The goal of this section is to prove the portion of Dahlquist's theorem that gives the convergence of LMMs. We do not consider the necessity of the conditions we state.

Consider solving

$$\begin{cases} u'(t) = f(u(t)), \\ u(0) = \eta \in \mathbb{R}^n, \end{cases}$$

on the domain $\mathcal{D} = \{(u,t) : \|u-\eta\| \leq a, |t| \leq T\}$. We will have two norms to consider for which we will use the same notation

$$\|u\|_\infty = \max_{0\leq t\leq T} |u(t)|, \quad \|f\|_\infty = \max_{\eta - a \leq u \leq \eta + a} \|f(u)\|_2.$$

Which norm is being used will be clear from context. The method we will use is a LMM

$$\sum_{j=0}^{r} \alpha_j U^{n+j} = k \sum_{j=0}^{r} \beta_j f(U^{n+j}), \quad \alpha_r = 1. \tag{0.1}$$

Suppose that the LTE satisfies

$$\|\tau^n_{\text{LMM}}\|_2 \leq C_1 \|u^{(p+1)}\|_\infty k^p,$$

for a constant $C_1 > 0$. Suppose further that

$$\begin{aligned} f^{(j)}, \quad j = 0,1,2,\ldots,p, &\text{ is continuous on } [\eta - a, \eta + a], \\ \|f^{(p)}\|_\infty \leq C_2, & \end{aligned}$$

for another constant C_2.

0.1 ▪ A proof in the scalar case

We will assume that $u(t) \in \mathbb{R}$ and comment on how one extends the arguments. Define

$$\underline{U}^n = \begin{bmatrix} U^{n+r-1} \\ \vdots \\ U^{n+1} \\ U^n \end{bmatrix}, \quad f(\underline{U}^n) = \begin{bmatrix} f(U^{n+r-1}) \\ \vdots \\ f(U^{n+1}) \\ f(U^n) \end{bmatrix}.$$

Then the LMM (0.1) can be written as

$$\underline{U}^{n+1} = \underbrace{\begin{bmatrix} -\alpha_{r-1} & -\alpha_{r-2} & \cdots & & -\alpha_0 \\ 1 & & & & 0 \\ & 1 & & & 0 \\ & & \ddots & & \vdots \\ & & & 1 & 0 \end{bmatrix}}_{A} \underline{U}^n + kF(\underline{U}^{n+1}, \underline{U}^n), \tag{0.2}$$

where

$$F(\underline{U}^{n+1}, \underline{U}^n) = \left[\beta_r e_1^T f(\underline{U}^{n+1}) + \begin{bmatrix} \beta_{r-1} & \beta_{r-2} & \cdots & \beta_0 \end{bmatrix} f(\underline{U}^n)\right] e_1,$$

and e_1 is the first standard basis vector. For the case where U^n is a vector, this would have to be a block matrix with blocks of the same form down the diagonal. We also use the notation

$$\underline{u}(t_n) = \begin{bmatrix} u(t_{n+r-1} \\ \vdots \\ u(t_n) \end{bmatrix}.$$

The assumption on the LTE implies that

$$\underline{u}(t_{n+1}) = A\underline{u}(t_n) + kF(\underline{u}(t_{n+1}), \underline{u}(t_n)) + k\tau^n_{\text{LTE}} e_1.$$

Now, define

$$\underline{E}^n = \underline{U}^n - \underline{u}(t_n),$$

which satisfies the iteration

$$\underline{E}^{n+1} = A\underline{E}^n + \underbrace{k\left(F(\underline{U}^{n+1}, \underline{U}^n) - F(\underline{u}(t_{n+1}), \underline{u}(t_n))\right) - k\tau^n_{\text{LMM}} e_1.}_{k\Delta^n_F}$$

At this point, it might be tempting to write

$$\|\underline{E}^{n+1}\|_2 \le \|A\|_2 \|\underline{E}^n\|_2 + k\|\Delta^n_F\|_2,$$

but we can see by example that this will not be sufficient. For Adams-Bashforth

$$A = \begin{bmatrix} -1 & 0 & \cdots & & 0 \\ 1 & 0 & \cdots & & 0 \\ & 1 & & \vdots & \\ & & \ddots & & \\ & & & 1 & 0 \end{bmatrix},$$

which satisfies $\|A\|_2 = \sqrt{2}$. It turns out that this would give the estimate,

$$\|\underline{E}^n\|_2 \le 2^{n/2} \|\tau^n_{\text{LMM}}\|_\infty.$$

This will not give convergence. This matrix A has eigenvalues in the closed unit disk but the eigenvector matrix has a condition number that is larger than 1, and this causes the problem. So, one could either find a new norm in which to measure things, not the 2-norm, or change the approach slightly. We will take the latter approach. This approach involves working with equalities for as along as possible:

$$\begin{aligned}
&\underline{E}^0 \text{ given} \\
&\underline{E}^1 = A\,\mathrm{e}^0 + k\Delta^0_F, \\
&\underline{E}^2 = A^2\,\mathrm{e}^0 + kA\Delta^0_F + k\Delta^1_F, \\
&\underline{E}^3 = A^3\,\mathrm{e}^0 + kA^2\Delta^0_F + kA\Delta^1_F + k\Delta^2_F, \\
&\quad\vdots \\
&\underline{E}^n = A^n\underline{E}^0 + k\sum_{j=0}^{n-1} A^{n-1-j}\Delta^j_F.
\end{aligned}$$

Now, to avoid arguments invoking the implicit function theorem, we will assume that $\beta_r = 0$ and the method under consideration is explicit. Then

$$F(\underline{U}^{n+1}, \underline{U}^n) = G(\underline{U}^n) = \left(\begin{bmatrix}\beta_{r-1} & \beta_{r-2} & \cdots & \beta_0\end{bmatrix} f(\underline{U}^n)\right) e_1.$$

Note that if f is Lipschitz with constant L then so is G, in the 2-norm, with constant

$$L' = L \sum_{j=1}^{r-1} |\beta_j|.$$

This implies that

$$\|\Delta_F^j\|_2 \le L' \|\underline{E}^n\| + \|\tau_{\text{LMM}}^n\|_\infty,$$

where $\|\tau_{\text{LMM}}^n\|_\infty$ gives the largest truncation error that is encountered in the course of the iteration, up to $n = N$. We then arrive at

$$\|\underline{E}^n\|_2 \le \|A^n\|_2 \|\underline{E}^0\|_2 + k \sum_{j=0}^{n-1} \|A^{n-1-j}\|_2 (L' \|\underline{E}^j\|_2 + \|\tau_{\text{LMM}}^n\|_\infty). \tag{0.3}$$

Now, suppose the method under consideration is zero-stable. We compute, using a cofactor expansion across the first row

$$\begin{aligned}
\det(\lambda I - A) &= \det \begin{bmatrix} \lambda + \alpha_{r-1} & \alpha_{r-2} & \cdots & & \alpha_0 \\ -1 & \lambda & & & 0 \\ & -1 & \lambda & & 0 \\ & & \ddots & & \vdots \\ & & & -1 & \lambda \end{bmatrix} \\
&= (\lambda + \alpha_{r-1})\lambda^{r-1} - \alpha_{r-1}(-\lambda^{r-2}) + \alpha_{r-2}\lambda^{r-2} + \cdots \\
&= \rho(\lambda).
\end{aligned}$$

Note that we used that $\alpha_r = 1$ in this calculation. Note that zero-stability then implies that all the eigenvalues of A are within the unit disk, $|\lambda| \le 1$, and if $|\lambda| = 1$ then the eigenvalue is simple. We now prove a simple but important lemma.

Lemma 0.1. *Suppose $A \in \mathbb{R}^n$ satisfies:*

- *If λ is an eigenvalue of A then $|\lambda| \le 1$, and*
- *if $|\lambda| = 1$ is an eigenvalue of A then λ is simple.*

Then there exists a constant C such

$$\|A^n\|_2 \le C,$$

for all $n \ge 0$.

Proof. This proof makes use of the ideas in Appendix **??**. Let $A = RJR^{-1}$ the matrix decomposition into Jordan canonical form. We can order the columns of R so that

$$J = \left[\begin{array}{cccc|c} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_m & \\ \hline & & & & J_1 \end{array}\right],$$

where $\lambda_1, \dots, \lambda_m$ are distinct eigenvalues with modulus 1. Then the remaining eigenvalues, the diagonal entries of J_1, must lie within the disk $D_\epsilon = \{z : |z| \le 1 - \epsilon\}$ for some $\epsilon > 0$. Then it is clear that

$$J^n = \left[\begin{array}{cccc|c} \lambda_1^n & & & & \\ & \lambda_2^n & & & \\ & & \ddots & & \\ & & & \lambda_m^n & \\ \hline & & & & J_1^n \end{array}\right]$$

We estimate $\|J_1^n\|_2$ by considering

$$J_1^n = \frac{1}{2\pi\,\mathrm{i}} \int_{\partial D_\epsilon} z^n (zI - J_1)^{-1} \mathrm{d}z,$$
$$\|J_1^n\|_2 \le \frac{|\partial D_\epsilon|}{2\pi} \max_{z \in \partial D_\epsilon} |z^n| \|(zI - J_1)^{-1}\|_2.$$

We can bound

$$\frac{|\partial D_\epsilon|}{2\pi} < 1,$$
$$|z^n| = 1 - \epsilon,$$
$$\max_{z \in \partial D_\epsilon} \|(zI - J_1)^{-1}\|_2 = C_3,$$

where C_3 is some finite constant. There are many such ways to combine these estimates, but one way is to write

$$J^n = \left[\begin{array}{c|c} & \\ \hline & J_1^n \end{array}\right] + \left(J^n - \left[\begin{array}{c|c} & \\ \hline & J_1^n \end{array}\right]\right),$$

and bound

$$\left\| \left[\begin{array}{c|c} & \\ \hline & J_1^n \end{array}\right] \right\|_2 \le C_3 (1-\epsilon)^n, \quad \left\| J^n - \left[\begin{array}{c|c} & \\ \hline & J_1^n \end{array}\right] \right\|_2 = 1.$$

These combine to give

$$\|A^n\| \le \|R\|_2 \|R^{-1}\|_2 (1 + C_3 (1-\epsilon)^n),$$

which gives the result.

Returning to (0.3), let $n \le N, kN = T$

$$\begin{aligned} \|\underline{E}^n\|_2 &\le C\|\underline{E}^0\|_2 + k \sum_{j=0}^{n-1} \left[CL' \|\underline{E}^j\|_2 + \|\tau^n_{\mathrm{LMM}}\|_\infty \right] \\ &\le C\|\underline{E}^0\|_2 + CT\|\tau^n_{\mathrm{LMM}}\|_\infty + kCL' \sum_{j=0}^{n-1} \|\underline{E}^j\|_2. \end{aligned}$$

This does not give us what we need yet as it is only a bound on $\underline{E}^n$ in terms of the previous $\underline{E}^j$'s. So, define a sequence that dominates $\|\underline{E}^n\|_2$

$$\begin{aligned} Z^0 &= C\|\underline{E}^0\|_2 + CT\|\tau^n_{\mathrm{LMM}}\|_\infty, \\ Z^n &= C\|\underline{E}^0\|_2 + CT\|\tau^n_{\mathrm{LMM}}\|_\infty + kDT \sum_{j=0}^{n-1} Z^j. \end{aligned}$$

By induction $\|\underline{E}^n\|_2 \leq Z^n$. And we find

$$\begin{aligned} Z^{n+1} - Z^n = kCL'Z^n \Rightarrow Z^{n+1} &= (1 + kCL')Z^n, \\ Z^n &= (1 + kCL')^n Z^0, \\ Z^n &\leq \mathrm{e}^{nkCL'} Z^0 \leq \mathrm{e}^{TDL} Z_0. \end{aligned}$$

So, this tells us that

$$\max_{0 \leq j \leq N} \|\underline{E}^j\|_2 \leq \mathrm{e}^{TDL} \left(C\|\underline{E}^0\|_2 + CT\|\tau^n_{\mathrm{LMM}}\|_\infty \right),$$

which proves the (uniform) convergence of the method at the same rate as the LTE.

Remark 0.2. *This method also proves the convergence of onestep methods on the same system, by setting $r = 1$. The extension to systems is straightforward following the remarks made here.*