^1 &\to (1, 0, \ldots, 0) \\
\widetilde{\omega}^2 &\to (0, 1, \ldots, 0) \\
&\vdots \\
\widetilde{\omega}^n &\to (0, 0, \ldots, 1)
\end{aligned} \tag{2.7}$$

where the components are in the one-form basis. Make sure you understand why this is 'obvious'.

So what is the value of the expression $\widetilde{p}(\overline{A})$ in components? By linearity,

$$\widetilde{p}(\overline{A}) = p_i \widetilde{\omega}^i (A^j \overline{e}_j) = p_i A^j \widetilde{\omega}^i (\overline{e}_j) = p_i A^j \delta^i{}_j = p_i A^i.$$

This is the *contraction* of $\widetilde{p}$ with $\overline{A}$. Note particularly that, since $\widetilde{p}$ and $\overline{A}$ are basis-independent, geometrical objects – or quite separately, since $\widetilde{p}(\overline{A})$ is a pure number – the *number* $p_i A^i$ is basis-independent also, even though the numbers p_i and A^i are separately basis-dependent.

Similarly, contracting the $\binom{2}{1}$ tensor T with one-forms $\widetilde{p}, \widetilde{q}$ and vector $\overline{A}$, we obtain the number

$$\mathsf{T}(\widetilde{p}, \widetilde{q}; \overline{A}) = p_i q_j A^k T^{ij}{}_k.$$

If we contract it instead with just the one-forms, we obtain the object $\mathsf{T}(\widetilde{p}, \widetilde{q}; \overline{\cdot})$, which is a one-form (since it maps a single vector to a number) with components

$$\mathsf{T}(\widetilde{p}, \widetilde{q}; \overline{\cdot})_k = p_i q_j T^{ij}{}_k$$

and the solitary unmatched lower index k on the right-hand-side indicates (or rather confirms) that this object is a one-form. The indexes are staggered so that we keep track of which argument they correspond to. The two tensors $\mathsf{T}(\overline{\cdot}, \widetilde{\cdot})$ and $\mathsf{T}(\widetilde{\cdot}, \overline{\cdot})$ are different tensors: it *may* be the case that $\mathsf{T}(\overline{e}_i, \widetilde{\omega}^j) = \mathsf{T}(\widetilde{\omega}^j, \overline{e}_i)$, and thus that $T_i{}^j = T^j{}_i$, but we cannot simply assume this.

2.6 The metric tensor

One thing we do not have yet is any notion of distance, but we can supply that very easily, by picking a symmetric $\binom{0}{2}$ tensor g, and calling that the *metric tensor*. metri

This definition allows us to define a 'distance function' on two vectors, $d(\overline{A}, \overline{B})$, as simply $d(\overline{A}, \overline{B}) = \mathsf{g}(\overline{A}, \overline{B})$, and a 'length-squared' of a vector $d(\overline{A}, \overline{A})$. This notion is hugely important to us when we study GR, and we will return to it in due course. We can (and in this course will) usefully define the inner product on the set of vectors using the metric, as simply $\overline{A} \cdot \overline{B} = \mathsf{g}(\overline{A}, \overline{B})$.

We can find the components of the metric tensor in the same way we can find the components of any tensor above:

$$\mathsf{g}(\overline{e}_i, \overline{e}_j) = g_{ij}. \tag{2.8}$$

As well as giving us a notion of distance, the metric tensor allows us to define a mapping between vectors and one-forms. Since it is a $\binom{0}{2}$ tensor, it is a thing which takes two vectors and turns them into a number. If we only supply a single vector $\overline{A} = A^i \overline{e}_i$ to the metric, we have a thing which takes one further vector and turns it into a number; but this is just a one-form, which we will write as $\widetilde{A}$:

$$\widetilde{A} = \mathsf{g}(\overline{A}, \cdot\,) = \mathsf{g}(\,\cdot\,, \overline{A}).$$

That is, for any vector $\overline{A}$, we have found a way of picking out a single associated one-form, written $\widetilde{A}$. What are the components of this one-form? Easy:

$$\begin{aligned} A_i = \widetilde{A}(\overline{e}_i) &= \mathsf{g}(\overline{e}_i, \overline{A}) && \text{(2.9a)}\\ &= \mathsf{g}(\overline{e}_i, A^j \overline{e}_j) && \text{(2.9b)}\\ &= A^j \mathsf{g}(\overline{e}_i, \overline{e}_j) && \text{(2.9c)}\\ &= g_{ij} A^j, && \text{(2.9d)} \end{aligned}$$

from Eq. (2.8) above. That is, the metric tensor can also be regarded as an '*index lowering*' operator. Ex.2 index

Can we do this trick in the other direction, defining a $\binom{2}{0}$ tensor which takes two one-forms as arguments and turns them into a number? Yes we can, and the natural way to do it is via the tensor's components.

The set of numbers g_{ij} is, at one level, just a matrix. Thus if it is non-singular (and we will always assume that the metric is non-singular), this matrix has an inverse, and we can take the components of the tensor we're looking for, g^{ij}, to be the components of this inverse. That means nothing other than

$$g^{ij} g_{jk} = \delta^i{}_k, \tag{2.10}$$

which immediately defines the components of the tensor we're looking for. We will refer to the tensors corresponding to g^{ij}, $g^i{}_j$ and g_{ij} indiscriminately as 'the metric'. Note that, since the effect of Eq. (2.10) is simply to lower one index of the first g^{ij}, by Eq. (2.9a), it is *always* true that

$$g^i{}_j = \delta^i{}_j.$$

What happens if we apply g^{ij} to the one-form components A_j?

$$g^{ij} A_j = g^{ij} g_{jk} A^k = \delta^i{}_k A^k = A^i, \tag{2.11}$$

so that the metric can raise components as well as lower them.

There is nothing in the above discussion that says that the tensor g has the same value at each point in spacetime. In general, g is a tensor field, and the different values of the metric at different points in spacetime are associated with the curvature of that spacetime. This is where the physics comes in.

2.7 Changing basis

The last very general set of properties we must discover about tensors is what happens when you change your mind about the sets of basis vectors and one-forms (you can't change your mind about just one of them, if they are to remain dual to each other).

We (now) know that if we have a set of basis vectors $\{\overline{e}_i\}$, then we can find the components of an arbitrary vector $\overline{A}$ to be $A^i = \overline{A}(\widetilde{\omega}^i)$, where the $\{\widetilde{\omega}^i\}$ are the set of basis one-forms which are dual to the vectors $\{\overline{e}_i\}$.

But there is nothing special about this basis, and we could equally well have chosen a completely different set $\{\overline{e}_{\bar{j}}\}$ with dual $\{\widetilde{\omega}^{\bar{j}}\}$ (note that we distinguish the vectors by putting a bar on the index j rather than the base symbol $\overline{e}$ – this does look odd, I know, but it ends up being notationally tidier). With respect to this basis, the *same* vector $\overline{A}$ can be written

$$\overline{A} = A^{\bar{i}}\overline{e}_{\bar{i}},$$

where, of course, the components are the set of numbers

$$A^{\bar{i}} = \overline{A}(\widetilde{\omega}^{\bar{i}}).$$

Since both these sets of components represent the *same* underlying object $\overline{A}$, we naturally expect that they are related to each other, and it is easy to write down that relation. From before

$$\begin{aligned} A^{\bar{i}} &= \overline{A}(\widetilde{\omega}^{\bar{i}}) \\ &= A^i\overline{e}_i(\widetilde{\omega}^{\bar{i}}) \\ &= \Lambda^{\bar{i}}_{i} A^i, \end{aligned} \tag{2.12}$$

where we have written the *transformation matrix* Λ as trans

matri

$$\Lambda^{\bar{i}}_{i} \equiv \overline{e}_i(\widetilde{\omega}^{\bar{i}}) \equiv \widetilde{\omega}^{\bar{i}}(\overline{e}_i). \tag{2.13}$$

Note that Λ is a *matrix*, not a tensor – there's no underlying geometrical object, and we have consequently not staggered its indexes. Also, note that indexes i and $\bar{\imath}$ are completely distinct from each other, and arbitrary, and we are using the similarity of symbols just to emphasise the symmetry of the operation. Exactly analogously, the components of a one-form $\widetilde{p}$ transform as Ex.2

$$p_{\bar{\imath}} = \Lambda^{i}_{\bar{\imath}} p_i, \tag{2.14}$$

where the transformation matrix is

$$\Lambda^{i}_{\bar{\imath}} \equiv \widetilde{\omega}^i(\overline{e}_{\bar{\imath}}). \tag{2.15}$$

Since the vector $\overline{A}$ is the same in both coordinate systems, we must have

$$\begin{aligned} \overline{A} = A^i \overline{e}_i &= A^{\bar{\imath}} \overline{e}_{\bar{\imath}} \\ &= \Lambda^{\bar{\imath}}_{\ j} E^k_{\bar{\imath}} A^j \overline{e}_k, \end{aligned} \tag{2.16}$$

where we write $E^k_{\bar{\imath}}$ as the (initially unknown) components of vector $\overline{e}_{\bar{\imath}}$ in the basis $\{\overline{e}_k\}$ (ie $\overline{e}_{\bar{\imath}} = E^k_{\bar{\imath}} \overline{e}_k$). This immediately requires that $\Lambda^{\bar{\imath}}_{\ j} E^k_{\bar{\imath}} = \delta^k_j$, and thus that the matrix E must be the inverse of Λ.

Moving on, the components of our $\binom{2}{1}$ tensor T transform as

$$T^{\bar{\imath}\bar{\jmath}}{}_{\bar{k}} = \Lambda^{\bar{\imath}}_{i} \Lambda^{\bar{\jmath}}_{j} \Lambda^{k}_{\bar{k}} T^{ij}{}_{k}. \tag{2.17}$$

Let us now look at the contraction $\widetilde{\omega}^i(\overline{e}_j)$ in two coordinate systems:

$$\delta^i_j = \widetilde{\omega}^i(\overline{e}_j) = \Lambda^{i}_{\bar{\imath}} \Lambda^{\bar{\jmath}}_{j} \widetilde{\omega}^{\bar{\imath}}(\overline{e}_{\bar{\jmath}}) = \Lambda^{i}_{\bar{\imath}} \Lambda^{\bar{\jmath}}_{j} \delta^{\bar{\imath}}_{\bar{\jmath}} = \Lambda^{i}_{\bar{\imath}} \Lambda^{\bar{\imath}}_{\ j}, \tag{2.18}$$

using Eqns. (2.3a), (2.13) and (2.15), or more specifically the results of example 2.8. That is the matrices $\Lambda^{i}_{\bar{\imath}}$ and $\Lambda^{\bar{\imath}}_{\ j}$ are inverses of each other.

Here, it has been convenient to introduce basis transformations by focusing on the transformation of the *components* of vectors and one-forms, in Eq. (2.12). We could alternatively introduce them by focusing on the transformation of the basis vectors and one-forms themselves, and this is the approach used in the discussion of basis transformations in Sect. 1.4 of part 3. Looking again at Eq. (2.16), we can identify $E^i_{\bar{\imath}}$ with the inverse matrix $\Lambda^i_{\bar{\imath}}$, and thus deduce the basis vector transformation to be

$$\overline{e}_{\bar{\imath}} = \Lambda^i_{\bar{\imath}} \overline{e}_i . \tag{2.19}$$

Ex.2.

2.8 Some misconceptions about coordinates and vectors

This section has a dangerous-bend marker, not because it introduces extra material, but because it should be approached with caution. It addresses some misconceptions about coordinates that have cropped up at various times. The danger of mentioning these, of course, is that if you learn about a misconception that hadn't occurred to you, it could make things worse! Also, I make forward references to material introduced in part 3, so this is a section which is probably more useful at revision time, than on your first read through. That said, if you're getting confused about the transformation material above (and you're not alone), then the following remarks might help.

There might be a temptation to write down something that appears to be the 'coordinate form':

$$x^{\bar{\imath}} = \Lambda^{\bar{\imath}}_{i} x^i \qquad \text{(meaningless).} \tag{2.20}$$

This looks a bit like Eq. (2.19), and a bit like Eq. (2.12); it feels like it should be describing the transformation of the $\{x^i\}$ coordinates into the $\{x^{\bar{\imath}}\}$ ones, so that it may appear to be the analogue of Eq. (2.21) below. It's none of these things, however.

Notice that we haven't, so far, talked of *coordinates* at all. When we talked of components in Sect. 2.5, and of changing bases in Sect. 2.7 (which of course we understand to be a change of coordinates), we did so by talking about *basis vectors*, $\overline{e}_i$. This is like talking about cartesian coordinates by talking exclusively about the basis vectors **i**, **j** and **k**, and avoiding talking about $\{x, y, z\}$.

In Sect. 1.1 of part 3, we introduce coordinates as *functions* on the manifold, $\{x^i : M \to \mathbb{R}\}$, and in Eq. (3.4) we define the basis vectors associated with them as $\overline{e}_i = \partial/\partial x^i$ (see also the discussion at the end of Sect. 1.4 of part 3). Thus Eq. (2.20) is suggesting that these coordinate functions are linear combinations of each other; that will generally not be true, and it is possible to get very confused in example 3.2, for example, by thinking in this way. It is tempting to look at Eq. (2.20) and interpret the x^i as components of the basis vectors, or something like that, but the real relationship is the other way around: the basis vectors are derived from the coordinate functions, and show the way in which the coordinate functions change as you move around the manifold. The components of the basis vectors are very simple – see Eq. (2.6).

It's also worth stressing, once we're talking about misconceptions, that *neither position vectors, nor connecting vectors, are 'vectors' in the sense introduced in this part of the notes*. In a flat space, such as the euclidean space of our intuitions or the flat space of Special Relativity, the difference between them disappears or, to put it another way, there is a one-to-one correspondence between 'a vector in the space' and 'the difference between two positions' (which is what a difference vector is). In a curved space, it's useful to talk about the former (and we do, at length), but the latter won't often have much physical meaning. It is because of this correspondence that we can 'parallel transport' vectors everywhere in a flat space (see Sect. 3.2 of part 3), without thinking about it, which means we can define vector differentiation without having to think very hard about it.

If you think of a *vector space* – that is, a field of vectors, such as you tend to imagine in the case of the electric field – then the things you imagine existing at each point in space-time are straightforwardly vectors. That is, they're a thing with size and direction, defined at each point.

3 Examples of bases and transformations

So far, so abstract. By now, we are *long* overdue for some illustrations.

3.1 Flat cartesian space

Consider the natural vectors on the euclidean plane – that is, the vectors you learned about in school. The obvious thing to do is to pick our basis to be the unit vectors along the x and y axes: $\overline{e}_1 = \overline{e}_x$ and $\overline{e}_2 = \overline{e}_y$. That means that the vector $\overline{A}$, say, which points from the origin to a point two units along and one up, can be written as $\overline{A} = 2\overline{e}_1 + 1\overline{e}_2$, or to have components $A^1 = 2$, $A^2 = 1$. We have chosen these basis vectors to be the usual orthonormal ones: however, we are *not* required to do this by anything in Sect. 2, and indeed we *cannot* even say this at this stage, because we have not (yet) defined a metric, and so we have no inner product, so that the ideas of 'orthogonal' and 'unit' do not yet exist.

What are the one-forms in this space? Possibly surprisingly, there is nothing in Sect. 2 which tells us what they are, so that we can pick anything we like as a one-form in the euclidean plane, as long as that one-form-thing obeys the axioms of a vector space (Sect. 1.1), and as long as whatever rule we devise for contracting vectors and one-form-things conforms to the constraint of Eq. (2.1).

For one-forms, then, we'll choose sets of planes all parallel to each other, with the property that if we 'double' a one-form, then the spacing between the planes halves (a picture helps here!). For our contraction rule, $\langle \cdot, \cdot \rangle$, we'll choose: "the number $\langle \widetilde{p}, \overline{A} \rangle$ is the number of planes of the one-form $\widetilde{p}$ which the vector $\overline{A}$ passes through". If the duality property Eq. (2.3a) is to hold, then this fixes the 'planes' of $\widetilde{\omega}^1$ to be lines perpendicular to the x-axis, one unit apart, and the planes of $\widetilde{\omega}^2$ to be similarly perpendicular to the y-axis.

For our metric, we can choose simply $g_{ij} = g^{ij} = g^i{}_j = \delta^i_j$ (in *this* coordinate system). This means that the length-squared of the vector $\overline{A} = 2\overline{e}_1 + 1\overline{e}_2$ is

$$\mathsf{g}(\overline{A}, \overline{A}) = g_{ij} A^i A^j = A^1 A^1 + A^2 A^2 = (2)^2 + (1)^2 = 5,$$

which corresponds to our familiar value for this, from Pythagoras' theorem.

The other interesting thing about this metric is that, when we use it to lower the indexes of an arbitrary vector $\overline{A}$, we find that $A_i = g_{ij} A^j = A^i$. In other words, for this metric (the natural one for this flat space, with orthonormal basis vectors) one-forms and vectors have the *same* components, so that we cannot any longer really tell the difference between them, and have to work hard to think of them as being separate. *This* is why you have never had to deal with one-forms before, because the distinction between the two things in the space of our normal (mathematical) experience is invisible.

Note that in this section we have distinguished *cartesian space* from a *euclidean space*. carte
The first is the system of rectilinear coordinates you learned about in school, and the second eucli
is the flat space of our normal experience (where the ratio of the circumference of a circle to its diameter is π, and so on). Here, we have been talking about euclidean space described by cartesian coordinates. In the next section we will also discuss euclidean space, but not in cartesian coordinates, and in the following section on Minkowski space we use cartesian coordinates to describe a non-euclidean space, which does not have the same metric as the two previous examples.

3.2 Polar coordinates

An alternative way of devising vectors for the Euclidean plane is to use polar coordinates. It is convenient to introduce these using the transformation equation Eq. (2.12). The radial and tangential basis vectors are

$$\overline{e}_{\bar{1}} = \overline{e}_r = \cos\theta\,\overline{e}_1 + \sin\theta\,\overline{e}_2 \tag{2.21a}$$

$$\overline{e}_{\bar{2}} = \overline{e}_\theta = -r\sin\theta\,\overline{e}_1 + r\cos\theta\,\overline{e}_2, \tag{2.21b}$$

where $\overline{e}_1 = \overline{e}_x$ and $\overline{e}_2 = \overline{e}_y$ as before. Note that these basis vectors vary over the plane, and that although they are orthogonal (though we 'don't know that yet' since we haven't defined a metric), they are not orthonormal. Thus we can write

$$\overline{e}_{\bar{1}} = \overline{e}_r = \Lambda^1_{\ \bar{1}}\overline{e}_1 + \Lambda^2_{\ \bar{1}}\overline{e}_2$$
$$\overline{e}_{\bar{2}} = \overline{e}_\theta = \Lambda^1_{\ \bar{2}}\overline{e}_1 + \Lambda^2_{\ \bar{2}}\overline{e}_2,$$

and so discover that

$$\Lambda^i_{\ \bar{\imath}} = \begin{pmatrix} \Lambda^1_{\ \bar{1}} & \Lambda^1_{\ \bar{2}} \\ \\ \Lambda^2_{\ \bar{1}} & \Lambda^2_{\ \bar{2}} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}. \tag{2.21c}$$

From Eq. (2.18) therefore,

$$\Lambda^{\bar{\imath}}_{\ i} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix}. \tag{2.21d}$$

We therefore know how to transform the components of vectors and one-forms between cartesian and plane polar coordinates. What does the metric tensor ($g_{ij} = \delta_{ij}$ in cartesian coordinates, remember) look like in these new coordinates? The components are just

$$g_{\bar{\imath}\bar{\jmath}} = \Lambda^i_{\ \bar{\imath}}\Lambda^j_{\ \bar{\jmath}}g_{ij},$$

and writing these components out in full, we find

$$g_{\bar{\imath}\bar{\jmath}} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}. \tag{2.22}$$

We see that, even though coordinates (x, y) and (r, θ) are describing the *same* flat space, the metric looks a little more complicated in the polar-coordinate coordinate system than it does in the plain cartesian one, and looking at this ‘cold’, we would have difficulty identifying the space described by Eq. (2.22) as flat Euclidean space.

We will see a lot more of this in the parts to come.

The definition of polar coordinates in Eq. (2.21) is *not* the one you are probably familar with, due to the presence of the r in the transformations. The expressions can be obtained very directly, by *choosing* as the transformation

$$\overline{e}_r \equiv \Lambda^x_r \overline{e}_x + \Lambda^y_r \overline{e}_y = \frac{\partial x}{\partial r}\overline{e}_x + \frac{\partial y}{\partial r}\overline{e}_y$$
$$\overline{e}_\theta \equiv \Lambda^x_\theta \overline{e}_x + \Lambda^y_\theta \overline{e}_y = \frac{\partial x}{\partial \theta}\overline{e}_x + \frac{\partial y}{\partial \theta}\overline{e}_y,$$

and similarly for $\overline{e}_\theta$. This is known as a *coordinate basis*, since it is generated directly from the relationship between the coordinate functions, $r^2 = x^2 + y^2$ and $\tan\theta = y/x$. Although this is a very natural definition of the new basis vectors, and is the type of transformation we will normally prefer, these basis vectors $\overline{e}_r$ and $\overline{e}_\theta$ are not of unit length, and indeed are of different lengths at different points in the coordinate plane. That is why the usual definition of polar basis vectors is chosen to be $\overline{e}_\theta = (1/r)\partial x/\partial\theta\,\overline{e}_x + (1/r)\partial y/\partial\theta\,\overline{e}_y$, which is a *non-coordinate basis*. See Schutz §§5.2 and 5.5 for further discussion.

coord

3.3 Minkowski space: vectors in special relativity

The final very important example of the ideas of this part is the space of Special Relativity: *Minkowski space*. See Schutz, ch.2, though he comes at the transformation matrix Λ from a slightly different angle.

Mink spac

Here there are four dimensions rather than two, and a basis for the space is formed from $\overline{e}_0 = \overline{e}_t$ and $\overline{e}_{1,2,3} = \overline{e}_{x,y,z}$. As is conventional in SR, we will now use greek indices for the vectors and one-forms, with the understanding that greek indices run over $\{0, 1, 2, 3\}$. The metric on this space, in these coordinates, is

$$g_{\mu\nu} \equiv \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1) \tag{2.23}$$

(note that this convention for the metric is the same as Schutz).

Vectors in this space are $\overline{A} = A^\mu \overline{e}_\mu$, and we can use the metric to lower the indexes and form the components in the dual space (which we define in a similar way to the way we defined the dual space in Sect. 3.1). Thus the contraction between a one-form $\widetilde{A}$ and vector $\overline{B}$ is just

$$\begin{aligned}\left\langle \widetilde{A}, \overline{B} \right\rangle &= A_\mu B^\mu = A_0 B^0 + A_1 B^1 + A_2 B^2 + A_3 B^3 \\ &= \eta_{\mu\nu} A^\mu B^\nu = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3.\end{aligned}$$

This last expression should be very familiar to you, since it is exactly the definition of the *norm* of two vectors which was so fundamental, and which seemed so peculiar, in special relativity.

We can define transformations from these Minkowski-space coordinates to new ones in the same space, by specifying the elements of a transformation matrix Λ (cf. Schutz §2.2). We can do this however we like, but there is a subset of these transformations which are particularly useful, since they result in the metric in the new coordinates having the *same* form as the metric in the old one, namely $g_{\bar{\mu}\bar{\nu}} = \eta_{\bar{\mu}\bar{\nu}}$. One of the simplest sets of such transformation matrices (parameterised by $0 \leq v \leq 1$) is

$$\Lambda^{\bar{\mu}}_{\ \mu}(v) = \begin{pmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.24}$$

where $\gamma = 1/\sqrt{1-v^2}$. Again, this should be rather familiar.

Examples

Some of the examples below are taken from the earlier presentations of this course by Martin Hendry; these are noted by “(MAH)” and the example sheet where they occurred. Exercises 1–12 of Schutz’s chapter 3 are also useful.

Example 2.1 (section 1.1)

Demonstrate that the set of ‘ordinary vectors’ does indeed satisfy these axioms. Demonstrate that the set of all functions also satisfies them and is thus a vector space. Demonstrate that the subset of functions $\{e^{ax} : a \in \mathbb{R}\}$ is a vector space (hint: think about what the ‘vector addition’ operator should be in this case). Can you think of other examples?

Example 2.2 (section 1.1)

Prove that if $\{b_i\}$ are the components of an arbitrary vector B with respect to an orthonormal basis $\{e_i\}$, then $b_i = B \cdot e_i$.

Example 2.3 (section 2.5)

Tensor components A^{ij} and B^{ij} are equal in one coordinate frame. By considering the transformation law for a $(2, 0)$ tensor (introduced later in this part, in Sect. 2.7), show that they must be equal in any coordinate trame. Show that if A^{ij} is symmetric in one coordinate frame, it is symmetric in any frame. (MAH, 3.2)

Example 2.4 (section 2.6)

Justify each of the steps in Eq. (2.9a).

Example 2.5 (section 2.6)

(a) Given that T is a $\binom{1}{2}$ tensor, $\overline{A}$ and $\overline{B}$ are vectors, $\widetilde{p}$ is a one-form, and g is the metric, give the $\binom{M}{N}$ rank of each of the following objects, where as usual $\cdot$ represents an unfilled argument to the function T (not all the following are valid; if not, say why):

1. $\overline{A}(\widetilde{\cdot})$	2. $\widetilde{p}(\overline{\cdot})$	3. $\mathsf{T}(\widetilde{\cdot}, \overline{\cdot}, \overline{\cdot})$	4. $\mathsf{T}(\widetilde{p}, \overline{\cdot}, \overline{\cdot})$
5. $\mathsf{T}(\widetilde{p}, \overline{A}, \overline{\cdot})$	6. $\mathsf{T}(\widetilde{\cdot}, \overline{A}, \overline{\cdot})$	7. $\mathsf{T}(\widetilde{\cdot}, \overline{\cdot}, \overline{B})$	8. $\mathsf{T}(\widetilde{\cdot}, \overline{A}, \overline{B})$
9. $\mathsf{T}(\widetilde{p}, \overline{A}, \overline{B})$	10. $\overline{A}(\overline{\cdot})$	11. $\widetilde{p}(\widetilde{\cdot})$	

(b) State which of the following are valid expressions representing the components of a tensor. For each of the expressions which is a tensor, state the $\binom{M}{N}$ type of the tensor; and for each expression which is not, explain why not.

1. g_{ii}	2. $g_{ij} T^j{}_{kl}$	3. $g_{ik} T^j{}_{kl}$	4. $T^i{}_{ij}$
5. $g_{ij} A^i A^j$	6. $g_{ij} A^k A^k$		

If you're looking at this after studying part 3, then how about (7) $A^i{}_{,j}$ and (8) $A^i{}_{;j}$?

What about $A^i = \overline{A}(\widetilde{\omega}^i)$? [Objective 2; based on Class Test, 2002]

Example 2.6 (section 2.7)

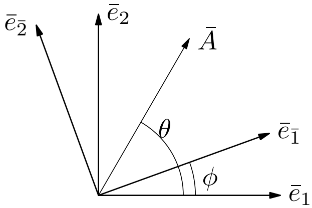

Figure 3

Figure 3 shows a vector $\overline{A}$ in two different coordinate systems. We have $\overline{A} = \cos\theta\,\overline{e}_1 + \sin\theta\,\overline{e}_2 = \cos(\theta-\phi)\overline{e}_{\bar{1}} + \sin(\theta-\phi)\overline{e}_{\bar{2}}$. Obtain $\overline{e}_{1,2}$ in terms of $\overline{e}_{\bar{1},\bar{2}}$ and vice versa, and obtain $A^{\bar{\imath}}$ in terms of $\cos\theta$, $\cos\phi$, $\sin\theta$ and $\sin\phi$. Thus identify the components of the matrix $\Lambda^{i}_{\bar{\imath}}$ (see also Eq. (2.19)).

Example 2.7 (section 2.7)

By repeating the logic which led to Eq. (2.12), prove Eq. (2.14) and Eq. (2.17). Alternatively, and more directly, use the results of example 2.8. [Objective 3]

Example 2.8 (section 2.7)

Make a table showing all the transformations $\overline{e}_i \leftrightarrow \overline{e}_{\bar{\imath}}$, $A^i \leftrightarrow A^{\bar{\imath}}$, $\widetilde{\omega}^i \leftrightarrow \widetilde{\omega}^{\bar{\imath}}$ and $p_i \leftrightarrow p_{\bar{\imath}}$, patterned after Eq. (2.12) and Sect. 1.4. You will need to use the fact that the Λ matrices are inverses of each other, and that $\overline{A} = A^i \overline{e}_i = A^{\bar{\imath}} \overline{e}_{\bar{\imath}}$. This is a repetitive, slightly tedious, but *extremely* valuable exercise. [Objective 3, Objective 4]

Example 2.9 (section 3.2)

Consider the vector $\overline{A} = 1\overline{e}_1 + 1\overline{e}_2$. Using Eq. (2.12) and the appropriate transformation matrix for polar coordinates, determine the components of this vector in the polar basis, at the points $(r, \theta) = (1, \pi/4)$, $(1, 0)$, $(2, \pi/4)$ and $(2, 0)$? Use the metric for these coordinates, Eq. (2.22) to find the length of $\overline{A}$ at each of these points. What happens at the origin? [Objective 3, Objective 4]

Example 2.10 (section 3.3)

The metric in the new coordinates is $\eta_{\bar{\mu}\bar{\nu}} = \Lambda^{\mu}_{\bar{\mu}}\Lambda^{\nu}_{\bar{\nu}}\eta_{\mu\nu}$. Confirm that $\eta_{\bar{\mu}\bar{\nu}} = \eta_{\mu\nu}$.

What is the inverse, $\Lambda^{\mu}_{\bar{\mu}}$ of Eq. (2.24)? Use it to write down the components of the metric tensor Eq. (2.23) after transformation by this Λ, and verify that $g_{\bar{\mu}\bar{\nu}} = \eta_{\bar{\mu}\bar{\nu}}$.

Consider the transformation matrix

$$\Lambda^{\mu}_{\bar{\mu}} = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{i}$$

which is the simplest transformation which 'mixes' the x- and t-coordinates. By *requiring* that $g_{\bar{\mu}\bar{\nu}} = \eta_{\bar{\mu}\bar{\nu}}$ after transformation by this Λ, find constraints on the parameters a, b, c, d (you can freely add the constraint $b = c$; why?), and so deduce the matrix Eq. (2.24). [Objective 3, Objective 4]