HOMEWORK, WEEK 12

This assignment is due Monday, April 12. Handwritten solutions are acceptable but LaTeX solutions are preferred. You must write in full sentences (abbreviations and common mathematical shorthand are fine).

(1) This problem set will define the tensor product $V \otimes W$ of two vector spaces V and W; we need some preliminary constructions first.

Let V be a vector space and let W be a subspace of V. Define an equivalence relation on V by stating that $v_1 \sim v_2$ if $v_1 - v_2 \in W$. Let V/W (or $\frac{V}{W}$) denote the set of equivalence classes under this relation. If $v \in V$, let $[v]$ denote its equivalence class.

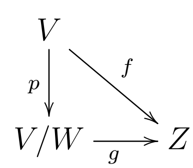
For two elements α, β of V/W, define $\alpha + \beta$ to be $[v_1 + v_2]$ where $\alpha = [v_1]$ and $\beta = [v_2]$. Also, for $c \in \mathbb{R}$, define $c\alpha$ to be $[cv]$ where $\alpha = [v]$.

Problem. Prove that $\alpha + \beta$ and $c\alpha$ are well-defined (i.e. independent of the choice of v_1 and v_2 representing α and β).

Hint: For addition, you want to show (why?) that if v_1, v_2 and v_1', v_2' are such that $v_1 - v_1' \in W$ and $v_2 - v_2' \in W$, then $(v_1 + v_2) - (v_1' + v_2')$ is also in W. Use that W is a subspace of V; scalar multiplication is similar.

Remark. One could check the vector space axioms for V/W to see that V/W is a vector space under the above addition and scalar multiplication. The space V/W is called a *quotient vector space.* If V has dimension n and W has dimension m, then V/W has dimension $n - m$.

(2) Given the setup of the previous problem, let Z be another vector space and let $f : V \to Z$ be a linear map such that $f(w) = 0$ for all $w \in W$. Let $p : V \to V/W$ denote the linear map sending v to $[v]$. Prove that there exists a unique linear map $g : V/W \to Z$ such that $f = g \circ p$, i.e. that the following diagram commutes:



Hint: For existence, define $g([v]) = f(v)$; you want to show that g is well-defined and linear (it follows immediately that the diagram commutes). If $v - v' \in W$, use the fact that $f(v - v') = 0$ by assumption. For linearity, use that $[v] + [v'] := [v + v']$ as defined above; show that g sends such a sum to $g([v]) + g([v'])$ (and similarly for scalar multiples).

For uniqueness, note that any $\alpha \in V/W$ is equal to $[v] = p(v)$ for some $v \in V$. Show that if $g, \tilde{g}$ are two maps making the diagram commute, then $g(\alpha) = \tilde{g}(\alpha)$ for all $\alpha \in V/W$.

(3) Let V and W be vector spaces over $\mathbb{R}$; consider the free vector space $F(V \times W)$ on the Cartesian product $V \times W$ of V and W. Given an element (v, w) of $V \times W$, we view (v, w) as an element of $F(V \times W)$ via the inclusion map $i : V \times W \to F(V \times W)$. Any element of $F(V \times W)$ is a finite linear combination of such elements (v, w).

Warning. $F(V \times W)$ disregards the vector space structures on V and W, and just treats $V \times W$ as the set of ordered pairs (v, w) where $v \in V$ and $w \in W$. For example, if $v \neq 0$ in V, then $(v, 0)$ and $(2v, 0)$ are linearly independent in $F(V \times W)$, even though v and $2v$ are not linearly independent in V.

Let S be the subset of $F(V \times W)$ consisting of the following elements:
- $(v, w) + (v', w) - (v + v', w)$ for all $v, v' \in V$ and $w \in W$
- $(v, w) + (v, w') - (v, w + w')$ for all $v \in V$ and $w, w' \in W$
- $c(v, w) - (cv, w)$ for all $v \in V$, $w \in W$, and $c \in \mathbb{R}$
- $c(v, w) - (v, cw)$ for all $v \in V$, $w \in W$, and $c \in \mathbb{R}$.

Definition. Define
$$V \otimes W := \frac{F(V \times W)}{\operatorname{span}(S)}.$$

Given an element (v, w) of $F(V \times W)$, its class $[(v, w)]$ in the quotient space $V \otimes W$ will be denoted by $v \otimes w$.

Warning. In general, not every element of $V \otimes W$ can be written as $v \otimes w$ for some $v \in V$ and $w \in W$. Elements like $v \otimes w$ are called "pure tensors." In general, an element of $V \otimes W$ is a linear combination of pure tensors, but it might not be a pure tensor itself. In quantum mechanics, impure tensors correspond to *entangled* states.

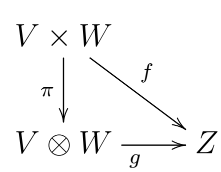
Recall that if V, W, Z are vector spaces over $\mathbb{R}$, a map $f : V \times W \to Z$ is said to be *bilinear* if:
- $f(v + cv', w) = f(v, w) + cf(v', w)$ for all $v, v' \in V$, $w \in W$, and $c \in \mathbb{R}$
- $f(v, w + cw') = f(v, w) + cf(v, w')$ for all $v \in V$, $w, w' \in W$, and $c \in \mathbb{R}$.

Problem. Show that the map $\pi : V \times W \to V \otimes W$ sending (v, w) to $v \otimes w$ is bilinear.

Hint: To see that $\pi(v + cv', w) = \pi(v, w) + c\pi(v', w)$, show that $(v + cv', w) - (v, w) - c(v', w)$ is in the span of the set S. The other equation is similar (you don't need to write out the proof again).

(4) In this problem you will prove the "universal property" of tensor products: roughly speaking, a bilinear map from $V \times W$ to Z is the same as a linear map from $V \otimes W$ to Z.

Problem. Let V, W, and Z be vector spaces over $\mathbb{R}$ and let $f : V \times W \to Z$ be a bilinear map. Show that there exists a unique linear map $g : V \otimes W \to Z$ such that $f = g \circ \pi$, i.e. such that the following diagram commutes:



Hint: For existence, use the previous homework to upgrade the (set-theoretic) map $f : V \times W \to Z$ into a linear map from $F(V \times W)$ to Z. Show that this linear map vanishes on elements of S, so that it descends to a linear map from $V \otimes W$ to Z.

For uniqueness, show that if g and $\tilde{g}$ are two linear maps satisfying the above conditions, then g and $\tilde{g}$ agree when applied to any pure tensor $v \otimes w$. Indeed, we have $v \otimes w = [(v, w)]$, so $g(v \otimes w) = g(\pi(v, w)) = f(v, w)$, and similarly for $\tilde{g}$. Conclude (using that any tensor is a linear combination of pure tensors) that $g = \tilde{g}$.

Remark. It's a good exercise to show, using this universal property, that we have $V \otimes W \cong W \otimes V$ for any vector spaces V and W (as well as other standard properties like $(V \otimes W) \otimes Z \cong V \otimes (W \otimes Z)$). One can also deduce relations like $V \otimes (W_1 \oplus W_2) \cong (V \otimes W_1) \oplus (V \otimes W_2)$, i.e. "tensor products distribute over direct sums" (up to isomorphism). The vector space $\mathbb{R} = \mathbb{R}^1$ acts as a "unit" for the tensor product operation; for any V, we have $V \otimes \mathbb{R}^1 \cong V$. (You don't need to prove any of these properties on the homework, but they are good exercises.)

It follows that if V has dimension n, so $V \cong \mathbb{R}^n \cong \mathbb{R} \oplus \cdots \oplus \mathbb{R}$ (n times), and W has dimension m, then $V \otimes W \cong (\mathbb{R} \oplus \cdots \oplus \mathbb{R}) \otimes (\mathbb{R} \oplus \cdots \oplus \mathbb{R})$, which expands out as a direct sum of nm copies of $\mathbb{R}$. Thus, $V \otimes W$ has dimension nm. Concretely, if $\{e_i\}$ forms a basis for V and $\{f_j\}$ forms a basis for W, then $\{e_i \otimes f_j\}$ forms a basis for $V \otimes W$.

Remark. Mathematicians often define a vector to be an element of a vector space. Correspondingly, one can define a *tensor of rank k* to be an element of a vector space of the form $V_1 \otimes \cdots \otimes V_k$ for some $k \geq 1$ (i.e. a tensor is a "vector in the space'$V_1 \otimes \cdots \otimes V_k$"). In quantum mechanics, if we have k isolated systems that don't communicate with each other and each has V_i as its (complex) Hilbert space of states (for $1 \leq i \leq k$), then the Hilbert space for the composite system is $V_1 \otimes \cdots \otimes V_k$. One may be able to prepare this system in an "entangled state," i.e. a state that can't be written as a single pure tensor $v_1 \otimes \cdots \otimes v_k$ (it must be a nontrivial sum of pure tensors); for example, you could start by allowing the k different physical systems to interact and then separate them later.

A useful example to keep in mind is in quantum computing, where k is the "number of qubits" and each V_i is the two-dimensional vector space $\mathbb{C}^2$ (with its standard inner product). Following Dirac's "bra-ket" notation, one often labels the two basis vectors of $\mathbb{C}^2$ as $|0\rangle$ and $|1\rangle$ (think of these as the quantum analogues of the two possible values "zero" and "one" of a classical bit). Then, for example, $\mathbb{C}^2 \otimes \mathbb{C}^2$ has basis

$$\{|0\rangle \otimes |0\rangle, |1\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |1\rangle\},$$

which is often written in shorthand as

$$|00\rangle, |10\rangle, |01\rangle, |11\rangle.$$

The state

$$\frac{1}{\sqrt{2}}|10\rangle + \frac{1}{\sqrt{2}}|01\rangle$$

of this 2-qubit system is an entangled state; it cannot be written as a pure tensor. If we simultaneously measure the values of the first and second qubits (even after they have been separated by an arbitrary distance), we will get either $(1,0)$ or $(0,1)$, not $(0,0)$ or $(1,1)$.

Remark. In some contexts (see e.g. `https://www.tensorflow.org/guide/tensor`), it is common to present tensors as being "higher-dimensional analogues of vectors and matrices." If a vector is a one-dimensional array of numbers and a matrix is a two-dimensional array, a tensor of rank k should be a k-dimensional array. One can connect this view on tensors with our definitions above as follows: given bases $\{e_{i,j}\}_j$ for each V_i, one gets a basis for $V_1 \otimes \cdots \otimes V_k$ by taking tensor products of basis vectors as discussed above. An arbitrary element of $V_1 \otimes \cdots \otimes V_k$ can be expanded as a linear combination of basis vectors; one has a unique coefficient $c_{j_1,\dots,j_k}$ on each basis vector $e_{1,j_1} \otimes \cdots \otimes e_{k,j_k}$. One organizes these coefficients into an array of the appropriate dimensionality:

- For $k = 1$, the coefficients c_j form a one-dimensional array (vector),
- For $k = 2$, the coefficients c_{j_1,j_2} form a two-dimensional array (matrix),
- For $k = 3$, the coefficients c_{j_1,j_2,j_3} form a three-dimensional array (rank-3 tensor), etc.

Once we introduce dual vector spaces on next week's HW, we will often prefer to view a matrix as representing a linear transformation between vector spaces $T : V \to W$, and then interpret this transformation as an element of $W \otimes V^*$ (still a rank-2 tensor, but involving a dual on one of the factors). See also the next remark.

Remark. The above definitions make no distinction between "covariant," "contravariant," and "mixed" tensors. This extra distinction is natural when we have fixed vector spaces $V_1, \dots, V_k$ and we consider tensor products of the V_i and/or their dual spaces V_i^*. In this setting, an element x of $V_1 \otimes \cdots \otimes V_k$ is called a "covariant tensor" and we write its coefficient on a basis vector $e_{1,j_1} \otimes \cdots \otimes e_{k,j_k}$ as $x^{j_1,\dots,j_k}$ (i.e. we use "up indices"). We also have "contravariant tensors" which are elements x of $V_1^* \otimes \cdots \otimes V_k^*$; we write its coefficient on the dual basis vector $e^{1,j_1} \otimes \cdots \otimes e^{k,j_k}$ as $x_{j_1,\dots,j_k}$ (i.e. we use "down indices"). In between these extreme cases, we have various types of mixed tensors whose coefficients have both up and down indices.

This up/down convention is part of "Einstein notation," developed by Albert Einstein for use in relativity (this notation is widespread in physics, and it's good for mathematicians to be familiar with too). Another part of the convention is to write only the coefficients and not the basis vectors they're multiplied by when specifying a tensor. For example:

- A vector $x = \sum_i x^i e_i$ in V gets written as x^i,
- A dual vector $x = \sum_i x_i e^i \in V^*$ gets written as x_i,
- A rank-2 covariant tensor $\sum_{i,j} x^{i,j} e_i \otimes f_j$ in $V \otimes W$ gets written as $x^{i,j}$ (where $\{f_j\}$ is the given basis for W),
- A rank-2 mixed tensor $\sum_{i,j} x^i_j e_i \otimes f^j$ in $W \otimes V^*$ gets written as x^i_j,
- A rank-2 contravariant tensor $\sum_{i,j} x_{i,j} e^i \otimes f^j$ in $V^* \otimes W^*$ gets written as $x_{i,j}$.

Often one will consider certain sums of these coefficients; the convention is that one sums over repeated indices (one must be up and the other down), and omits the sum symbol. For example, let a^i_j represent an element of $W \otimes V^*$ (corresponding to a linear transformation T from V to W; in fact, in the assumed bases, a^i_j is the usual matrix entry of T in row i and column j). Mathematically, we would write a^i_j as

$\sum_{i,j} a^i_j e_i \otimes f^j$. Let v^i represent a vector in V; mathematically, we would write v^i as $\sum_i v^i e_i$. Applying the transformation a^i_j to the vector v^i gives

$$\sum_{i,j} a^i_j v^j f_i$$

(a vector in W), and the convention is to write this vector as simply

$$a^i_j v^j$$

(the sum over j is implicit, as is the sum over i with the "invisible" basis vectors f_i). You know this quantity represents a vector since it has one "free" index i and this index is up.

Remark. The meaning of "covariant" and "contravariant" above is based on how the coefficients of the tensors transform under changes of basis on the underlying vector spaces. For notational simplicity, assume the V_i are all the same vector space V, with basis $\{e_i\}$, and suppose we are changing to a new basis $\{e'_i\}$. We can write $e'_j = \sum_i a^i_j e'_i$ where the coefficients a^i_j form an invertible matrix. Correspondingly, if

$$x = \sum_{i_1,\dots,i_k} x^{i_1,\dots,i_k} e_{i_1} \otimes e_{i_k},$$

then we equivalently have

$$x = \sum_{j_1,\dots,j_k,i_1,\dots,i_k} x^{j_1,\dots,j_k} a^{i_1}_{j_1} \cdots a^{i_k}_{j_k} e'_{j_1} \otimes e'_{j_k}.$$

In Einstein notation, $x^{i_1,\dots,i_k}$ is transformed to $x^{j_1,\dots,j_k} a^{i_1}_{j_1} \cdots a^{i_k}_{j_k}$. If you have some coefficients $x(i_1,\dots,i_k)$ and they transform this way under changes of basis, then they should be written with "up indices" like $x^{i_1,\dots,i_k}$. Similarly, write $e'_i = \sum_j (a^{-1})^j_i e'_j$; if $x(i_1,\dots,i_k)$ transforms to

$$x(j_1,\dots,j_k)(a^{-1})^{j_1}_{i_1} \cdots (a^{-1})^{j_k}_{i_k},$$

then x can be viewed as a contravariant tensor of rank k and one writes $x_{i_1,\dots,i_k}$ in Einstein notation.

Remark. A typical situation is where each V_i is $\mathbb{R}^n$ for some (fixed) n, equipped with some basis (more generally, each V_i is $(TM)_p$ for some smooth manifold M and some p). In both mathematics and physics, one often considers "tensor–valued functions" or "tensor fields," i.e. maps from some $U \subset \mathbb{R}^n$ into $\mathbb{R}^n \otimes \cdots \otimes \mathbb{R}^n$ (more generally, one considers sections of tensor products of the tangent bundle of M). Such a function f can be thought of as an indexed collection of real-valued functions $f^{j_1,\dots,j_k}$ (this is still true locally in the general setting of manifolds and vector bundles). If we allow $(\mathbb{R}^n)^*$ (or the cotangent bundle T^*M) in our tensor products, then the functions f may have both upper and lower indices. For example, differential forms can be viewed as certain sections of tensor powers of the cotangent bundle T^*M (so their indices are down in Einstein notation). In this context, the transformation rules for tensors are often viewed in terms of changes of basis coming from coordinate changes on the manifold M. When dealing specifically with tensor fields on spacetime, Einstein notation uses the additional convention that Roman indices are "spacial indices" ($i, j \in \{1, 2, 3\}$) while Greek indices are "spacetime indices" ($\mu, \nu \in \{0, 1, 2, 3\}$).