PROOF OF A LINEAR ALGEBRA LEMMA, WITH LINKS TO THE CORRESPONDING LEAN 4 CODE

PAUL D. NELSON

We formalize in Lean 4 a proof of [3, Theorem 1.8]:

Theorem 1. *(`MainConcrete'` in MainConcrete)*
Let M_n denote the space of complex $n \times n$ matrices, included in the space M_{n+1} of $(n+1) \times (n+1)$ matrices as the upper-left block, e.g., for $n = 2$, as

$$\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let τ be an element of M_{n+1} with the property that no eigenvalue of τ is also an eigenvalue of the upper-left $n \times n$ submatrix τ_0 of τ. Let $x \in M_{n+1}$ with $[x, \tau] = 0$, where $[a, b] := ab - ba$. Let z denote the image in M_{n+1} of the identity element of M_n, thus $z = \mathrm{diag}(1, \dots, 1, 0)$ with n ones. Suppose that

$$[x, [z, \tau]] = [y, \tau]$$

for some $y \in M_n$. Then x is a scalar matrix.

In this note, we record a proof and indicate which parts correspond to which Lean files and theorems.

Remark 2. The original proof reduced to a determinental identity [3, Theorem 17.2] that was verified using some geometric invariant theory. We decided to formalize a related but more elementary argument noted in [2, Remark 5.15].

Theorem 1 extends (with the same proof) to algebraically closed fields of characteristic $\neq 2$ (`MainConcrete` in MainConcrete). It generalizes further to commutative rings in which 2 is a unit, and may be formulated in the language of endomorphisms of finite free modules (see [2, Remark 5.15]):

Theorem 3. *(`MainAbstract` in MainAbstract)*
Let R be a nontrivial commutative ring in which 2 is a unit. Let V be a finite free module over R. Let $\tau \in \mathrm{End}_R(V \times R)$, with projection $\tau_0 \in \mathrm{End}_R(V)$. Assume that the characteristic polynomials of τ and τ_0 are coprime.

Let $x \in \mathrm{End}_R(V \times R)$ with $[x, \tau] = 0$, let z denote the image under the extension-by-zero map $\mathrm{End}_R(V) \to \mathrm{End}_R(V \times R)$ of the identity endomorphism of V (corresponding to the matrix $\mathrm{diag}(1, \dots, 1, 0)$), and suppose that

$$[x, [z, \tau]] = [y, \tau]$$

for some $y \in \mathrm{End}_R(V)$, extended by zero to an element of $\mathrm{End}_R(V \times R)$. Then x is a scalar endomorphism.

Date: 18 January 2024.

Remark 4. We have formalized only the special case of Theorem 3 in which R is a field. This case suffices for Theorem 1, which was our motivating goal. To deduce the general case would require one further formalization, namely that of Lemma 9, below.

To verify that the hypotheses of Theorem 1 translate to those of Theorem 3, we use the following:

Lemma 5. *(*`coprime_of_disjoint_roots` *in CoprimeOfDisjointRoots)*
Let R be an algebraically closed field. Let $p, q \in R[X]$ be nonzero polynomials with no common root. Then p and q are coprime, that is to say, we can write $1 = ap + bq$ with $a, b \in R[X]$.

Formalizing the passage from Theorem 3 to 1 requires checking that various operations are compatible with the passage from $V \times R$ to R^{n+1} obtained by choosing a basis of V. This is done, laboriously, in MainConcrete.

It remains to explain the proof of (the field case of) Theorem 1. This requires a couple further lemmas.

Lemma 6. *(*`cyclic_e_of_coprime_charpoly` *and* `cyclic_e'_of_coprime_charpoly` *in CyclicOfCoprime)*
Let R be a nontrivial commutative ring. Let V be a finite free R-module. Let $\tau \in \mathrm{End}_R(V \times R)$ have the property that its and its projection to $\mathrm{End}_R(V)$ have coprime characteristic polynomials. Then the vector $(0, 1) \in V \times R$ and the dual vector $(0, 1)^t \in (V \times R)^*$ are cyclic with respect to τ, i.e., we can write every vector in the respective module as a polynomial in τ applied to those vectors.

Remark 7. For the same reasons as in Remark 4, Lemma 6 has only been fully formalized over a field.

Proof. Let us show that $(0, 1)^t$ is cyclic; a similar argument gives the other assertion. We may reduce to the case that R is a field using Lemma 9, below (that lemma has not yet formalized, but is irrelevant if we restrict from the outset to the field case).

Let $W \leq (V \times R)^*$ denote the $R[\tau]$-submodule generated by $(0, 1)^t$; our task is to show that $W = (V \times R)^*$.

Let $U \leq V \times R$ denote the coannihilator of W (i.e., the kernel of the natural map $V \times R \to W^*$). By duality for finite free modules over fields, it is equivalent to check that $U = 0$. Since W contains $(0, 1)^t$, we see that U is contained in the first summand $V \hookrightarrow V \times R$. Furthermore, it is clear by construction that U is an $R[\tau]$-submodule of $V \times R$.

We have reduced to verifying the following (`no_invariant_subspaces_of_coprime_charpoly` in CyclicOfCoprime):

- if U is an $R[\tau]$-submodule of $V \times R$ contained in the first summand $V \hookrightarrow V \times R$, then $U = 0$.

To that end, let $u \in U$; we must show that $u = 0$. Let $\tau_0 \in \mathrm{End}_R(V)$ denote the projection of τ, and let $p, p_0 \in R[X]$ denote the characteristic polynomials of τ, τ_0. By hypothesis, we may write $1 = ap_\tau + bp_{\tau_0}$ for some $a, b \in R[X]$. By evaluating this abstract identity of polynomials on τ and appealing to the consequence $p(\tau) = 0$ of the Cayley–Hamilton theorem, we see that $u = b(\tau)p_0(\tau)u$. On the other hand, since u lies in the $R[\tau]$-submodule U of $V \hookrightarrow V \times R$, we see by induction on the degree of p that $p_0(\tau)u = p_0(\tau_0)u$. By another appeal to Cayley–Hamilton, we have $p_0(\tau_0) = 0$. It follows that $u = 0$, as required. □

Lemma 8. *(*`injective_of_cyclic` *and* `injective_of_cyclic'` *in InjectiveOfCyclic)*
Let R be a nontrivial commutative ring. Let V be a finite free R-module (in practice, this will play the role of the module "$V \times R$" appearing above). Let $\tau \in \mathrm{End}_R(V)$

(i) If $e \in V$ is cyclic with respect to τ, then the map

$$\{x \in \mathrm{End}_R(V) : [x, \tau] = 0\} \to V$$

$$x \mapsto xe$$

is injective.

(ii) Dually, if $e^* \in V^*$ is cyclic with respect to τ, then the map

$$\{x \in \mathrm{End}_R(V) : [x, \tau] = 0\} \to V^*$$

$$x \mapsto e^* x$$

is injective.

Proof. We verify (i), as the proof of (ii) is similar. It is enough to show that, for $x \in \mathrm{End}_R(V)$ with $[x, \tau] = 0$ and $xe = 0$, we have $x = 0$. It is enough to show that $xv = 0$ for each $v \in V$. By hypothesis, we may write $v = p(\tau)e$ for some $p \in R[X]$. Since x commutes with τ, we see by induction on the degree of p that it also commutes with $p(\tau)$. Thus $xv = xp(\tau)e = p(\tau)xe = 0$, as required. □

Having established Lemmas 5 and 6, we are now in position to complete the proof of Theorem 3.

Proof of Theorem 3. (see MainAbstract)
Set $t := [x, z] - y$. Then $[x, z] = t + y$, and by 3, we have $[t, \tau] = 0$.
Recall that $z \in \mathrm{End}_R(V \times R)$ is the extension by zero of the identity endomorphism of the first summand V (in matrix language, this is $\mathrm{diag}(1, \dots, 1, 0)$). Let $e := (0, 1) \in V \times R$ and $e^* := (0, 1)^t \in (V \times R)^*$ denote, respectively, the inclusion of the identity element in the second summand and the projection onto the second factor. We may identify $ee^* := e \otimes e^*$ with the endomorphism of $V \times R$ given by extension by zero of the identity endomorphism of the second summand R (in matrix language, this is $\mathrm{diag}(0, \dots, 0, 1)$). It follows that $z + ee^*$ is the identity endomorphism of $V \times R$, and in particular, commutes with x. We thus have $[x, ee^*] = -t - y$. Let us evaluate this last identity on the vectors e and e^*. Recalling that y is the extension by zero of an endomorphism of V, we have $ye = 0$ and $e^* y = 0$, hence

$$[x, ee^*]e = -te$$

and

$$e^*[x, ee^*] = -e^* t.$$

On the other hand, by expanding commutator brackets and using that $e^* e = 1$, we compute that

$$[x, ee^*]e = x(ee^*)e - (ee^*)xe = (xe) - e(e^* xe) = (x - e^* xe)e,$$

where here $e^* xe \in R$ is identified with a scalar endomorphism of $V \times R$, and similarly

$$e^*[x, ee^*] = e^* x(ee^*) - e^*(ee^*)x = (e^* xe)e^* - e^* x = -e^*(x - e^* xe).$$

By Lemmas 5 and 6, the maps defined on the centralizer of τ given by multiplication against e or e^* are injective. Since both $x - e^*xe$ and t centralize τ, we deduce from the above four identities that

$$x - e^*xe = -t = -(x - e^*xe),$$

which implies that $2x = 2(e^*xe)$. By our hypothesis that 2 is a unit, we conclude that x is a scalar endomorphism. □

This completes our discussion of the proof of the field case of Theorem 3, hence that of Theorem 1.

The following would suffice to extend the formalization of Theorem 3 to a general commutative ring in which 2 is a unit:

Lemma 9. *Let R be a commutative ring. Let M be a finitely-generated R-module, and let $N \leq M$ be a submodule. Assume that $N + \mathfrak{m}M = M$ for each maximal ideal $\mathfrak{m}$ of R. Then $N = M$.*

Proof. We can check whether $N = M$ after localizing at each maximal ideal $\mathfrak{m}$ of R [1, Prop 3.9], so it suffices to consider the case of a local ring R with maximal ideal $\mathfrak{m}$. We can then appeal to Nakayama's lemma [1, Cor 2.7]. □

Remark 10. Mathlib contains suitable formulations of Nakayama's lemma. The main outstanding ingredient needed to formalize Lemma 9 is the generalization of `ideal_eq_bot_of_localization` from ideals to modules.

References

[1] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Series in Mathematics. Westview Press, Boulder, CO, economy edition, 2016. For the 1969 original see [MR0242802].

[2] Yueke Hu and Paul D Nelson. Subconvex bounds for $u_{n+1} \times u_n$ in horizontal aspects. 09 2023. arXiv:2309.06314.

[3] Paul D. Nelson. Spectral aspect subconvex bounds for $U_{n+1} \times U_n$. *Invent. Math.*, 232(3):1273–1438, 2023.