A brief explanation of the Steinitz Exchange Lemma

Suppose we have a set of linearly independent vectors $I = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\}$, and a set of spanning vectors $S = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$. Then we may replace, one at a time, vectors in S with vectors in I, until we get a spanning set $T = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{s}\}$, where $\mathbf{s}$ is one of the $\mathbf{s}_i$.

—

Since S spans, we know that $\mathbf{i}_1$ may be written as a linear combination of elements of S, say $\mathbf{i}_1 = \lambda_1\mathbf{s}_1 + \lambda_2\mathbf{s}_2 + \lambda_3\mathbf{s}_3 + \lambda_4\mathbf{s}_4$. One of the λ_i must be non-zero, or else we have $\mathbf{i}_1 = 0$, which contradicts I being linearly independent. So, by renaming the $\mathbf{s}_i$ if necessary, we may assume that $\lambda_1 \neq 0$.

We have $\mathbf{s}_1 = \dfrac{1}{\lambda_1}\big(\mathbf{i}_1 - \lambda_2\mathbf{s}_2 - \lambda_3\mathbf{s}_3 - \lambda_4\mathbf{s}_4\big)$.

Then $S' = \{\mathbf{i}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$ spans. Why? We know that S spans, so any vector is some combination of $\mathbf{s}_1, \ldots, \mathbf{s}_4$, and we can now replace the occurrence of $\mathbf{s}_1$ by the above expression in $\mathbf{i}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4$.

We now have the linearly independent set $I' = \{\mathbf{i}_2, \mathbf{i}_3\}$ and the spanning set $S' = \{\mathbf{i}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$.

Since S' spans, we know that $\mathbf{i}_2$ may be written as a linear combination of elements of S', say $\mathbf{i}_2 = \mu_1\mathbf{i}_1 + \mu_2\mathbf{s}_2 + \mu_3\mathbf{s}_3 + \mu_4\mathbf{s}_4$. This time, we know that one of μ_2, μ_3, μ_4 must be non-zero, or else we have $\mathbf{i}_2 = \mu_1\mathbf{i}_1$, which contradicts the original I being linearly independent. So, by renaming the $\mathbf{s}_i$ if necessary, we may assume that $\mu_2 \neq 0$.

We have $\mathbf{s}_2 = \dfrac{1}{\mu_2}\big(\mathbf{i}_2 - \mu_1\mathbf{i}_1 - \mu_3\mathbf{s}_3 - \mu_4\mathbf{s}_4\big)$.

Then $S'' = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{s}_3, \mathbf{s}_4\}$ spans. Why? Like before, we know that S' spans, so that any vector is some combination of $\mathbf{i}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4$, and we can now replace the occurrence of $\mathbf{s}_2$ with the above expression in $\mathbf{i}_1, \mathbf{i}_2, \mathbf{s}_3, \mathbf{s}_4$.

We now have the linearly independent set $I'' = \{\mathbf{i}_3\}$ and the spanning set $S'' = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{s}_3, \mathbf{s}_4\}$.

Since S'' spans, we know that $\mathbf{i}_3$ may be written as a linear combination of elements of S'', say $\mathbf{i}_3 = \nu_1\mathbf{i}_1 + \nu_2\mathbf{i}_2 + \nu_3\mathbf{s}_3 + \nu_4\mathbf{s}_4$. This time, we know that one of ν_3, ν_4 must be non-zero, or else we have $\mathbf{i}_3 = \nu_1\mathbf{i}_1 + \nu_2\mathbf{i}_2$, which contradicts the original I being linearly independent. So, by renaming the $\mathbf{s}_i$ if necessary, we may assume that $\nu_3 \neq 0$.

We have $\mathbf{s}_3 = \dfrac{1}{\nu_3}\big(\mathbf{i}_3 - \nu_1\mathbf{i}_1 - \nu_2\mathbf{i}_2 - \nu_4\mathbf{s}_4\big)$.

Then $S''' = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{s}_4\}$ spans, for reasons just like before. And this is the set T we claimed existed at the start.

—

What would have happened if we'd started with the linearly independent set $I = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\}$ and the spanning set $S = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$?

We would do one more step in the process above, and end up replacing all elements in S with those in I, and we would conclude that I itself spans.

What if we'd started with the linearly independent set $I = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4, \mathbf{i}_5\}$ and the spanning set $S = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$?

Well, after four steps in the above process we would have reached the point where $I'''' = \{\mathbf{i}_5\}$ and $S'''' = \{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4\}$ were the sets under consideration. But this is impossible, for if this S'''' spanned, then we could write $\mathbf{i}_5$ as a linear combination of $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{i}_4$, which would contradict I being linearly independent in the first place.

—

It is clear that if I is a finite linearly independent set and S is a finite spanning set, then the procedure above could be performed. In particular, we would find that $|I| \leqslant |S|$.

An immediate corollary is that if a vector space has a finite basis, then any two bases for it have the same size. For if B_1 and B_2 are two bases for it, then taking $I = B_1$ and $S = B_2$, we see that $|B_1| \leqslant |B_2|$, and taking $I = B_2$ and $S = B_1$, we see that $|B_2| \leqslant |B_1|$.

This means that 'dimension' is well-defined.

We can also see that, in a finite-dimensional vector space, any linearly independent set may be extended to a basis: let I be the linearly independent set and let S be any basis you like, and at the end of the process we have reached a basis containing I.

—

Note the use of the word 'finite' a few times above. The process described did assume that I and S were finite. We could have taken S to be infinite, and replaced $|I|$ many elements of S with the elements of I. That's okay, because the process would still terminate after $|I|$ many steps.

It's not obvious what would happen if I were infinite. If the exchange process turns out to be invalid, then the corollaries might not follow. Is 'dimension' even well-defined for huge spaces, and does every vector space actually have a basis? For the answers, you'll have to go to *Logic & Set Theory* in the third year and throw Zorn's Lemma at the problem.

Please let me know of any corrections: glt1000@cam.ac.uk