Fundamental Theorem of Algebra via Linear Algebra

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

Recall the fundamental theorem of algebra, FTA for short.

Theorem 1. *Let $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ be a polynomial with complex coefficients. Then it has a zero in $\mathbb{C}$, that is, there exists a $\lambda \in C$ such that $P(\lambda) = 0$.* □

One encounters the fundamental theorem of algebra in a Linear Algebra course while trying to prove the following result.

Theorem 2. *Let V be a finite dimensional vector space over $\mathbb{C}$. Let $A\colon V \to V$ be linear. Then A has an eigen-value, that is, there exists a nonzero vector $v \in V$ and a complex number $\lambda \in \mathbb{C}$ such that $Av = \lambda v$.* □

The way one proves this theorem is to rewrite the equation $Av = \lambda v$ as $(A - \lambda I)v = 0$ and interpret it as the kernel of $(A - \lambda I)$ is nontrivial. It follows that $(A - \lambda I)$ is singular or what is the same $\det(A - \lambda I) = 0$. Thus, λ is an eigen value of A iff it is a root of the characteristic polynomial of A. As $\mathbb{C}$ is algebraically closed, we deduce that we can always find such a λ. Thus the theorem is a consequence of the fundamental theorem of algebra.

What is equally well-known is that if we can prove Theorem 2 somehow independent of the FTA, then FTA can be deduced from it. This is easily seen as follows. Given a monic polynomial (that is the coefficient of the leading term is 1), $X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, one considers the so-called companion matrix of the polynomial:

$$A := \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}.$$

Then $\det(XI - A)$ is the given polynomial, as can seen by induction on n. (This may also be seen as follows. Consider the matrix $(XI - A)$:

$$\begin{pmatrix} X & 0 & \dots & 0 & a_0 \\ -1 & X & \dots & 0 & a_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & -1 & X + a_{n-1} \end{pmatrix}.$$

Multiply the last row by X and add it to the $(n-1)$-th row, and in the resulting matrix, multiply the $(n-1)$-th row by X and add it to the $(n-2)$-th row and so on. At the $(n-1)$-th step, we end up with a matrix of the form

$$\begin{pmatrix} 0_{1\times(n-1)} & p(X) \\ -I_{(n-1)\times(n-1)} & * \end{pmatrix}.$$

By the properties of determinants, it follows that $\det(XI - A)$ is the determinant of the above matrix which is clearly $p(X)$.)

Now if we assume that Theorem 2 is true, applying it to A we deduce FTA.

In fact, Derksen proves a seemingly stronger result and deduces Theorem 2 and hence FTA.

Theorem 3. *If $A_1, \dots, A_r$ are commuting endomorphisms of a finite dimensional vector space over $\mathbb{C}$, then they have a common eigen-vector, that is, there exists a nonzero vector $v \in V$ and $\lambda_j \in \mathbb{C}$ for $1 \le j \le r$ such that $A_j v = \lambda_j v$ for $1 \le j \le r$.* □

To achieve this, he uses a slightly more involved induction. In stead, we use his argument to prove Theorem 2 in a more direct way. Theorem 3 can be deduced from Theorem 2.

Like any proof of FTA, this also needs some results from analysis/topology.

Lemma 4. *Any polynomial of odd degree with real coefficients has a real zero.*

Proof. It is enough to prove that a monic polynomial

$$P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1 X + a_0, (a_j \in \mathbb{R}, 0 \le j \le n-1),$$

of odd degree has a real zero. If we set $\alpha := 1 + |a_0| + \dots + |a_{n-1}|$, then it is easy to see that $P(\alpha) > 0$ and $P(-\alpha) < 0$. Intermediate value theorem assures us of the existence of some $\lambda \in (-\alpha, \alpha)$ such that $P(\lambda) = 0$.

A more qualitative argument would run as follows. Write

$$P(X) = X^n(1 + \frac{a_{n-1}}{X} + \dots + \frac{a_0}{X^n}).$$

Then as $X \to \pm\infty$, the terms of the form $\frac{a_j}{X^{n-j}} \to 0$ for $0 \le j \le n-1$. Since they are finite in number, for $|X|$ sufficiently large, we have an estimate for the sum $1 + \frac{a_{n-1}}{X} + \dots + \frac{a_0}{X^n}$:

$$1/2 \le 1 + \frac{a_{n-1}}{X} + \dots + \frac{a_0}{X^n} \le 3/2.$$

Consequently, $P(X) \le X^n/2 < 0$ if X is sufficiently large negative number and $P(X) \ge X^n/2 > 0$ if X is sufficiently large positive number. Now complete the proof as before. □

Lemma 5. *Let n be odd. Let V be an n-dimensional vector space over $\mathbb{R}$. Assume that $A, B\colon V \to V$ be two commuting endomorphims. Then they have a common eigen vector.*

Proof. We prove the result by induction on the odd dimension of V. If $\dim V = 1$, then the result is obvious, as any nonzero vector is an eigen vector for both the operators.

Let n be odd. Assume that the result is true for all vector spaces of odd dimension less than n. As a consequence of Lemma 4, A will have an eigen value, say, λ. Let

$$K = \mathrm{Ker}\,(A - \lambda I) \text{ and } R = \mathrm{Im}\,(A - \lambda I).$$

Both these spaces remain invariant under any endomorphism B of V that commutes with A. (Note that $(A - \lambda I)$ commutes with A and that any such B commutes with $A - \lambda I$.) For, if $x \in R$, then $x = (A - \lambda I)v$ for some $v \in V$. Now,

$$Bx = B(Av - \lambda v) = A(Bv) - \lambda Bv = (A - \lambda I)Bv \in R.$$

Similarly, if $x \in K$, then

$$(A - \lambda I)Bx = B(A - \lambda I)x = B0 = 0.$$

Note that $\dim K \geq 1$. Now, either $K = V$ or K is a proper subspace of V. In the first case, since K is odd dimensional, B will have an eigen vector in K. This vector is an eigen vector of A also with eigen value λ.

Let $\dim K < n$. By rank-nullity theorem, $\dim K + \dim R = n$ and so exactly one of $\dim K$ and $\dim R$ must be odd. Whichever has dimension odd, its dimension will be strictly less than n. Hence by induction, both A and B will have a common eigen vector on that subspace. □

Lemma 6. *Let $\mathbb{F}$ be any field. Fix a positive integer $d > 1$. Assume that any endomorphism on any vector space over the field $\mathbb{F}$ whose dimension is not divisible by d has an eigen vector. Then any pair of commuting linear maps on V whose dimension is not divisible by d has a common eigen vector.*

Proof. The proof is exactly similar to that of the last lemma and hence omitted. □

We are now ready to prove Theorem 2.

Proof. The proof will be by induction on the highest power of 2 dividing the dimension of the vector space. Thus, we would have proved the result for vector spaces of dimension 101 even before we prove the result for $n = 2$!

As the starting point, we consider the case when V is odd dimensional complex vector space. This corresponds to $k = 0$, that is, 2^0 is the highest power of 2 dividing the dimension n. Let us consider the *real* vector space $\mathbf{H}_n$ of all hermitian matrices of order n. Its dimension is $2 \times \left(\frac{n(n-1)}{2}\right) + n = n^2$. Hence it is odd dimensional. We consider two linear maps

$$T_1(B) := \frac{AB + BA^*}{2} \text{ and } T_2(B) := \frac{AB - BA^*}{2i}.$$

One easily verifies that T_1 and T_2 are linear maps on $\mathbf{H}_n$ and that they commute. Hence by Lemma 5, they have a common eigen vector. Let B be a (non-zero) common eigen vector, say, $T_1 B = \lambda B$ and $T_2 B = \mu B$. It follows that

$$AB = T_1(B) + iT_2 B = (\lambda + i\mu)B.$$

Since B is nonzero, when considered as a linear transformation on V (with respect to a choice of an ordered basis), there exists a vector $v \in V$ such that $Bv \neq 0$. But then, $ABv = (\lambda + i\mu)Bv$ says that Bv is an eigen vector of A. (Alternatively, if we take as v any non-zero column vector of B, then it is an eigen vector of A with eigen value $\lambda + i\mu$.)

We now proceed to the inductive step. Let $k \geq 1$. The inductive hypothesis is that for any n, where the highest power of 2 dividing n is less than 2^k, any linear map on any n dimensional complex vector space has an eigen vector. As a consequence of Lemma 5, it follows that any pair of commuting linear maps on a vector space of such a dimension will have a common eigen vector.

Let V be an n-dimensional vector space and $T\colon V \to V$ be linear. Assume that 2^k is the maximum power of 2 that divides n. We need to show that T has an eigen vector. Fix an ordered basis of V. Let A be the matrix of T with respect to this basis.

We modify the proof in the starting inductive step. We now consider the set $\mathbf{S}_n$ of complex symmetric matrices of order n. It is a complex vector space of dimension $\frac{n(n+1)}{2}$. We consider two maps

$$L_1(B) = AB + BA^t \qquad \text{and} \qquad L_2(B) = ABA^t,$$

where C^t denotes the transpose of the matrix C. It is easy to verify that they are linear maps on $\mathbf{S}_n$ and that they commute. Thanks to our hypothesis on n, 2^{k-1} is the maximum power of 2 that divides the dimension of $\mathbf{S}_n$. Hence by what we said above, there exists a common eigen vector $B \in \mathbf{S}_n$ of L_1 and L_2, say,

$$L_1 B = AB + BA^t = \lambda B \qquad \text{and} \qquad L_2 B = ABA^t = \mu B.$$

Using the value of BA^t from the first equation, we get

$$\mu B = L_2 B = A(\lambda B - AB).$$

Hence $(A^2 B - \lambda AB + \mu B) = 0$ or what is the same

$$(A^2 - \lambda A + \mu I)B = 0.$$

Consider the quadratic equation $X^2 - \lambda X + \mu = 0$ in $\mathbb{C}$. By the standard formula for the roots of quadratic equation and since square roots of any complex number exists in $\mathbb{C}$, we can factorize $(X^2 - \lambda X + \mu) = (X - \alpha)(X - \beta)$. If we plug in A in place of X, we see that

$$(A - \alpha I)(A - \beta I)B = 0.$$

If $(A-\beta I)B = 0$, then any nonzero column of B will be an eigen vector of A. If $(A-\beta I)B \neq 0$, then any of its nonzero column is an eigen vector of A. $\square$

Remark 7. Theorem 3 can be now proved in a way similar to the proof in Lemma 5.

Reference: H.Derksen, "The fundamental Theorem of Algebra and Linear Algebra", Amer. Math. Monthly, **110**, (2003), 620-623.

Derksen's article is available from:
`http://www.math.lsa.umich.edu/ hderksen/preprint.html`