REMARKS CONCERNING THE RELATION BETWEEN VECTORS AND COORDINATES, RESP. BETWEEN LINEAR MAPS AND MATRICES

JÜRGEN TOLKSDORF

NOVEMBER 14, 2012

1. Abstract vector spaces versus coordinate spaces

Let V be a real vector space of finite dimension $n \geq 1$.

For instance, $V := \mathbb{R}^n$. In the sequel we always regard $\mathbb{R}^n$ as an "abstract" vector space. That is, as a set $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ is the $n-$fold cartesian product of the field of real numbers. The linear structure on this set is then defined in terms of the linear structure of $\mathbb{R}$: For $\vec{x} \equiv (x_1, x_2, \ldots, x_n)$, $\vec{y} \equiv (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$ one sets: $\vec{x} + \lambda\vec{y} := (x_1 + \lambda y_1, x_2 + \lambda y_2, \ldots, x_n + \lambda y_n) \in \mathbb{R}^n$. It is straightforward to prove that with respect to this defined linear structure the set $\mathbb{R}^n$ becomes a real vector space of dimension n. However, $\mathbb{R}^n$ is not just an $n-$dimensional vector space. It also is endowed with a *distinguished basis*: $\vec{e}_1 \equiv (1, 0, \ldots, 0)$, $\vec{e}_2 \equiv (0, 1, \ldots, 0), \ldots, \vec{e}_n \equiv (0, 0, \ldots, 1) \in \mathbb{R}^n$. Indeed, by construction of $\mathbb{R}^n$ any vector $\vec{x} \equiv (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ may be uniquely decomposed as

$$\begin{aligned} \vec{x} \equiv (x_1, x_2, \ldots, x_n) &= x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n \\ &\equiv \sum_{1\leq k\leq n} x_k\vec{e}_k \,. \end{aligned} \tag{1}$$

In this context, the real numbers $x_1, x_2, \ldots, x_n \in \mathbb{R}$ are called the *cartesian coordinates* of the vector $\vec{x} \in \mathbb{R}^n$ *with respect to the basis* $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \in \mathbb{R}^n$. Apparently, we may also choose an arbitrary basis $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n \in \mathbb{R}^n$, such that

$$\vec{x} = (x_1, x_2, \ldots, x_n) = \sum_{1\leq k\leq n} \lambda_k\vec{b}_k \,. \tag{2}$$

Here, the real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ are referred to as the *coordinates* of the vector $\vec{x} \in \mathbb{R}^n$ *with respect to the basis* $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n \in \mathbb{R}^n$. Again, these coordinates are uniquely determined by the choice of the basis.

Notice that both coordinates $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ represent the *same vector* $\vec{x} \in \mathbb{R}^n$, however, with respect to *different basis* in $\mathbb{R}^n$, i.e.

$$\sum_{1\leq k\leq n} x_k\vec{e}_k = \sum_{1\leq k\leq n} \lambda_k\vec{b}_k \,. \tag{3}$$ decomp1

Clearly, every vector $\vec{b}_k \in \mathbb{R}^n$ can be uniquely written as

$$\vec{b}_k \equiv (b_{1k}, b_{2k}, \ldots, b_{nk}) = \sum_{1\leq i\leq n} b_{ik}\vec{e}_i \,, \quad (k = 1, \ldots, n) \,. \tag{4}$$ decomp2

Likewise, since $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n \in \mathbb{R}^n$ is supposed to form a *basis*, it follows that for all $k = 1, \ldots, n$ every vector $\vec{e}_k$ can be uniquely decomposed as

$$\vec{e}_k = \sum_{1 \leq i \leq n} a_{ik} \vec{b}_i \,. \tag{5}$$

Since on both sides of (11) the decomposition is unique the coordinates $x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ of the vector $\vec{x} \in \mathbb{R}^n$ are related by

$$\begin{aligned} x_k &= \sum_{1 \leq j \leq n} b_{ki} \lambda_i \,, \\ \lambda_k &= \sum_{1 \leq j \leq n} a_{ki} x_i \,. \end{aligned} \tag{6}$$

Furthermore, the real numbers $a_{ij}, b_{ij} \in \mathbb{R}$ $(1 \leq i, j \leq n)$ fulfill: $\sum_{1 \leq k \leq n} a_{ik} b_{kj} = \delta_{ij}$. Again, this is because of the unique decomposition of a vector with respect to a chosen basis.

Another example of a real vector space of dimension $n = kl$, consider the vector space of real matrices of size $k \times l$. That is,

$$V := \mathbb{R}^{k \times l} \equiv \left\{ \vec{x} \equiv \begin{pmatrix} x_{11} & x_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & x_{1l} \\ x_{21} & x_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & x_{2l} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{k1} & x_{k2} & \cdot & \cdot & \cdot & \cdot & \cdot & x_{kl} \end{pmatrix} \,\Big|\, x_{ij} \in \mathbb{R},\ 1 \leq i \leq k,\ 1 \leq j \leq l \right\}, \tag{7}$$

where the linear structure $\vec{x} + \lambda \vec{y}$ is provided by the usual adding of real matrices and the usual multiplication of a real matrix with a real number.

Again, due to its construction the real vector space $\mathbb{R}^{k \times l}$ comes equipped with a distinguished basis: $\vec{e}_{11}, \vec{e}_{12}, \ldots, \vec{e}_{1l}, \ldots, \vec{e}_{k1}, \ldots, \vec{e}_{kl}$, where $\vec{e}_{ij}$ is the matrix with entry 1 at the position i, j and zero otherwise, for all $1 \leq i \leq k$ and $1 \leq j \leq l$.

However, not every vector space comes equipped with respect to a distinguished basis. Indeed, the fact that both vector spaces $\mathbb{R}^n$ and $\mathbb{R}^{k \times l}$ allow for a "natural choice of a basis" strongly depends on how the respective underlying sets are endowed with a linear structure. In the case of an abstract vector space V, however, the underlying set is always supposed to have a linear structure. Therefore, an abstract vector space does not allow for a distinguished basis, in general, which can be used as a "reference basis". However, one has the following statement:

Proposition 1.1. *Every real vector space of finite dimension $n \geq 0$ is isomorphic to $\mathbb{R}^n$.*

Proof. Let $\vec{v}_1, \ldots, \vec{v}_n \in V$ be a basis. The mapping

$$\begin{aligned} \beta : V &\longrightarrow \mathbb{R}^n \\ \vec{x} = \sum_{1 \leq k \leq n} x_k \vec{v}_k &\mapsto (x_1, \ldots, x_n) \end{aligned} \tag{8}$$

coordisom1

is linear and bijective with the inverse being given by $\beta^{-1}(x_1, \ldots, x_n) := \sum_{1 \leq k \leq n} x_k \vec{v}_k$. □

The point to be stressed here is that the isomorphism β strongly depends on the chosen basis $\vec{v}_1, \dots, \vec{v}_n \in V$. This is, because the coordinates $x_1, \dots, x_n \in \mathbb{R}$ of $\vec{x} \in V$ do so! In fact, the isomorphism (8) is fully determined by the identification

$$V \ni \vec{v}_k \leftrightarrow \vec{e}_k \in \mathbb{R}^n \quad (k = 1, \dots, n)\,. \tag{9}$$

Since there is no preferred basis available in V one may choose another basis $\vec{u}_1, \dots, \vec{u}_n \in V$ to obtain a correspondingly alternative isomorphism

$$\begin{aligned} \alpha :\, V &\longrightarrow \mathbb{R}^n \\ \vec{x} = \sum_{1 \le k \le n} y_k \vec{u}_k &\mapsto (y_1, \dots, y_n)\,. \end{aligned} \tag{10}$$

coordisom2

Notice again that similar to (11)

$$\sum_{1 \le k \le n} x_k \vec{v}_k = \sum_{1 \le i \le n} y_i \vec{u}_i\,. \tag{11}$$

decomp1

It may thus not come as a big surprise that the *coordinate transformation*

$$\begin{aligned} \alpha \circ \beta^{-1} : \mathbb{R}^n &\xrightarrow{\simeq} \mathbb{R}^n \\ (x_1, \dots, x_n) &\mapsto (y_1, \dots, y_n) \end{aligned} \tag{12}$$

coordtrans1

is given by

$$y_i = \sum_{1 \le k \le n} a_{ik} x_k\,. \tag{13}$$

Here, the real numbers $a_{11}, \dots, a_{nn} \in \mathbb{R}$ are fully determined by

$$\vec{v}_k =: \sum_{1 \le i \le k} a_{ik} \vec{u}_i\,, \quad (k = 1, \dots, n)\,, \tag{14}$$

since both $\vec{u}_1, \dots, \vec{u}_n \in V$ and $\vec{v}_1, \dots, \vec{v}_n \in V$ form a basis in V. Accordingly, the inverse of the coordinate transformation (12) is provided by

$$x_k = \sum_{1 \le j \le n} b_{ki} y_i\,, \tag{15}$$

where the real numbers $b_{11}, \dots, b_{nn} \in \mathbb{R}$ are now fully determined by

$$\vec{u}_i =: \sum_{1 \le k \le n} b_{ki} \vec{v}_k\,, \quad (k = 1, \dots, n)\,. \tag{16}$$

Again, it follows that $\sum_{1 \le k \le n} a_{ik} b_{kj} = \delta_{ij}$, $(1 \le i, j \le n)$, since $(\alpha \circ \beta^{-1}) \circ (\alpha \circ \beta^{-1})^{-1} = (\alpha \circ \beta^{-1}) \circ (\beta \circ \alpha^{-1})$ equals the identity map on $\mathbb{R}^n$.

Notation: The (basis-dependent) isomorphism (8) is called a *coordinate map*. Accordingly, $(x_1, \dots, x_n) \in \mathbb{R}^n$ are referred to as the *coordinates* of the vector $\vec{x} \in V$ with respect to the chosen basis $\vec{v}_1, \dots, \vec{v}_n \in V$. Similarly, the (again basis dependent) isomorphism (10) provides another coordinate map. It allows to identify the same vector $\vec{x} \in V$ with the new coordinates $(y_1, \dots, y_n) \in \mathbb{R}^n$. The coordinates $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ are related to each other by the coordinate transformation $\alpha \circ \beta^{-1}$ and and its inverse $\beta \circ \alpha^{-1}$, both being isomorphisms on the *coordinate space* $\mathbb{R}^n$. Using this notation, $\mathbb{R}^n$ may be interpreted both as an abstract vector space and as its coordinate space, where the natural isomorphism is given by the identity map:

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \vec{x} \equiv (x_1, \ldots, x_n) = \sum_{1\leq k\leq n} x_k \vec{e}_k &\mapsto (x_1, \ldots, x_n) \in \mathbb{R}^n\,. \end{aligned} \tag{17}$$ coordisom3

Of course, with respect to the standard basis $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ the distinction between the interpretation of $\mathbb{R}^n$ as abstract vector space (similar to V) and its coordinate space becomes somehow redundant. However, notice that also on the vector space $\mathbb{R}^n$ one may choose an arbitrary basis $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{R}^n$. Then, (17) reads:

$$\begin{aligned} \alpha : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \vec{x} \equiv (x_1, \ldots, x_n) = \sum_{1\leq k\leq n} \lambda_k \vec{b}_k &\mapsto (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n\,. \end{aligned} \tag{18}$$ coordisom4

2. Linear maps and matrices

Let W be another real vector space of finite dimension $m \geq 1$, maybe different from $n = dimV$.

First, let us call in mind that every linear map $f : V \longrightarrow W$ is fully determined by the image of a chosen basis $\vec{v}_1, \ldots, \vec{v}_n \in V$ with respect to f. Indeed, let $\vec{w}_k := f(\vec{v}_k) \in W$ be given for all $k = 1, \ldots, n$. Notice that includes the case $\vec{w}_k = \vec{0} \in W$ for some (or even all) $k \in \{1, \ldots, n\}$. In fact, let $\vec{x} = \sum_{1\leq k\leq n} x_k \vec{v}_k \in V$ be arbitray. We may set $f(\vec{x}) := \sum_{1\leq k\leq n} x_k \vec{w}_k \in W$. By construction, the thus defined map $f : V \to W$ is linear. Let now $g : V \to W$ be any linear map which fulfills $g(\vec{v}_k) = \vec{w}_k \in W$ for all $k = 1, \ldots, n$. It follows that for all $\vec{x} = \sum_{1\leq k\leq n} x_k \vec{v}_k \in V$ the map

$$\begin{aligned} (g - f)(\vec{x}) := g(\vec{x}) - f(\vec{x}) &= \sum_{1\leq k\leq n} x_k g(\vec{v}_k) - \sum_{1\leq k\leq n} x_k \vec{w}_k \\ &= \vec{0} \in W\,. \end{aligned} \tag{19}$$

In other words, $g-f$ equals the "null-map" (i.e. the unique mapping which maps *every* $\vec{x} \in V$ into the zero-vector $\vec{0} \in W$). Therefore, $g = f$, which was to be demonstrated.

Notice that despite its definition, the linear map f does not depend on the choice of the basis $\vec{v}_1, \ldots, \vec{v}_n \in V$. Also notice that the constant mapping $f(\vec{v}_k) := \vec{w}$, where $\vec{w} \in W$ denotes some fixed vector, does not give rise to a *linear* map unless $\vec{w} = \vec{0}$.

In order to characterize a linear map $f : V \to W$, one has to distinguish the following two cases:

(1) $n \leq m$: In this case, f is *injective* if and only if $f(\vec{v}_1), \ldots, f(\vec{v}_n) \in W$ are linear independent. In fact, $\vec{x} = \sum_{1\leq k\leq n} x_k \vec{v}_k \in V \in ker(f)$ implies that $\vec{0} = f(\vec{x}) = \sum_{1\leq k\leq n} x_k f(\vec{v}_k) \in W$. Hence, it follows from the very definition of linear independency that $x_1 = \cdots = x_n = 0 \in \mathbb{R}$. That is, $f(\vec{x}) = \vec{0}$ implies $\vec{x} = \vec{0} \in V$. Consequently, the linear independency of $f(\vec{v}_1), \ldots, f(\vec{v}_n) \in W$ yields $ker(f) = \{\vec{0}\} \subset V$.

Conversely, let us assume that $ker(f) = \{\vec{0}\} \subset V$. That is, $\vec{0} \in V$ is the only vector which is mapped into $\vec{0} \in W$ by f. Therefore, $\vec{0} = \sum_{1\leq k\leq n} x_k f(\vec{v}_k) = f\big(\sum_{1\leq k\leq n} x_k \vec{v}_k\big)$. This implies that $\sum_{1\leq k\leq n} x_k \vec{v}_k = \vec{0} \in V$. Since $\vec{v}_1, \ldots, \vec{v}_n \in$

V is a basis, one infers that $x_1 = \cdots = x_n = 0$. This in turn proves that $f(\vec{v}_1), \ldots, f(\vec{v}_n) \in W$ are linear independent.

(2) $n > m$: We prove that f is *surjective* if and only if

$$span_{\mathbb{R}}(f(\vec{v}_1), \ldots, f(\vec{v}_n)) := \left\{ \vec{y} = \sum_{1 \le k \le n} \lambda_k f(\vec{v}_k) \,|\, \lambda_k \in \mathbb{R} \right\} = W\,. \tag{20}$$

surjectivity

First, notice that in the case considered $f(\vec{v}_1), \ldots, f(\vec{v}_n) \in W$ cannot be linear independent.

Assume that (20) holds true. We may put $\vec{x} := \sum_{1 \le k \le n} \lambda_k \vec{v}_k \in V$ to get $f(\vec{x}) = \vec{y} \in W$ for every $\vec{y} \in W$. Conversely, assume that for all $\vec{y} \in W$ there exists $\vec{x} \in V$, such that $\vec{y} = f(\vec{x})$. In this case, the decomposition $\vec{x} = \sum_{1 \le k \le n} \lambda_k \vec{v}_k$ yields (20).

As a consequence, one infers that a linear map $f : V \to W$ is an *isomorphism* if and only if $m = n$ and f is injective.

Let, respectively, $\alpha : V \xrightarrow{\simeq} \mathbb{R}^n$, $\vec{x} = \sum_{1 \le k \le n} x_k \vec{v}_k \mapsto (x_1, \ldots, x_n)$ and $\beta : W \xrightarrow{\simeq} \mathbb{R}^m$, $\vec{y} = \sum_{1 \le k \le m} y_k \vec{w}_k \mapsto (y_1, \ldots, y_m)$ be coordinate maps with respect to the chosen basis $\vec{v}_1, \ldots, \vec{v}_n \in V$ and $\vec{w}_1, \ldots, \vec{w}_m \in W$.

A linear map $f : V \longrightarrow W$ may then be represented by the linear map

$$\begin{aligned} f_{\alpha\beta} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x_1, \ldots, x_n) &\mapsto (y_1, \ldots, y_m) := \beta(f(\alpha^{-1}(x_1, \ldots, x_n)))\,, \end{aligned} \tag{21}$$

coordinatemap

i.e.

$$f_{\alpha\beta} = \beta \circ f \circ \alpha^{-1}\,. \tag{22}$$

Notice that the *coordinate map* $f_{\alpha\beta}$ depends on the choice of basis in V and W. In contrast, the mapping f itself does not refer to any such choice.

The advantage of the representation of f by $f_{\alpha\beta}$ is provided by the following isomorphism between the real vector space $\mathbb{R}^n$ and the real vector space $\mathbb{R}^{n\times 1}$ of matrices of size $n \times 1$:

$$\begin{aligned} \sigma_n : \mathbb{R}^n &\xrightarrow{\simeq} \mathbb{R}^{n\times 1} \\ (x_1, \ldots, x_n) &\mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \end{aligned} \tag{23}$$

columisomorphism

Clearly, for every finite $n > 0$ this mapping is linear and injective. Furthermore, both real vector spaces have the same dimension. Hence, according to the forgoing remark, the mapping σ_n provides an isomorphism. This holds true for any $0 < n < \infty$.

We put $\mathbf{e}_k := \sigma_n(\vec{e}_k) \in \mathbb{R}^{n\times 1}$ for all $1 \le k \le n$ and every $0 < n < \infty$. That is, there is no notational distinction made between $\mathbf{e}_k \in \mathbb{R}^{n\times 1}$ and $\mathbf{e}_k \in \mathbb{R}^{m\times 1}$ even if $n \neq m$. Also we write $\sigma_\alpha : V \xrightarrow{\simeq} \mathbb{R}^{n\times 1}$, $\vec{x} = \sum_{1 \le k \le n} x_k \vec{v}_k \mapsto \sigma_n(x_1, \ldots, x_n) = \sum_{1 \le k \le n} x_k \mathbf{e}_k$. Likewise, we write $\sigma_\beta : W \xrightarrow{\simeq} \mathbb{R}^{m\times 1}$, $\vec{y} = \sum_{1 \le k \le m} y_k \vec{w}_k \mapsto \sigma_m(y_1, \ldots, y_m) = \sum_{1 \le k \le m} y_k \mathbf{e}_k$.

When taking advantage of these mappings, the action of the coordinate map (21) (and thus also of f) can be realized in terms of *matrix multiplication*. In fact, one has

$$\vec{y} = f(\vec{x}) \Leftrightarrow f_{\alpha\beta}(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \Leftrightarrow \sigma_\beta(\vec{y}) = \mathbf{F}\sigma_\alpha(\vec{x})\,, \tag{24}$$

where the matrix $\mathbf{F} \in \mathbb{R}^{m\times n}$ is given by

$$\mathbf{F} := \begin{pmatrix} f_{11} & f_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & f_{1n} \\ f_{21} & f_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & f_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_{m1} & f_{m2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & f_{mn} \end{pmatrix} \tag{25}$$

and

$$f(\vec{v}_j) = f(\sigma_\alpha^{-1}(\mathbf{e}_j)) =: \sum_{1\leq i\leq m} f_{ij}\vec{w}_i \; = \sum_{1\leq i\leq m} f_{ij}\sigma_\beta^{-1}(\mathbf{e}_i) \quad (1 \leq j \leq n)\,. \tag{26}$$

The advantage to realize the action of a linear map by matrix multiplication is indeed the main motivation to identify the vector space $\mathbb{R}^n$, consisting of real $n-$tuples, with the vector space $\mathbb{R}^{n\times 1}$ of real matrices of size $n \times 1$ according to (23). It just simplifies calculations!

Notice that for all $j = 1, \ldots, n$:

$$\sigma_\beta(f(\vec{v}_j)) = \sigma_\beta\big(f(\sigma_\alpha^{-1}(\mathbf{e}_j))\big) = \sum_{1\leq i\leq m} f_{ij}\mathbf{e}_i = \begin{pmatrix} f_{1j} \\ \vdots \\ f_{mj} \end{pmatrix} \in \mathbb{R}^{m\times 1}\,. \tag{27}$$

Hence, the matrix representative of f may be written as

$$\mathbf{F} = (\sigma_\beta(f(\vec{v}_1)), \sigma_\beta(f(\vec{v}_2)), \ldots, \sigma_\beta(f(\vec{v}_n))) \in \mathbb{R}^{m\times n}\,. \tag{28}$$ matrixofmap

Once more, the matrix representative $\mathbf{F}$ of the linear mapping f always refers to the basis chosen in V and W. In contrast, the linear mapping itself does not refer to such a choice. The main reason to introduce $\mathbf{F}$ is that the abstract action of a linear map f can then be simply realized by matrix multiplication.

Example:
Let $V := \{\vec{x} := x_1 + x_2 t + x_3 t^2 \,|\, x_1, x_2, x_3 \in \mathbb{R}\}$ be the three-dimensional real vector space of real polynomials of second order in one variable t. Likewise, we may also consider the two-dimensional real vector space $W := \{\vec{y} := y_1 + y_2 s \,|\, y_1, y_2 \in \mathbb{R}\}$ of real polynomials of first order in one variable s.

Let

$$\begin{aligned} f : V &\longrightarrow W \\ x_1 + x_2 t + x_3 t^2 &\mapsto (x_1 + x_3) + x_2 s\,. \end{aligned} \tag{29}$$ linearmapex

Hence, the linear map is fully determined by

$$f(1) = f(t^2) := 1,\; f(t) := s\,. \tag{30}$$ mapdefex

Clearly, the map f is surjective and $ker(f) = \{\lambda(1 - t^2) \,|\, \lambda \in \mathbb{R}\} \subset V$.

We may set $\sigma_\alpha(1) := \mathbf{e}_1$, $\sigma_\alpha(t) := \mathbf{e}_2$, $\sigma_\alpha(t^2) := \mathbf{e}_3 \in \mathbb{R}^{3\times 1}$ and $\sigma_\beta(1) := \mathbf{e}_1$, $\sigma_\beta(s) := \mathbf{e}_2 \in \mathbb{R}^{2\times 1}$. With respect to this choice the linear map (29) is easily seen to be represented by the matrix

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 3}\,. \tag{31}$$ matrixrepex

Indeed, one gets

$$\mathbf{F}\mathbf{e}_1 = \mathbf{F}\mathbf{e}_3 = \mathbf{e}_1,\; \mathbf{F}\mathbf{e}_2 = \mathbf{e}_2 \in \mathbb{R}^{2\times 1}\,, \tag{32}$$

which corresponds to (30).

We still have to clarify what happens if one changes the basis in V and/or W. For this let $\vec{v}_1, \dots, \vec{v}_n \in V$ and $\vec{v}'_1, \dots, \vec{v}'_n \in V$ be a basis in V. Likewise, let $\vec{w}_1, \dots, \vec{w}_n \in W$ and $\vec{w}'_1, \dots, \vec{w}'_n \in W$ be a basis in W. These choices allow for the following isomorphisms:

$$\begin{aligned} \alpha : V &\xrightarrow{\simeq} \mathbb{R}^n \\ \vec{x} = \sum_{1 \le k \le n} x_k \vec{v}_k &\mapsto (x_1, \dots, x_n)\,, \\ \alpha' : V &\xrightarrow{\simeq} \mathbb{R}^n \\ \vec{x} = \sum_{1 \le k \le n} x'_k \vec{v}'_k &\mapsto (x'_1, \dots, x'_n)\,. \\ \beta : W &\xrightarrow{\simeq} \mathbb{R}^m \\ \vec{y} = \sum_{1 \le k \le m} y_k \vec{w}_k &\mapsto (y_1, \dots, y_m)\,, \\ \beta' : W &\xrightarrow{\simeq} \mathbb{R}^m \\ \vec{y} = \sum_{1 \le k \le n} y'_k \vec{w}'_k &\mapsto (y'_1, \dots, y'_m)\,. \end{aligned} \tag{33}$$

coordisom5

The respective coordinate transformations

$$\begin{aligned} \alpha' \circ \alpha^{-1} : \mathbb{R}^n &\xrightarrow{\simeq} \mathbb{R}^n \\ (x_1, \dots, x_n) &\mapsto (x'_1, \dots, x'_n) \\ \beta' \circ \beta^{-1} : \mathbb{R}^m &\xrightarrow{\simeq} \mathbb{R}^m \\ (y_1, \dots, y_n) &\mapsto (y'_1, \dots, y'_n) \end{aligned} \tag{34}$$

are given by

$$\begin{aligned} x'_i &= \sum_{1 \le j \le n} a_{ij} x_j \qquad (i = 1, \dots, n)\,, \\ y'_i &= \sum_{1 \le j \le m} b_{ij} y_j \qquad (i = 1, \dots, m)\,. \end{aligned} \tag{35}$$

The expansion coefficients a_{ij}, $b_{ij} \in \mathbb{R}$ are defined as

$$\begin{aligned} \vec{v}_j &=: \sum_{1 \le i \le n} a_{ij} \vec{v}'_i \qquad (j = 1, \dots, n)\,, \\ \vec{w}_j &=: \sum_{1 \le i \le m} b_{ij} \vec{w}'_i \qquad (j = 1, \dots, m)\,. \end{aligned} \tag{36}$$

Let again $f : V \longrightarrow W$ be a linear map. According to (21) it can be represented either as

$$\begin{aligned} f_{\alpha\beta} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x_1, \ldots, x_n) &\mapsto (y_1, \ldots, y_m) := \beta(f(\alpha^{-1}(x_1, \ldots, x_n)))\,, \end{aligned} \tag{37}$$

coordinatemaps1

i.e.

$$f_{\alpha\beta} = \beta \circ f \circ \alpha^{-1}\,, \tag{38}$$

or as

$$\begin{aligned} f_{\alpha'\beta'} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x'_1, \ldots, x'_n) &\mapsto (y'_1, \ldots, y'_m) := \beta'(f(\alpha'^{-1}(x'_1, \ldots, x'_n)))\,, \end{aligned} \tag{39}$$

coordinatemaps2

i.e.

$$\begin{aligned} f_{\alpha'\beta'} &= \beta' \circ f \circ \alpha'^{-1} \\ &= \left(\beta' \circ \beta^{-1}\right) \circ f_{\alpha\beta} \circ \left(\alpha' \circ \alpha^{-1}\right)^{-1}. \end{aligned} \tag{40}$$

mapcomp

Similarly to (28) one thus obtains the following two matrices representing f with respect to the chosen basis:

$$\mathbf{F} = (\sigma_\beta(f(\sigma_\alpha^{-1}(\mathbf{e}_1))), \sigma_\beta(f(\sigma_\alpha^{-1}(\mathbf{e}_2))), \ldots, \sigma_\beta(f(\sigma_\alpha^{-1}(\mathbf{e}_n)))) \in \mathbb{R}^{m\times n}\,. \tag{41}$$

matrixofmap1

$$\mathbf{F}' = (\sigma_{\beta'}(f(\sigma_{\alpha'}^{-1}(\mathbf{e}_1))), \sigma_{\beta'}(f(\sigma_{\alpha'}^{-1}(\mathbf{e}_2))), \ldots, \sigma_{\beta'}(f(\sigma_{\alpha'}^{-1}(\mathbf{e}_n)))) \in \mathbb{R}^{m\times n}\,. \tag{42}$$

matrixofmap2

These to matrices $\mathbf{F}, \mathbf{F}' \in \mathbb{R}^{m\times n}$ are different, in general, although they represent the *same* linear map f. In fact, $\mathbf{F}, \mathbf{F}' \in \mathbb{R}^{m\times n}$ are related to each other as

$$\mathbf{F}' = \mathbf{B}\mathbf{F}\mathbf{A}^{-1}\,, \tag{43}$$

matrixmul

where the matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{m\times m}$ are given by

$$\mathbf{A} := \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{m2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix}, \quad \mathbf{B} := \begin{pmatrix} b_{11} & b_{12} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & b_{1m} \\ b_{21} & b_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & b_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & b_{mm} \end{pmatrix}. \tag{44}$$

That is, the matrix $\mathbf{A}$ represents the coordinate transformation $\alpha' \circ \alpha^{-1}$ with respect to the standard basis $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ and $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^{n\times 1}$. Likewise, the matrix $\mathbf{B}$ represents the coordinate transformation $\beta' \circ \beta^{-1}$ with respect to the standard basis $\vec{e}_1, \ldots, \vec{e}_m \in \mathbb{R}^m$ and $\mathbf{e}_1, \ldots, \mathbf{e}_m \in \mathbb{R}^{m\times 1}$. When this matrix representation is taken into account, the composition of maps (40) is realized by matrix multiplication according to (43).

Let us return to our previous **example**. We may put $\vec{v}_1 := 1, \vec{v}_2 := t, \vec{v}_3 := t^2 \in V$ and $\vec{w}_1 := 1, \vec{w}_2 := s \in W$ as a basis. Alternatively, we may consider the basis $\vec{v}'_1 := \vec{v}_1, \vec{v}'_2 := \vec{v}_2 + \vec{v}_3, \vec{v}'_3 := \vec{v}_2 - \vec{v}_3 \in V$ and $\vec{w}'_1 := 2\vec{w}_1, \vec{w}'_2 := \vec{w}_1 + \vec{w}_2 \in W$.

To determine the coordinate transformation one has to calculate the coefficients $a_{ij}, b_{kl} \in \mathbb{R}$. From

$$\begin{aligned} &\vec{v}_1 = \vec{v}'_1\,, \quad \vec{v}_2 = \frac{1}{2}(\vec{v}'_2 + \vec{v}'_3)\,, \quad \vec{v}_3 = \frac{1}{2}(\vec{v}'_2 - \vec{v}'_3)\,. \\ &\vec{w}_1 = \frac{1}{2}\vec{w}'_1\,, \quad \vec{w}_2 = \vec{w}'_2 - \frac{1}{2}\vec{w}'_1 \end{aligned} \tag{45}$$

one infers that

$$
\begin{aligned}
&a_{11} = 1\,, \quad a_{12} = a_{13} = 0\,,\\
&a_{21} = 0\,, \quad a_{22} = a_{23} = 1/2\,,\\
&a_{31} = 0\,, \quad a_{32} = -a_{33} = 1/2\,.\\
&b_{11} = 1/2\,, \quad b_{12} = -1/2\,,\\
&b_{21} = 0\,, \qquad b_{22} = 1\,.
\end{aligned} \tag{46}
$$

Hence,

$$
\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix} \in \mathbb{R}^{3\times 3}\,, \quad \mathbf{B} = \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}\,. \tag{47}
$$

One therefore obtains for the matrix $\mathbf{F}' \in \mathbb{R}^{2\times 3}$:

$$
\begin{aligned}
\mathbf{F}' &= \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 1/2 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}\,.
\end{aligned} \tag{48}
$$

Apparently, this result is quite different from $\mathbf{F}$. Yet both matrices $\mathbf{F}$ and $\mathbf{F}'$ represent the same linear map (29).

Notice that $\mathbf{A}^{-1} \in \mathbb{R}^{3\times 3}$ can be directly read off from the definition of the basis $\vec{v}'_1 \ldots, \vec{v}'_3 \in V$ in terms of the basis $\vec{v}_1 \ldots, \vec{v}_3 \in V$. Also, notice that

$$
\mathbf{F}'\mathbf{e}_1 = \frac{1}{2}\mathbf{e}_1\,, \quad \mathbf{F}'\mathbf{e}_2 = \mathbf{e}_2\,, \quad \mathbf{F}'\mathbf{e}_3 = \mathbf{e}_2 - \mathbf{e}_1 \in \mathbb{R}^{2\times 1}\,, \tag{49}
$$

where, again, the same notation is used for both the standard basis in $\mathbb{R}^{n\times 1}$ and $\mathbb{R}^{m\times 1}$.

Let us introduce one more **notation**: We set $\mathbf{x} := \sigma_\alpha(\vec{x}) \in \mathbb{R}^{n\times 1} = \sum_{1\le k\le n} x_k \mathbf{e}_k$, resp. $\mathbf{x}' := \sigma_{\alpha'}(\vec{x}) = \sum_{1\le k\le n} x'_k \mathbf{e}_k \in \mathbb{R}^{n\times 1}$. Notice that both vectors $\mathbf{x}$ and $\mathbf{x}'$ of the real $n-$dimensional vector space $\mathbb{R}^{n\times 1}$ of $(n\times 1)-$matrices represent the *same* (abstract) vector $\vec{x} \in V$. Likewise, we set $\mathbf{y} := \sigma_\beta(\vec{y}) = \sum_{1\le k\le m} y_k \mathbf{e}_k \in \mathbb{R}^{m\times 1}$, resp. $\mathbf{y}' := \sigma_{\beta'}(\vec{y}) = \sum_{1\le k\le m} y'_k \mathbf{e}_k \in \mathbb{R}^{m\times 1}$. One has

$$
\vec{y} = f(\vec{x}) \quad \Leftrightarrow \quad \left\{ \begin{array}{lcl} \mathbf{y} & = & \mathbf{F}\mathbf{x}\,, \\ \mathbf{y}' & = & \mathbf{F}'\mathbf{x}'\,. \end{array} \right. \tag{50}
$$

twomatequiv

Furthermore, since $\mathbf{x}' = \mathbf{A}\mathbf{x}$ and $\mathbf{y}' = \mathbf{B}\mathbf{y}$ one obtains (43).

Concerning our previous **example**, one gets for the coordinate transformations

$$
\begin{aligned}
\mathbf{x}' &= \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ (x_2 + x_3)/2 \\ (x_2 - x_3)/2 \end{pmatrix} \in \mathbb{R}^{3\times 1}\,, \\
\mathbf{y}' &= \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} (y_1 - y_2)/2 \\ y_2 \end{pmatrix} \in \mathbb{R}^{2\times 1}\,.
\end{aligned} \tag{51}
$$

Indeed, it is straightforward to verify that

$$\begin{aligned}\vec{y} &= \sigma_{\beta'}^{-1}(\mathbf{y}') = \beta'^{-1}(y'_1, y'_2) \\ &= y'_1\vec{w}'_1 + y'_2\vec{w}'_2 = \frac{1}{2}(y_1 - y_2)\vec{w}'_1 + y_2\vec{w}'_2 \\ &= (y_1 - y_2)\vec{w}_1 + y_2(\vec{w}_1 + \vec{w}_2) = y_1\vec{w}_1 + y_2\vec{w}_2 \\ &= \beta^{-1}(y_1, y_2) = \sigma_{\beta}^{-1}(\mathbf{y}) \\ &= (x_1 + x_3)\vec{w}_1 + x_2\vec{w}_2 = f(x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3) \\ &= f(\vec{x})\,,\end{aligned} \tag{52}$$

for all $\vec{x} \in V$. This demonstrates once more that the linear map (29) between *(abstract) vector spaces* V and W can be represented by the different linear maps between the corresponding *coordinate spaces*:

$$\mathbf{F}, \mathbf{F}' : \mathbb{R}^{3\times 1} \longrightarrow \mathbb{R}^{2\times 1}\,. \tag{53}$$

One therefore should be careful when identifying linear maps with matrices! Such an identification always refers to a *choice* of basis!! In more abstract terms, there is no *natural* way, in general, to identify the real vector space $Hom_{\mathbb{R}}(V, W)$ of all linear maps from a real $n-$dimensional vector space V into a real $m-$dimensional vector space W and the real $nm-$dimensional vector space $\mathbb{R}^{m\times n}$ of $(m \times n)-$matrices. In fact, every isomorphism:

$$\begin{aligned} Hom_{\mathbb{R}}(V, W) &\xrightarrow{\simeq} \mathbb{R}^{m\times n} \\ f &\mapsto \mathbf{F}\,, \end{aligned} \tag{54}$$

depends on the arbitrariness of a basis chosen in V and/or W.

Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
E-mail address: Juergen.Tolksdorf@mis.mpg.de