f(A) Part 2

Suppose $f : \mathbb{C} \to \mathbb{C}$ is a continuous function. We want to investigate here how it makes sense to "extend" f to a function $f : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$. We know that f in the former sense is a different kind of thing than in the latter sense, but we will use f do denote them both, nonetheless.

This explaination is based on the *Jordan canonical form* (JCF) of a square matrix. A *Jordan block* is a $k \times k$ (say) matrix of the form

$$J = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{bmatrix} = \lambda I + N.$$

This matrix has one eigenvalue, λ of multiplicity k, and an associated eigenspace of dimension 1 (only one eigenvector for λ). N is *nilpotent*: $N^k = 0$. Every square matrix A is similar to a block diagonal matrix with Jordan blocks on its diagonal, called the JCF of A, i.e. there exists $V \in \mathbb{C}^{n \times n}$ such that

$$V^{-1}AV = J = \begin{bmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & J_{m-1} & 0 \\ 0 & 0 & \dots & 0 & J_m \end{bmatrix}.$$

If all of the Jordan blocks are 1×1, then the JCF is the diagonal matrix we dealt with in part 1. But in order get a general definition for $f(A)$, we need to define $f(J)$ for an arbitrary Jordan block. Again, we are led by polynomials. We will need this lemma (which you should prove yourself (the summand index could begin at k)):

$$p(x) = \sum_{j=0}^{n} a_j x^j \quad \Longrightarrow \quad \frac{d^k}{dx^k}\, p(x) \equiv p^{(k)}(x) = k! \sum_{j=0}^{n} \binom{j}{k} a_j x^{j-k}.$$

Now suppose $J = \lambda I + N$ is $k \times k$. Then $J^m = (\lambda I + N)^m = \sum_{i=0}^{m} \binom{m}{i} \lambda^{m-i} N^i$, and

$$\begin{aligned} p(J) &= \sum_{j=0}^{n} a_j \,[\, \sum_{i=0}^{j} \binom{j}{i} \lambda^{j-i} N^i \,] \\ &= \sum_{i=0}^{m} N^i \,[\, \sum_{j=0}^{n} \binom{j}{i} a_j \lambda^{j-i} \,] \\ &= \sum_{i=0}^{k} N^i \,[\, p^{(i)}(\lambda)/i! \,]. \end{aligned}$$

This matrix is upper triangular and constant along each diagonal. The constant along the ith diagonal $p^{(i)}(\lambda)/i!$. Of course, for $f(J)$ we should replace p with f.

We now see that $f : \mathbb{C} \to \mathbb{C}$ must have n derivatives for $f(A)$ to make sense on the set of all $n \times n$ matrices. We have shown that if f is analytic (differentiable on all of $\mathbb{C}$), then $f(A)$ defined above is a continuous function of the entries of A.