Crystal's Proof about the kth root of a matrix PhL 2.1.12

I struggled, as Geoff would say, with rolling my own proof but could not make it go. I then found the following 2007 theorem and proof in a master's thesis by one Crystal Monterz Gordon of Georgia State U. It uses a certain polynomial ingredient that I think makes the uniqueness proof work.

[I wonder if she concocted this proof herself or found the proof somewhere. Yes, she says she adapted her proof from Ref 8 which is Horn and Johnson, *Matrix Analysis*, 1985 which I am now downloading just for my reserve stacks. It is a text searchable djvu. In fact Crystal's proof is not "adapted", it is cribbed word for word and symbol for symbol from pp 405-406 of that book.]

I will paste and comment:

First, here is a statement of her theorem:

> **Theorem 4.3:** Let $A \in M_n$ be positive semidefinite and let $k \geq 1$ be a given integer. Then there exists a unique positive semidefinite Hermitian matrix B such that $B^k = A$. We also have
>
> (a) $BA = AB$ and there is a polynomial $p(t)$ such that $B = p(A)$;
>
> (b) rank B = rank A, so B is a positive definite if A is; and
>
> (c) B is real if A is real.

So in her notation, B is the small matrix, so to speak. I only cared about k = 2, but fine to do general k. The key idea here for me is the addition of uniqueness to the theorem. So we start out:

> **Proof:** We know that the Hermitian matrix A can be unitarily diagonalized as $A = UDU^*$ with $D = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ and all $\lambda_i \geq 0$. We define $B = UD^{\frac{1}{k}}U^*$, where $D^{1/k} \equiv \text{diag}(\lambda_1^{1/k}, \lambda_2^{1/k}, ..., \lambda_n^{1/k})$, and the unique nonnegative kth root is taken in each case. Clearly, $B^k = A$ and B is Hermitian and positive semidefinite. Also, $AB = UDU^*UD^{\frac{1}{k}}U^* = UDD^{\frac{1}{k}}U^* = UD^{\frac{1}{k}}DU^* = UD^{\frac{1}{k}}U^*UDU^* = BA$, and B is positive semidefinite because all λ_i (and hence their kth roots) are nonnegative.

Comments:

(1) As with most non-QM papers, the Hermitian adjoint is indicated by * and not † and it is CC with transpose. I agree with the opening statement that Hermitian A can be diagonalized by unitary U, and I know in fact that such U could be made from columns as normalized eigenfunctions of A. So

$A = UDU^*$ her D is my Λ $U^* = U^{-1}$ unitary

$A = A^*$ Hermitian

(2) If A is positive semi-definite, then yes, all the eigenvalues are positive or zero, so $\lambda_n \geq 0$ as she says.

If we define

$$B \equiv UD^{1/k}U^* = UD^{1/k}U^{-1}$$

Then certainly we get

$$B^k = UD^{1/k}\,U^{-1}\,UD^{1/k}\,U^{-1}\;\; UD^{1/k}U^* \;\ldots\; = \; U\,DU^* = A$$

so right off the bat we have found at least one root of A, it is $B \equiv UD^{1/k}U^*$. We show in a second that this root B is positive definite and Hermitian.

Of course to write this B out, we have to actually diagonalize Hermitian A and *find* a matrix U and *find* the eigenvalues λ_i which make up the matrix D.

Now notice that

$$B = UD^{1/k}U^* \quad \Rightarrow \quad B^* = U^{**}D^{1/k*}U^* = U\,D^{1/k}\,U^* = B \quad \Rightarrow \; B = B^*$$

so our root B is a Hermitian matrix. But why is it positive definite? Well,

$$B = UD^{1/k}U^{-1} \quad \Rightarrow \quad BU = UD^{1/k}$$

$$\Rightarrow \; B\,\{\ldots x_i \ldots\} = \{\ldots x_i \ldots\}\,D^{1/k} \; = \; \{\;\ldots\ldots\; \lambda_i{}^{1/k} x_i \ldots\}$$

$$\Rightarrow B\,x_i = \lambda_i{}^{1/k} x_i \qquad i = 1..N$$

Thus, B has all non-negative eigenvalues (we assume we took $\lambda_i{}^{1/k}$ non-negative) so B is positive definite.

So at this point we have found at least one Hermitian positive semidefinite k^{th} root of our Hermitian positive semidefinite matrix A. An excellent start.

Then we have

$$AB = (U\,DU^{-1})\,(UD^{1/k}U^{-1}) \; = U\,DD^{1/k}U^{-1} \; = U\,D^{1/k}DU^{-1}$$

$$BA = (UD^{1/k}U^{-1})\,(U\,DU^{-1}) = U\,D^{1/k}DU^{-1}$$

where we use the fact that two diagonal D type matrices compute. So we know that

[A,B] = 0

This was an extra thing she is proving along the way. So done with the above. Moving along,

> The rank of B is just the number of nonzero λ_i terms, which is also the rank of A. If A is real and positive semidefinite, then we know that U may be chosen to be a real orthogonal matrix, so it is clear that B can be chosen to be real in this case. It remains only to consider the question of uniqueness.

I know that both matrices B and $A=B^k$ have the same eigenfunctions: since B is Hermitian, we can say $B\,x_i = \lambda_i^{1/k} x_i$ as noted above, and then you just write

$$B^k x_i = B^{k-1} B x_i = B^{k-1}\ \lambda_i^{1/k} x_i = \ldots\ = \lambda_i\, x_i$$

If A has M non-zero EV's, then so does B since they are related to each other trivially, and therefore A and B have the same rank, one of her added items.

If A is purely real, then U will be real-orthogonal, and $B = UD^{1/k}U^*$ will also be pure real.

Now comes the uniqueness section. We begin with,

> Notice first that there is a polynomial $p(t)$ such that $p(A) = B$; we need only choose $p(t)$ the Lagrange interpolating poynomial for the set $\{(\lambda_1, \lambda_1^{\frac{1}{k}}), \ldots, (\lambda_n, \lambda_n^{\frac{1}{k}})\}$ to get $p(D) = D^{\frac{1}{k}}$ and $p(A) = p(UDU^*) = Up(D)U^* = UD^{\frac{1}{k}}U^* = B$. But

I had to go look up this "Lagrange interpolating formula". I then realized this thing appeared in my Sheid numerical math Schaum book in Chapter 8 there, where it was called " Lagrange's formula for the collocation polynomial" with spacing of the x_i that need not be equal. My notes are pretty clear there. In general, you can fit N points with a polynomial of degree N-1. So you see her list of N points above, and then p(t) is the solution polynomial, whatever it is. Thus, we have $p(\lambda_n) = \lambda_n^{1/k}$ for each root, since these are the points we have fit, and that then means that $p(D) = D^{1/k}$ if you just combine things. Then:

$$p(A) = p(UDU^*) = \Sigma_i\, a_i\, (UDU^{-1})^i = U\,[\Sigma_i\, a_i\, D]\, U^{-1} = U\,\rho(D)\,U^{-1} = U\,D^{1/k}\,U^{-1} = B$$

So we get this interesting fact that p(A) = B (where $A=B^k$) for our polynomial function p(t) which fits the N points as shown above. She does not mention the degree of p(t). If some of the EV's are the same, then there are fewer points to fit, and then p(t) will have degree lower than N-1, but the degree of p(t) really plays no role in what follows, so that is why she did not mention it.

We continue, where C is now going to be our "other" assumed root with the same properties as B:

then if C is any positive semidefinite Hermitian matrix such that $C^k = A$, we have $B = p(A) = p(C^k)$ so that $CB = Cp(C^k) = p(C^k)C = BC$. Since B and C are commuting Hermitian matrices, they may be simultaneously unitarily

So repeating the argument above we find that

$A = C^k \Rightarrow B = p(A) = p(C^k)$

$CB = C\,p(C^k) = p(C^k)\,C = BC$ // since C commutes with any polynomial of C

$\Rightarrow [B,C] = 0$

So I agree that if B and C are two *different* positive semidefinite Hermitian k^{th} roots of A, they must commute. This was the big fact that I was missing in my own proof attempt, I did not know how to show it. But my proof was only for k = 2 and positive definite, so throw out mine and take hers!

diagonalized; that is, there is some unitary matrix V and diagonal matrices D_1 and D_2 with nonnegative diagonal entries such that $B = VD_1V^*$ and $C = VD_2V^*$. Then from the fact that $B^k = A = C^k$ we deduce that $D_1^k = D_2^k$. But since the nonnegative kth root of a nonnegative number is unique, we conclude that $(D_1^k)^{\frac{1}{k}} = D_1 = D_2 = (D_2^k)^{\frac{1}{k}}$ and $B = C$. ■

I agree with the first two lines here just from my historical knowledge of QM and I surely have a proof of that fact somewhere. We then have

$B^k = (VD_1V^{-1})^k = V\,D_1^k\,V^{-1}$
$C^k = (VD_2V^{-1})^k = V\,D_2^k\,V^{-1}$

$\Rightarrow V\,D_1^k\,V^{-1} = V\,D_2^k\,V^{-1} \Rightarrow D_1^k = D_2^k$

For diagonal matrices where we always took positive roots, this means $D_1 = D_2$. But then B = C using formulas above, and the two assumed different matrices *must be the same*, so the solution is unique. QED