Linear functionals and dual spaces *(Secs. 31, 32, 14, 19)*

(See also Simmonds, *A Brief on Tensor Analysis*, Chap. 2)

Definitions: A *linear functional* is a linear operator whose codomain is $\mathcal{F}$ (a one-dimensional vector space). The set of such, $\mathcal{V}^* \equiv \mathcal{L}(\mathcal{V};\mathcal{F})$, is the *dual space of* $\mathcal{V}$.

The dimension of $\mathcal{V}^*$ is equal to that of $\mathcal{V}$. The elements of $\mathcal{V}^*$ are represented by row matrices, those of $\mathcal{V}$ by column matrices.

If $\dim \mathcal{V} = \infty$, one usually considers only the linear functionals which are *continuous* with respect to some topology. This space is called the *topological dual*, as opposed to the *algebraic dual*. The topological dual spaces of infinite-dimensional vector spaces are of even greater practical importance than those of finite-dimensional spaces, because they can contain new objects of a different nature from the vectors of the original spaces. In particular, the linear functionals on certain function spaces include *distributions*, or "generalized functions", such as the notorious Dirac δ. (Chapter 5 of Milne's book is one place to find an introduction to distributions.)

I will use the notation $\tilde{V}$, $\tilde{U}$, ... for elements of $\mathcal{V}^*$, since the textbook's notation "$\vec{v}^{\,*}$" could be misleading. (There is no particular $\vec{v} \in \mathcal{V}$ to which a given $\tilde{V} \in \mathcal{V}^*$ is necessarily associated.) Thus $\tilde{U}(\vec{v}) \in \mathcal{F}$. This notation is borrowed from B. Schutz, *Geometrical Methods of Mathematical Physics.*

Often one wants to consider $\tilde{U}(\vec{v})$ as a function of $\tilde{U}$ with $\vec{v}$ fixed. Sometimes people write

$$\left\langle \tilde{U}, \vec{v} \right\rangle \equiv \tilde{U}(\vec{v}).$$

Thus $\langle \ldots \rangle$ is a function from $\mathcal{V}^* \times \mathcal{V}$ to $\mathcal{F}$ (sometimes called a *pairing*.) We have

$$\left\langle \alpha\tilde{U} + \tilde{V}, \vec{v} \right\rangle = \alpha \left\langle \tilde{U}, \vec{v} \right\rangle + \left\langle \tilde{V}, \vec{v} \right\rangle,$$
$$\left\langle \tilde{U}, \alpha\vec{u} + \vec{v} \right\rangle = \alpha \left\langle \tilde{U}, \vec{u} \right\rangle + \left\langle \tilde{U}, \vec{v} \right\rangle.$$

(Note that there is no conjugation in either formula. The pairing is *bilinear*, not sesquilinear.)

It will not have escaped your notice that this notation conflicts with one of the standard notations for an inner product — in fact, the one which I promised to use in this part of the course. For that reason, I shall **not** use the bracket notation for the result of applying a linear functional to a vector; I'll use the function notation, $\tilde{U}(\vec{v})$.

Relation to an inner product

Suppose that $\mathcal{V}$ is equipped with an inner product. Let's use the notation $\langle \vec{u}, \vec{v} \rangle$, with

$$\langle \alpha\vec{u}, \vec{v} \rangle = \overline{\alpha}\, \langle \vec{u}, \vec{v} \rangle\,, \qquad \langle \vec{u}, \alpha\vec{v} \rangle = \alpha\, \langle \vec{u}, \vec{v} \rangle\,.$$

(Thus $\langle \vec{u}, \vec{v} \rangle \equiv \vec{v} \cdot \vec{u}$.)

Definition: The *norm* of a linear functional $\tilde{U}$ is the number

$$\|\tilde{U}\|_{\mathcal{V}^*} \equiv \sup_{\vec{0} \neq \vec{v} \in \mathcal{V}} \frac{\|\tilde{U}(\vec{v})\|_{\mathcal{F}}}{\|\vec{v}\|_{\mathcal{V}}}\,.$$

Riesz Representation Theorem (31.2). *Let $\mathcal{V}$ be a Hilbert space. (This includes any finite-dimensional space with an inner product.) Then*

(1) *Every $\vec{u} \in \mathcal{V}$ determines a $\tilde{U}_{\vec{u}} \in \mathcal{V}^*$ by*

$$\tilde{U}_{\vec{u}}(\vec{v}) \equiv \langle \vec{u}, \vec{v} \rangle\,.$$

(2) *Conversely, every [continuous] $\tilde{U} \in \mathcal{V}^*$ arises in this way from some (unique) $\vec{u} \in \mathcal{V}$.*

(3) *The correspondence $\underline{G}\colon \vec{u} \mapsto \tilde{U}_{\vec{u}} \equiv \underline{G}(\vec{u})$ is antilinear and preserves the norm:*

$$\tilde{U}_{\alpha\vec{u}+\vec{v}} = \overline{\alpha}\tilde{U}_{\vec{u}} + \tilde{U}_{\vec{v}}\,;$$

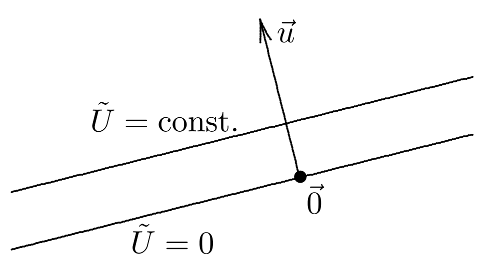
$$\|\tilde{U}_{\vec{u}}\|_{\mathcal{V}^*} = \|\vec{u}\|_{\mathcal{V}}\,.$$

Thus if $\mathcal{F} = \mathbf{R}$, then $\underline{G}$ is an isometric isomorphism of $\mathcal{V}$ onto $\mathcal{V}^*$. [Therefore, when there is an inner product, we can think of $\mathcal{V}$ and $\mathcal{V}^*$ as essentially the same thing.]

Proof: See Bowen & Wang, p. 206. The geometrical idea is that

$\vec{u} =$ normal vector to $\ker \tilde{U}$

$=$ gradient of the function $\tilde{U}(\vec{v})$.



(Here *gradient* is meant in the geometrical sense of a vector whose inner product with a unit vector yields the directional derivative of the function in that direction.)

Note: Closer spacing of the level surfaces is associated with a *longer* gradient vector.

Coordinate representation

Choose a basis. Recall that in an *ON* basis $\{\hat{e}_j\}$,

$$\langle \vec{u}, \vec{v} \rangle = \sum_{j=1}^{N} \overline{u^j}\, v^j .$$

Thus $\tilde{U}_{\vec{u}}$ is the linear functional with matrix $(\overline{u^1}, \ldots, \overline{u^N})$. (In particular, in an ON basis the gradient of a real-valued function is represented simply by the row vector of partial derivatives.)

If the basis (call it $\{\vec{d}_j\}$) is not ON, then

$$\langle \vec{u}, \vec{v} \rangle = \sum_{j,k=1}^{N} g_{jk}\, \overline{u^j}\, v^k \equiv g_{jk}\, \overline{u^j}\, v^k ,$$

where $g_{jk} \equiv \left\langle \vec{d}_j, \vec{d}_k \right\rangle$ (the "metric tensor" of differential geometry and general relativity). Note that g_{jk} is symmetric if $\mathcal{F} = \mathbf{R}$ (a condition henceforth referred to briefly as "the real case".) We see that $\tilde{U}_{\vec{u}}$ has now the matrix $\{g_{jk}\, \overline{u^j}\}$ (where j is summed over, and the free index k varies from 1 to N). Thus (in the real case) $\{g_{jk}\}$ is the matrix of $\underline{G}$.

Conversely, given $\tilde{U} \in \mathcal{V}^*$ with matrix $\{U_j\}$, so that

$$\tilde{U}(\vec{v}) = U_j v^j ,$$

then the corresponding $\vec{u} \in \mathcal{V}$ is given (in the real case) by

$$u^k = g^{kj} U_j ,$$

where $\{g^{jk}\}$ is the matrix of $\underline{G}^{-1}$ — i.e., the inverse matrix of $\{g_{jk}\}$. The reason for using the same letter for two different matrices — inverse to each other — will become clear later.

The dual basis

Suppose for a moment that we do not have an inner product (or ignore it, if we do). Choose a basis, $\{\vec{d}_j\}$, for $\mathcal{V}$, so that

$$\vec{v} = v^j \vec{d}_j = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^N \end{pmatrix} .$$

Then (**definition**) the *dual basis*, $\{\tilde{D}^j\}$, is the basis (**for** $\mathcal{V}^*$!) consisting of those linear functionals having the matrices

$$D^j = (0, 0, \ldots, 0, 1, 0, \ldots) \quad (1 \text{ in } j\text{th place}).$$

That is,

$$\tilde{D}^j(\vec{d}_k) \equiv \delta^j{}_k\,, \quad \forall j, k.$$

If $\tilde{U} = U_j \tilde{D}^j$, then

$$\tilde{U}(\vec{v}) = \left[U_j \tilde{D}^j\right]\left(v^k \vec{d}_k\right) = U_j v^j = (U_1, U_2, \ldots)\begin{pmatrix} v^1 \\ \vdots \\ v^N \end{pmatrix}.$$

The reciprocal basis

If $\mathcal{V}$ is a real inner product space, **define** the *reciprocal basis* $\{\vec{\overline{d}}^j\}$ (**in** $\mathcal{V}$!) to be the vectors in $\mathcal{V}$ corresponding to the dual-basis vectors $\tilde{D}^j$ under the Riesz isomorphism:

$$\vec{\overline{d}}^j \equiv \underline{G}^{-1} \tilde{D}^j.$$

Equivalent definition (Sec. 14): $\vec{\overline{d}}^j$ is defined by

$$\left\langle \vec{\overline{d}}^j, \vec{d}_k \right\rangle = \delta^j{}_k\,, \quad \forall j, k.$$

Note that the bar in this case does *not* indicate complex conjugation. (It covers just the symbol "d", not the superscript.) If $\{\vec{d}_j\}$ is ON, then $g_{jk} = \delta_{jk}$ and hence $\vec{\overline{d}}^j = \vec{d}_j$ for all j. We "discovered" the reciprocal basis earlier, while constructing projection operators associated with nonorthogonal bases. The reciprocal basis of the reciprocal basis is the original basis.

Given $\vec{v} \in \mathcal{V}$, we may expand it as

$$\vec{v} = v^j \vec{d}_j = v_j \vec{\overline{d}}^j.$$

Note that

$$v^j = \tilde{D}^j(\vec{v}) = \left\langle \vec{\overline{d}}^j, \vec{v} \right\rangle,$$

and similarly

$$v_j = \left\langle \vec{d}_j, \vec{v} \right\rangle;$$

to find the coordinates of a vector with respect to one basis you take the inner products with respect to the elements from the other basis. (Of course, if $\{d_j\}$ is ON, then the two bases are the same and all the formulas we're looking at simplify.) Now we see that

$$v_j = v^k \left\langle \vec{d}_j, \vec{d}_k \right\rangle = g_{jk} v^k;$$

the counterpart equation for the other basis is

$$v^j = g^{jk} v_k \,.$$

There follow

$$\langle \vec{u}, \vec{v} \rangle = u^j v_j = u_j v^j = g_{jk} u^j v^k = g^{jk} u_j v_k \,.$$

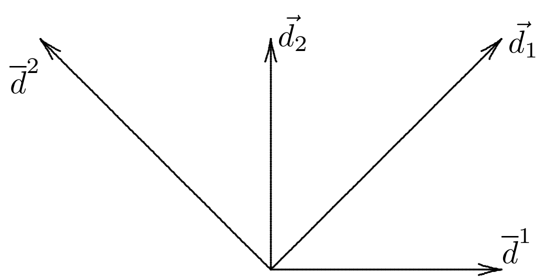
After practice, "raising and lowering indices" with g becomes routine; g (with indices up or down) serves as "glue" connecting adjacent vector indices together to form something scalar.

Geometry of the reciprocal basis

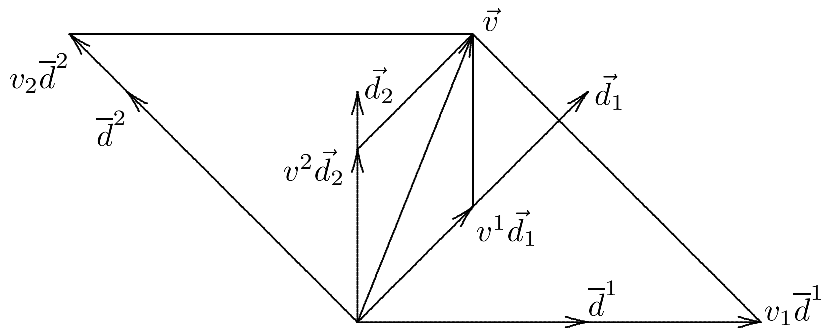
In a two-dimensional space, for instance, $\overleftarrow{d}^1$ must be orthogonal to $\vec{d}_2$, have positive inner product with $\vec{d}_1$, and have length inversely proportional to $\|\vec{d}_1\|$ so that

$$\left\langle \overleftarrow{d}^1, \vec{d}_1 \right\rangle = 1.$$

Corresponding remarks hold for $\overleftarrow{d}^2$, and we get a picture like this:



Here we see the corresponding *contravariant* (v^j) and *covariant* (v_j) components of a vector $\vec{v}$:



Application: Curvilinear coordinates. (See M. R. Spiegel, *Schaum's Outline of Vector Analysis*, Chaps. 7 and 8.)

Let $x^j \equiv f^j(\xi^1, \ldots, \xi^N)$. For example,

$$\begin{aligned} x &= r\cos\theta, \qquad & x^1 &= x, \qquad & x^2 &= y, \\ y &= r\sin\theta; \qquad & \xi^1 &= r, \qquad & \xi^2 &= \theta. \end{aligned}$$

Let $\mathcal{V} = \mathbf{R}^N$ be the vector space where the Cartesian coordinate vector $\vec{x}$ lives. It is equipped with the standard inner product which makes the natural basis ON. Associated with a coordinate system there are two sets of basis vectors at each point:

1) the normal vectors to the coordinate surfaces ($\xi^j =$ constant):

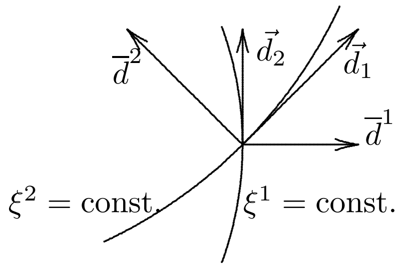
$$\nabla\xi^j = \left(\frac{\partial\xi^j}{\partial x^1}, \frac{\partial\xi^j}{\partial x^2}, \ldots\right) \equiv \overline{d}^j.$$

From a fundamental point of view, these are best thought of as vectors in $\mathcal{V}^*$, or "covectors". In classical vector analysis they are regarded as members of $\mathcal{V}$, however. In effect, the dual-space vectors have been mapped into $\mathcal{V}$ by $\underline{G}^{-1}$; they are a reciprocal basis.

2) the tangent vectors to the coordinate lines ($\xi^k =$ constant for $k \neq j$):

$$\frac{d\vec{x}}{d\xi^j} = \begin{pmatrix} \frac{\partial x^1}{\partial\xi^j} \\ \frac{\partial x^2}{\partial\xi^j} \\ \vdots \end{pmatrix} \equiv \vec{d}_j\,.$$

These are ordinary (nondual) vectors (members of $\mathcal{V}$), sometimes called "contravariant vectors".

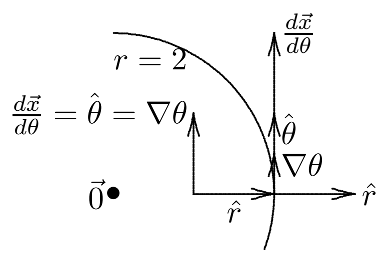


Note that $\left\langle (\nabla\xi^j), \left(\frac{d\vec{x}}{d\xi^k}\right)\right\rangle = \frac{\partial\xi^j}{\partial\xi^k} = \delta^j{}_k$. Thus the two sets of vectors form mutually reciprocal bases. (Another way of looking at this equation is that the inner product or pairing of a row of the Jacobian matrix of the coordinate transformation with a column of the inverse of the Jacobian matrix is the corresponding element (0 or 1) of the unit matrix.)

For polar coordinates, define $\hat{r}$ and $\hat{\theta}$ to be the usual unit vectors. Then you will find that

$$\frac{d\vec{x}}{dr} = \hat{r}, \qquad \frac{d\vec{x}}{d\theta} = r\hat{\theta}\,;$$
$$\nabla r = \hat{r}, \qquad \nabla\theta = \frac{\hat{\theta}}{r}\,.$$

(The geometrical interpretation of the two θ equations is that an increment in θ changes $\vec{x}$ little if r is small, much if r is large; and that θ changes rapidly with $\vec{x}$ if r is small, slowly if r is large.) In this case the two bases are OG but not ON, hence they are distinct. Their orthogonality makes it possible to define uniquely the ON basis $\{\hat{r}, \hat{\theta}\}$ sitting halfway between them.



Change of basis

Again we ignore the inner product for awhile and study the dual basis. Recall our earlier notation:

$$\{\vec{v}_j\} = \text{“old” basis}, \qquad \{\vec{w}_j\} = \text{“new” basis},$$
$$\vec{x} = \alpha^j\vec{v}_j = \beta^k\vec{w}_k \text{ is an arbitrary element in } \mathcal{V}.$$

Suppose the transformation (old basis $\mapsto$ new basis) is

$$\vec{w}_k = R^j{}_k\vec{v}_j\,. \tag{1}$$

(In our previous discussion of change of basis, R was called S^{-1}.) Then the transformation (old coordinates $\mapsto$ new coordinates) is

$$\beta^k = \left(R^{-1}\right)^k{}_j \alpha^j. \tag{2}$$

(One matrix is "contragredient" to the other — the inverse of its transpose.)

Now look at $\tilde{U} \in \mathcal{V}^*$ and the two dual bases:

$$\tilde{U} = \gamma_j \tilde{V}^j = \delta_k \tilde{W}^k.$$

Then

$$\tilde{U}(\vec{x}) = \gamma_j \alpha^j = \delta_k \beta^k = \delta_k (R^{-1})^k{}_j \alpha^j.$$

Thus

$$\gamma_j = (R^{-1})^k{}_j \delta_k\,.$$

This may be denoted the transformation (new* coordinates $\mapsto$ old* coordinates), the * standing for the dual space, $\mathcal{V}^*$. This result is more useful to us in the inverse direction:

$$\delta_k = R^j{}_k \gamma_j \tag{3}$$

(old* coordinates $\mapsto$ new* coordinates). We can rewrite (3) so as to untangle the indices into a normal matrix multiplication:

$$\delta_k = \gamma_j R^j{}_k \qquad \text{or} \qquad \delta_k = (R^*)_k{}^j \gamma_j\,.$$

Note that (3) "looks like" (1). Historically, vectors in the dual space $\mathcal{V}^*$ were called *covariant vectors*, because under a change of coordinate system (basis) their coordinates transform "along with" the basis vectors in $\mathcal{V}$. The vectors in the original space $\mathcal{V}$ were called *contravariant vectors*, because their coordinates transform "in the opposite direction from" the basis vectors, as shown by (2). [Two familiar examples of the latter phenomenon are (a) the result of a change of a unit of measurement, and (b) the relation between the "active" rotation of an observer and the "passive" rotation of his view of the world.] Nowadays in many quarters it is considered in poor taste to talk of vectors as "transforming" at all: Vectors are abstract objects which remain *the same* no matter what coordinate system is used to describe them! Dual vectors are still called *covectors*, but the "co" just means "dual" to the "ordinary" vectors in $\mathcal{V}$.

Change of basis as Leibnitz would write it

Writing β as x, α as ξ, and R^{-1} as S, we cast the linear variable change (2) into the form of the general nonlinear change of variables considered earlier:

$$x^k = S^k{}_j \xi^j.$$

Note that

$$\frac{\partial x^k}{\partial \xi^j} = S^k{}_j\,.$$

By the inverse function theorem,

$$\frac{\partial \xi^j}{\partial x^k} = (S^{-1})^j{}_k = R^j{}_k\,.$$

The point of this remark is that a handy way to remember the respective transformation laws of vectors and covectors is through the following prototypes of each:

$$\text{contravector:} \qquad \text{Tangent vector to a curve, } \frac{d\vec{x}(t)}{dt}$$

$$\text{covector:} \qquad \text{Gradient of a function, } \left\{\frac{\partial f}{\partial x^k}\right\}$$

The transformation laws then follow from the multivariable chain rule:

$$\frac{dx^k}{dt} = \frac{\partial x^k}{\partial \xi^j}\,\frac{d\xi^j}{dt} \ \Rightarrow\ \beta^k = \frac{\partial x^k}{\partial \xi^j}\,\alpha^j \tag{2'}$$

$$\frac{\partial f}{\partial x^k} = \frac{\partial f}{\partial \xi^j}\,\frac{\partial \xi^j}{\partial x^k} \ \Rightarrow\ \delta_k = \frac{\partial \xi^j}{\partial x^k}\,\gamma_j \tag{3'}$$

These equations remain meaningful for nonlinear coordinate transformations — but that is material for another course.

The dual operator

Given $\underline{A}\colon \mathcal{V} \to \mathcal{U}$, there is a unique, linear $\underline{A}^*\colon \mathcal{U}^* \to \mathcal{V}^*$ **defined** by

$$[\underline{A}^*\tilde{U}](\vec{v}) = \tilde{U}(\underline{A}\vec{v}), \quad \forall \vec{v} \in \mathcal{V}. \tag{1}$$

That is,

$$\underline{A}^*\tilde{U} \equiv \tilde{U} \circ \underline{A}. \tag{2}$$

(This explicit formula proves the uniqueness and linearity.) In the "pairing" notation which we recently outlawed, this would be written

$$\left\langle \underline{A}^*\tilde{U}, \vec{v}\right\rangle = \left\langle \tilde{U}, \underline{A}\vec{v}\right\rangle. \tag{3}$$

Version (3) looks suspiciously like the definition of the *adjoint* operator in a Hilbert space. Indeed, if $\mathcal{U}$ and $\mathcal{V}$ are inner-product spaces, then $\mathcal{U}^*$ is isomorphic to $\mathcal{U}$ and $\mathcal{V}^*$ to $\mathcal{V}$ (up to conjugation in the complex case), and *under these isomorphisms,* $\underline{A}^*\colon \mathcal{U}^* \to \mathcal{V}^*$ *coincides with* $\underline{A}^*\colon \mathcal{U} \to \mathcal{V}$.

Practical applications of the dual operator are presented by Milne in Secs. 2.7, 2.9, 3.5(end), and 3.10. Unfortunately, we do not have time to discuss them in the course.

Writing $\tilde{U}(\vec{v})$ as the pairing $\left\langle \tilde{U}, \vec{v} \right\rangle$ emphasizes that each $\vec{v} \in \mathcal{V}$ defines a linear functional on $\mathcal{V}^*$:

$$[\underline{J}\vec{v}](\tilde{U}) \equiv \tilde{U}(\vec{v}).$$

That is, $\mathcal{V}$ is isomorphic to a subspace $\underline{J}[\mathcal{V}] \subset (\mathcal{V}^*)^*$.

If $\dim \mathcal{V} < \infty$ (and for many infinite-dimensional spaces too), $\underline{J}[\mathcal{V}]$ is *equal* to $\mathcal{V}^{**}$ — there are no other linear functionals on $\mathcal{V}^*$. When this is true, $\mathcal{V}$ is called *reflexive*; $\mathcal{V}$ and $\mathcal{V}^{**}$ are "the same".

Note that the isomorphism $\underline{J}\colon \mathcal{V} \leftrightarrow \mathcal{V}^{**}$ is *fixed* — independent of a choice of basis or any other structure. In contrast, the isomorphism $\underline{G}\colon \mathcal{V} \leftrightarrow \mathcal{V}^*$ depends on the inner product. If there is no inner product, $\mathcal{V}$ is certainly isomorphic to $\mathcal{V}^*$ because they have the same dimension (speaking now of finite-dimensional spaces), but there is no *preferred* ("natural" or "canonical") isomorphism. (The apparently obvious mapping $\vec{v}_j \leftrightarrow \tilde{V}^j$ is basis-dependent. It disagrees with $\vec{v}_j \leftrightarrow \underline{G}\vec{v}_j$ if the basis is not ON.)