Jason Starr
Fall 2011

# MAT 544 Problem Set 6 Solutions

## Problems.

**Problem 1** Let $(W, \|\bullet\|_W)$ be a Banach space (the most important case is $W = \mathbb{R}^n$). Let $I$ be a bounded, open interval in $\mathbb{R}$ with closure $\overline{I}$, and let

$$F : \overline{I} \to L(W, W), \quad t \mapsto (F_t : W \to W)$$

be a bounded, continuous function. For every $t_0 \in I$, consider the initial value problem

$$\frac{dA}{dt}(t) = F_t \circ A(t), \quad A(t_0) = \mathrm{Id}_W$$

where $A$ is a continuously differentiable function from some open neighborhood of $t_0$ in $I$ to $L(W, W)$. As proved in lecture, there is a unique solution $A_{t_0}(t)$. Denote this by $A(t, t_0) = A_{t_0}(t)$, called a *Green's function.*

**(a)** For fixed $t_0, t_1 \in I$, check that both of the following functions

$$\tilde{A}(t) = A(t, t_0), \quad \widehat{A}(t) = A(t, t_1) \circ A(t_1, t_0).$$

solve the initial value problem

$$\frac{dA}{dt}(t) = F_t \circ A(t), \quad A(t_1) = A(t_1, t_0),$$

and thus are equal by uniqueness. In particular, conclude that $A(t_1, t_0)$ and $A(t_0, t_1)$ are inverse (bounded) linear operators.

**(b)** Let $U$ be an element in $L(W, W)$ which has an inverse $U^{-1}$ in $L(W, W)$. Check that $B(t, t_0) = U \circ A(t, t_0) \circ U^{-1}$ is a solution of the initial value problem

$$\frac{dB}{dt}(t) = (U \circ F_t \circ U^{-1}) \circ B(t), \quad B(t_0) = \mathrm{Id}_W.$$

**(c)** Now let $\vec{g} : \overline{I} \to W$ be a continuous function and consider the initial value problem

$$\frac{d\vec{x}}{dt} = F_t(\vec{x}(t)) + \vec{g}(t), \quad \vec{x}(t_0) = 0.$$

where $\vec{x}(t)$ is a continuous map $\overline{I} \to W$ which is continuously differentiable on $I$. This equation is called an *inhomogeneous* linear ODE. Check that the following formula gives one solution (which is unique)

$$\vec{x}(t) = A(t,t_0) \circ \int_{t_0}^{t} A(s,t_0)^{-1} \circ \vec{g}(s)ds = A(t,t_0) \circ \int_{t_0}^{t} A(t_0,s) \circ \vec{g}(s)ds =$$

$$\int_{t_0}^{t} A(t,s) \circ \vec{g}(s)ds.$$

**Problem 2** Let $(W, \| \bullet \|_W)$ be a Banach space, e.g., $W = \mathbb{R}^n$. Let $f(z) = \sum_{n=0}^{\infty} c_n z^n$ be an absolutely convergent series with positive radius of convergence $R$, i.e.,

$$\sum_{n=0}^{\infty} |c_n| R^n < \infty.$$

**(a)** For every $A$ in the closed ball $B_{\leq R}(0)$ in $L(W,W)$, prove that the sequence of partial sums

$$\sum_{n=0}^{N} c_n A^n$$

converges to a limit. Call this limit $f_{L(W,W)}(A)$.

**(b)** Prove further that the associated map

$$f_{L(W,W)} : B_{\leq R}(0) \to L(W,W), \quad A \mapsto f_{L(W,W)}(A)$$

is continuous. (Consider this as a uniform limit of polynomial functions.)

**(c)** For every $U$ in $L(W,W)$ with inverse $U^{-1}$ also in $L(W,W)$, prove that $f_{L(W,W)}(UAU^{-1})$ equals $Uf_{L(W,W)}(A)U^{-1}$.

**Problem 3** Read about how to find the Jordan normal form, particularly for $2 \times 2$, $3 \times 3$ and $4 \times 4$ matrices. One short review is available from the lecture notes at the following URL: `http://ocw.mit.edu/courses/mathematics/18-034-honors-differential-equations-spring-2004/`

**Problem 4** Consider the following second order differential equation with initial values.

$$\frac{d^2}{dt^2}x(t) - 6\frac{d}{dt}x(t) + 9x(t) = 0, \quad x(t_0) = b_0, \ x'(t_0) = b_1$$

**(a)** Find a $2 \times 2$ matrix $A$ such that for every choice of $b_0$,$b_1$, the unique solution of the initial value problem

$$\frac{d}{dt}\vec{x}(t) = A\vec{x}(t), \quad \vec{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \end{bmatrix}, \quad \vec{x}(t_0) = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}.$$

gives a solution of the second order differential equation by $x(t) = x_0(t)$.

**(b)** Find an invertible $2 \times 2$ matrix $U$ such that $AU = U\tilde{A}$ where $\tilde{A}$ is a $2 \times 2$ matrix of the form

$$\tilde{A} = \tilde{S} + \tilde{N}$$

where $\tilde{S}$ is a diagonal matrix, $\tilde{N}$ is a strictly upper triangular matrix (with zeroes on the diagonal), and $\tilde{S}\tilde{N} = \tilde{N}\tilde{S}$.

**(c)** Compute $\exp(\tilde{S}(t-t_0))$, $\exp(\tilde{N}(t-t_0))$ and

$$\exp(\tilde{A}(t-t_0)) = \exp(\tilde{S}(t-t_0))\exp(\tilde{N}(t-t_0)).$$

The compute $\exp(A(t-t_0))U = U\exp(\tilde{A}(t-t_0))$.

**(d)** Let $c_0$, $c_1$ be real numbers such that

$$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = U \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}.$$

Find the solution of the initial value problem with respect to $c_0$ and $c_1$, and use this to find the solution of the original second order differential equation with respect to $c_0$ and $c_1$.

**Problem 5** For the same differential equation as in **Problem 4**, use **Problem 1(c)** to solve the following inhomogeneous equation

$$\frac{d^2}{dt^2}x(t) - 6\frac{d}{dt}x(t) + 9x(t) = e^{3t}, \quad x(t_0) = x'(t_0) = 0.$$

There are other methods to solve this problem (such as the method of undetermined coefficients). You may use these methods to check your work, but please write up the solution using the Green's function as in **Problem 1(c)**.

**Solutions to Problems.**

**Solution to (1)** There are some results about linear maps. Each one is elementary. Some have been explicitly proved in lecture. It seems best to state them explicitly. Let $(U, \|\bullet\|_U)$, $(V, \|\bullet\|_V)$ and $(W, \|\bullet\|_W)$ be normed vector spaces.

**Lemma 0.1.** *(i) For every bounded linear map $S : U \to V$ and for every bounded linear map $T : V \to W$, the composition $T \circ S : U \to W$ is a bounded linear map and $\|T \circ S\|_{op} \leq \|T\|_{op} \cdot \|S\|_{op}$.*

*(ii) For every bounded linear map $S : U \to V$, the map $C_{\bullet,S} : L(V,W) \to L(U,W)$ by $C_{\bullet,S}(T) = T \circ S$ is a bounded linear map with $\|C_{\bullet,S}\|_{op} \leq \|S\|_{op}$. Similarly, for every bounded linear map $T : V \to W$, the map $C_{T,\bullet} : L(U,V) \to L(U,W)$ by $C_{T,\bullet}(S) = T \circ S$ is a bounded linear map with $\|C_{T,\bullet}\|_{op} \leq \|T\|_{op}$.*

*(iii) The linear map $C_{\bullet,*} : L(U,V) \to L(L(V,W), L(U,W))$ by $S \mapsto C_{\bullet,S}$ is a bounded linear map with operator norm $\leq 1$. Similarly the linear map $C_{*,\bullet} : L(V,W) \to L(L(U,V), L(U,W))$ by $T \mapsto C_{T,\bullet}$ is a bounded linear map with operator norm $\leq 1$.*

*(iv) The composition morphism*

$$C : L(V,W) \times L(U,V) \to L(U,W), \quad (T,S) \mapsto T \circ S$$

*is continuously differentiable with total derivative*

$$dC_{(T,S)}(\Delta T, \Delta S) = \Delta T \circ S + T \circ \Delta S.$$

*Proof.* **(i)** It is easy to see that a composition of linear maps is a linear map (and this is proved in linear algebra courses). Assume that $S$ and $T$ are bounded linear maps. For every $\vec{u} \in U$, we have $\|S(\vec{u})\|_V \leq \|S\|_{\text{op}} \|\vec{u}\|_U$. And for every $\vec{v} \in V$, also $\|T(\vec{v})\|_W \leq \|T\|_{\text{op}} \|\vec{v}\|_V$. Setting $\vec{v} = S(\vec{u})$ gives $\|(S \circ T)(\vec{u})\|_W \leq \|S\|_{\text{op}} \|T(\vec{u})\|_V \leq \|S\|_{\text{op}} \|T\|_{\text{op}} \|\vec{u}\|_U$. Therefore $T \circ S$ is a bounded linear map and $\|T \circ S\|_{\text{op}} \leq \|T\|_{\text{op}} \cdot \|S\|_{\text{op}}$.

**(ii)** It is straightforward to see that $C_{\bullet,S}$ is a linear map. This is essentially (one half of) distributivity of composition with addition and scalar multiplication. And by (i), $\|C_{\bullet,S}(T)\|_{\text{op}} = \|T \circ S\|_{\text{op}} \leq \|T\|_{\text{op}} \|S\|_{\text{op}}$. Therefore $C_{\bullet,S}$ is a bounded linear map with $\|C_{\bullet,S}\|_{\text{op}} \leq \|S\|_{\text{op}}$. A similar argument proves the analogous result for $C_{T,\bullet}$.

**(iii)** It is straightforward to see that $C_{\bullet,*}$ is a linear map. This is essentially the other half of distributivity. And by (ii), $\|C_{\bullet,S}\|_{\text{op}} \leq \|S\|_{\text{op}}$. Therefore $C_{\bullet,*}$ is a bounded linear map with $\|C_{\bullet,*}\|_{\text{op}} \leq 1$. A similar argument proves the analogous result for $C_{*,\bullet}$.

**(iv)** By Theorem 3.8.2 on p. 154, for every $(T,S) \in L(V,W) \times L(U,V)$, the derivative $dC_{(T,S)}$ exists and varies continuously in $(T,S)$ if and only if for every $(T,S)$ both partial derivatives $d(C_{\bullet,S})_T$ and $d(C_{T,\bullet})_S$ exist and vary continuously in $(T,S)$. By (ii), both $C_{\bullet,S}$ and $C_{T,\bullet}$ are bounded linear operators, hence differentiable with derivatives $d(C_{\bullet,S})_T = C_{\bullet,S}$ and $d(C_{T,\bullet})_S = C_{T,\bullet}$. By (iii), these both vary continuously in $(T,S)$. Hence $C$ is continuously differentiable, and

$$dC_{(T,S)}(\Delta T, \Delta S) = d(C_{\bullet,S})_T(\Delta T) + d(C_{T,\bullet})_S(\Delta S) = \Delta T \circ S + T \circ \Delta S.$$

□

**Corollary 0.2.** *Let $(R, \| \bullet \|_R)$ be a normed vector space, let $\tilde{R}$ be an open subset of $R$, let $S : \tilde{R} \to L(U,V)$ and $T : \tilde{R} \to L(V,W)$ be continuously differentiable functions. Then the function $T \circ S : \tilde{R} \to L(U,W)$ by $\vec{r} \mapsto T(\vec{r}) \circ S(\vec{r})$ is continuously differentiable and*

$$d(T \circ S)_{\vec{r}}(\Delta \vec{r}) = dT_{\vec{r}}(\Delta \vec{r}) \circ S(\vec{r}) + T(\vec{r}) \circ dS_{\vec{r}}(\Delta \vec{r}).$$

*In particular, if $R = \mathbb{R}$ with its absolute value norm, then we have*

$$\frac{d}{dt}(T(t) \circ S(t)) = \frac{dT}{dt}(t) \circ S(t) + T(t) \circ \frac{dS}{dt}(t).$$

*Proof.* This follows from Lemma 0.1 together with the Chain Rule, Theorem 3.6.2 on p. 143. $\square$

**Solution to (a)** Apply Corollary 0.2 with $R = \mathbb{R}$, $\tilde{R} = I$, $U = V = W$, with $T(t) = A(t, t_1)$ and with $S(t)$ the constant function $S(t) = A(t_1, t_0)$, which has zero derivative. This gives

$$\frac{d}{dt}(A(t, t_1) \circ A(t_1, t_0)) = \frac{d}{dt}A(t, t_1) \circ A(t_1, t_0).$$

Since $A(t, t_1)$ is a solution of the initial value problem,

$$\frac{d}{dt}A(t, t_1) = F_t \circ A(t, t_1).$$

Substituting this in gives

$$\frac{d}{dt}(A(t, t_1) \circ A(t_1, t_0)) = (F_t \circ A(t, t_1)) \circ A(t_1, t_0).$$

And by associativity of composition, this gives

$$\frac{d}{dt}(A(t, t_1) \circ A(t_1, t_0)) = F_t \circ (A(t, t_1) \circ A(t_1, t_0)),$$

i.e.,

$$\frac{d}{dt}\widehat{A}(t) = F_t \circ \widehat{A}(t).$$

Moreover we have

$$\widehat{A}(t_1) = A(t_1, t_1) \circ A(t_1, t_0).$$

Since $A(t, t_1)$ solves the initial value problem, $A(t_1, t_1)$ equals $\mathrm{Id}_W$. Therefore we have

$$\widehat{A}(t_1) = \mathrm{Id}_W \circ A(t_1, t_0) = A(t_1, t_0).$$

Therefore $\widehat{A}(t)$ solves the given initial value problem. On the other hand, $\tilde{A}(t) = A(t, t_0)$ clearly solves the initial value problem. Since the solution of the initial value problem is unique, it follows that $\widehat{A}(t) = \tilde{A}(t)$, i.e.,

$$A(t, t_1) \circ A(t_1, t_0) = A(t, t_0).$$

In particular, since $A(t_0, t_0) = \mathrm{Id}_W = A(t_1, t_1)$, we conclude that

$$A(t_0, t_1) \circ A(t_1, t_0) = \mathrm{Id}_W = A(t_1, t_0) \circ A(t_0, t_1).$$

Therefore $A(t_1, t_0)$ and $A(t_0, t_1)$ are inverse linear operators.

**Solution to (b)** First of all, for any continuously differentiable function $f : I \to L(W, W)$, by Corollary 0.2, $f(t) \circ U^{-1}$ is continuously differentiable, and

$$\frac{d}{dt}(f(t) \circ U^{-1}) = \frac{df}{dt}(t) \circ U^{-1}.$$

And for $f(t) = U \circ A(t,t_0)$, again by Corollary 0.2,

$$\frac{d}{dt}(U \circ A(t,t_0)) = U \circ \frac{d}{dt}A(t,t_0).$$

Putting this together gives,

$$\frac{d}{dt}(U \circ A(t,t_0) \circ U^{-1}) = U \circ \frac{dA(t,t_0)}{dt} \circ U^{-1}.$$

Since $A(t,t_0)$ solves the initial value problem,

$$\frac{dA(t,t_0)}{dt} = F_t \circ A(t,t_0),$$

we have that

$$\frac{d}{dt}B(t,t_0) = U \circ F_t \circ A(t,t_0) \circ U^{-1} =$$

$$U \circ F_t \circ U^{-1} \circ U \circ A(t,t_)) \circ U^{-1} = (U \circ F_t \circ U^{-1}) \circ B(t,t_0).$$

Finally, $B(t_0,t_0) = U \circ A(t_0,t_0) \circ U^{-1}$. Since $A(t_0,t_0)$ equals $\mathrm{Id}_W$, this gives

$$B(t_0,t_0) = U \circ \mathrm{Id}_W \circ U^{-1} = U \circ U^{-1} = \mathrm{Id}_W.$$

Therefore $B(t,t_0) = U \circ A(t,t_0) \circ U^{-1}$ solves the initial value problem.

**Solution to (c)** Denote by $T : I \to L(W,W)$ the function $T(t) = A(t,t_0)$. This is continuously differentiable with derivative

$$\frac{dT}{dt}(t) = F_t \circ T(t)$$

Denote by $S : I \to W = L(\mathbb{R},W)$ the function

$$S(t) = \int_{t_0}^{t} A(s,t_0)^{-1} \circ \vec{g}(s)ds.$$

By the Fundamental Theorem of Calculus, $S(t)$ is continuously differentiable with derivative equal to the continuous function

$$\frac{dS}{dt} = A(t,t_0)^{-1} \circ \vec{g}(t).$$

Therefore applying Corollary 0.2, $\vec{x} : I \to W = L(\mathbb{R},W)$, $\vec{x}(t) = T(t) \circ S(t)$, is continuously differentiable with derivative equal to

$$\frac{d\vec{x}}{dt}(t) = \frac{dT}{dt}(t) \circ S(t) + T(t) \circ \frac{dS}{dt}(t) =$$

$$F_t \circ T(t) \circ S(t) + A(t,t_0) \circ (A(t,t_0)^{-1} \circ \vec{g}(t)) = F_t \circ \vec{x}(t) + \vec{g}(t).$$

Therefore $\vec{x}(t)$ solves the ordinary differential equation. Moreover,

$$\vec{x}(t_0) = T(t_0) \circ S(t_0) = A(t_0, t_0) \circ 0_W = \mathrm{Id}_W(0_W) = 0_W.$$

Therefore $\vec{x}(t)$ solves the initial value problem.

**Solution to (2) (a) and (b)** Although the following is *not* called for in this problem, it did come up in lecture. For a real-valued power series $f(z) = \sum_{m=0}^{\infty} c_m z^m$ with radius of convergence $R$, not only is the power series $f_{L(W,W)}(T) := \sum_{m=0}^{\infty} c_m T^m$ convergent and continuous for $T \in B_R(0) \subset L(W,W)$, in fact it is continuously differentiable on $B_R(0)$. This requires more work than proving only that $f_{L(W,W)}$ is continuous. The proof uses analogues in $L(W,W)$ of some elementary algebraic facts about polynomials, e.g., the binomial theorem, and some elementary results about derivatives of polynomials, and the proof uses the Taylor approximation with remainder for real-valued power series. The basic idea is that for $T, \Delta T \in L(W,W)$ with $t := \|T\|_{\mathrm{op}}$, $\Delta t := \|\Delta T\|_{\mathrm{op}}$, one can bound the remainder term in the first-order Taylor approximation, $\|f_{L(W,W)}(T + \Delta T) - f_{L(W,W)}(T) - d(f_{L(W,W)})_T(\Delta T)\|_{\mathrm{op}}$, by the analogous "classical" term $|f(t + \Delta t) - f(t) - f'(t)\Delta t|$, which in turn is bounded by the Mean Value Theorem in terms of a uniform constant times $|\Delta t|^2$.

**Lemma 0.3.** *Let $(V, \| \bullet \|_V)$ be a Banach space. Let $R > 0$ be a real number. Let $(c_m)_{m=0,1,2,\dots}$ be a sequence of elements in $V$ such that the power series $\sum_{m=0}^{\infty} c_m t^m$ has radius of convergence $R$, i.e., for every $0 \le r < R$, the series $\sum_{m=0}^{\infty} \|c_m\|_V r^m$ converges. Then all of the following hold.*

*(i) For every real number $0 \le r < R$, the sequence $(f_n)_{n=0,1,2,\dots}$ of partial sums $f_n = \sum_{m=0}^{n} c_m t^m$ converges in $BC([-r,r], V)$. In particular, the pointwise limit $f_\infty$ is a continuous function $f_\infty : (-R, R) \to V$.*

*(ii) For every sequence $(a_m)_{m=0,1,2,\dots}$ of real numbers with $\limsup(\sqrt[m]{|a_m|}) \le 1$ and for every integer $k \ge 0$, also $(a_m c_{m+k})_{m=0,1,2,\dots}$ gives a power series of radius of convergence $R$.*

*(iii) In particular, the sequences $((m+1)c_{m+1})_{m=0,1,2,\dots}$ and $((m+2)(m+1)c_{m+2})_{m=0,1,2,\dots}$ give continuous functions $g_\infty(t) = \sum_{m=1}^{\infty} m c_m t^{m-1}$ and $h_\infty(t) = \sum_{m=2}^{\infty} m(m-1) c_m t^{m-2}$ defined on $(-R, R)$.*

*(iv) Denote by $\|h_\infty\|_V(t)$ the power series $\sum_{m=0}^{\infty} m(m-1)\|c_m\|_V t^{m-2}$ which is uniformly convergent, bounded and continuous on every $[-r, r]$ with $0 < r < R$ and which is continuous on $(-R, R)$. For every real number $0 < r < R$, for every $t \in (-r, r)$, and for every real $\Delta t$ with $|\Delta t| < r - |t|$, we have*

$$\|f_\infty(t + \Delta t) - f_\infty(t) - g_\infty(t) \cdot \Delta t\|_V \le \frac{1}{2}\|h_\infty\|_V(r) \cdot |\Delta t|^2.$$

*In particular, $f_\infty(t)$ is differentiable at $t$ with derivative equal to $g_\infty(t)$.*

*Proof.* **(i)** Because $V$ is a Banach space, also $BC([-r,r], V)$ is a Banach space with the uniform norm. Thus every absolutely convergent series in $BC([-r,r], V)$ is convergent. And the hypotheses exactly insure that the series $(f_n)_{n=0,1,2,\dots}$ is absolutely convergent,

$$\sum_{m=0}^{n} \|c_m t^m\|_{\mathrm{un}} \leq \sum_{m=0}^{n} \|c_m\|_V r^m.$$

Since $(-R,R)$ is the union of $[-r,r]$ over all $r$ with $0<r<R$, it follows that $f_\infty$ is defined and continuous on all of $(-R,R)$.

**(ii) and (iii)** This follows by the Root Test from single variable calculus.

**(iv)** Without loss of generality, assume that $\Delta t \geq 0$. As we have seen before, the function

$$I_{t,t+\Delta t} : BC([-r,r],V) \to V, \quad f(t) \mapsto \int_t^{t+\Delta t} f(s)ds$$

is a bounded linear functional with operator norm $\leq |\Delta t|$. In particular, since the partial sums $g_n = \sum_{m=1}^{n} m c_m t^{m-1}$ converge uniformly to $g_\infty$, the integrals converge as well. By construction,

$$\int_t^{t+\Delta t} g_n(s)ds = f_n(t+\Delta t) - f_n(t).$$

Similarly we have

$$\int_t^{t+\Delta t} g_n(s) - g_n(t)ds = f_n(t+\Delta t) - f_n(t) - g_n(t)\Delta t.$$

Since $(f_n)_{n=0,1,2,\dots}$ converges uniformly to $f_\infty$ and since $(g_n)_{n=0,1,2,\dots}$ converges uniformly to $g_\infty$, it suffices to prove that for every $n=0,1,2,\dots,$ we have

$$\| \int_t^{t+\Delta t} g_n(s) - g_n(t)ds\|_V \leq \frac{1}{2}\|h_\infty\|_W(r)|\Delta t\|^2.$$

Of course this integral is the same as

$$\sum_{m=1}^{n} m c_m \int_t^{t+\Delta t} s^{m-1} - t^{m-1}ds.$$

By the Mean Value Theorem (or the Binomial Theorem, etc.), $|s^{m-1}-t^{m-1}| \leq |s-t|(m-1)r^{m-2}$. Integrating gives,

$$\left\| \int_t^{t+\Delta t} g_n(s) - g_n(t)ds \right\|_V \leq \frac{1}{2}\sum_{m=2}^{n} m(m-1)\|c_m\|_W r^{m-2} \cdot |\Delta t|^2 \leq \frac{1}{2}\|h_\infty\|_W(r) \cdot |\Delta t|^2.$$

$\square$

For every integer $n \geq 2$, define a function

$$C_n : L(W,W)^n \to L(W,W)$$

recursively in $n$ by $C_2(T_1, T_2) = C(T_1, T_2)$, where $C$ is the function from Lemma 0.1, and by

$$C_{n+1}(T_1, \dots, T_n, T_{n+1}) = C(C_n(T_1, \dots, T_n), T_{n+1}).$$

Because composition is associative, the function $C_n$ can be unambiguously expressed by

$$C_n(T_1, \dots, T_n) = T_1 \circ \dots \circ T_n.$$

Define the *diagonal map* by,

$$\delta_n : L(W,W) \to L(W,W)^n, \quad T \mapsto (T, T, \dots, T).$$

And define the *power map* by $P_n = C_n \circ \delta_n$, i.e.,

$$P_n : L(W,W) \to L(W,W)^n, \quad T \mapsto T \circ T \circ \dots \circ T \text{ ($n$ times )}.$$

Let $n \geq 1$ be an integer. Denote by $J_n$ the set of all functions $j : \{1, \dots, n\} \to \{1, 2\}$. In particular denote by $c_{n,1}$, resp. $c_{n,2}$, the constant function with value 1, resp. with value 2. For each $j \in J_n$, define a function

$$\widehat{j} : L(W,W) \times L(W,W) \to L(W,W)^n, \quad (T_1, T_2) \mapsto (T_{j(1)}, T_{j(2)}, \dots, T_{j(n)}),$$

i.e., for every projection $\pi_k : L(W,W)^n \to L(W,W)$ we have $\pi_k \circ \widehat{j} = \pi_{j(k)}$. In particular, for $i = 1, 2$ we have

$$\widehat{c}_{n,i}(T_1, T_2) = (T_i, T_i, \dots, T_i) = P_n \circ \pi_i(T_1, T_1).$$

For every $j \in J_n$, denote by $\tilde{j}$ the composition $C_n \circ \widehat{j}$,

$$\tilde{j} : L(W,W) \times L(W,W) \to L(W,W), \quad (T_1, T_2) \mapsto T_{j(1)} \circ \dots \circ T_{j(k)} \circ \dots \circ T_{j(n)}.$$

For every integer $k = 0, \dots, n$, denote by $J_{n,k}$ the set of functions $j$ in $J_n$ with $j^{-1}(\{2\})$ of cardinality $k$, e.g., $J_{n,0} = \{c_{n,1}\}$ and $J_{n,n} = \{c_{n,2}\}$. Define a function $DP_n$ by

$$DP_n : L(W,W) \times L(W,W) \to L(W,W), \quad (DP_n)_{T_1}(T_2) = \sum_{j \in J_{n,1}} \tilde{j}(T_1, T_2) =$$

$$\sum_{k=1}^{n} C_n(T_1, \dots, T_1, T_2, T_1, \dots, T_1) \text{ ($k^{\text{th}}$ variable )}.$$

**Lemma 0.4.** *Notations are as above.*

*(i) For every $n \geq 2$, the function $C_n(T_1, \dots, T_n)$ is continuously differentiable with*

$$d(C_n)_{(T_1,\dots,T_n)}(\Delta T_1, \dots, \Delta T_n) = \sum_{k=1}^{n} C_n(T_1, \dots, T_{k-1}, \Delta T_k, T_{k+1}, \dots, T_n).$$

*The function $\delta_n(T)$ is a bounded linear map with operator norm $\leq n$. And $P_n(T)$ is continuously differentiable with $d(P_n)_T(\Delta T) = (DP_n)_T(\Delta T)$.*

*(ii) For every $n \geq 1$ and for every $j \in J_n$, the function $\widehat{j}(T_1, T_2)$ is a bounded linear map with operator norm $\leq n$. In particular, it is continuously differentiable.*

*(iii) For every $n \geq 1$ and for every $j \in J_n$, the function $\tilde{j}(T_1, T_2)$ is continuously differentiable with*

$$d\tilde{j}_{(T_1,T_2)}(\Delta T_1, \Delta T_2) = \sum_{k=1}^{n} C_n(T_{j(1)}, \dots, T_{j(k-1)}, \Delta T_{j(k)}, T_{j(k+1)}, \dots, T_{j(n)}).$$

*(iv) Each set $J_{n,k}$ is finite of size $\binom{n}{k}$. The set $J_n$ is finite of size $2^n$.*

*(v) For every $j \in J_{n,k}$ we have*

$$\|\tilde{j}(T_1, T_2)\|_{op} \leq \|T_1\|_{op}^{n-k} \|T_2\|_{op}^{k}.$$

*In particular, we have*

$$\|d(P_n)_T\|_{op} \leq n\|T\|_{op}^{n-1}.$$

*(vi) For every integer $n \geq 1$ and for every $(T, \Delta T) \in L(W, W) \times L(W, W)$, we have*

$$P_n(T + \Delta T) = (T + \Delta T)^n = \sum_{j \in J_n} \tilde{j}(T, \Delta T) = \sum_{k=0}^{n} \sum_{j \in J_{n,k}} \tilde{j}(T, \Delta T) =$$

$$P_n(T) + (dP_n)_T(\Delta T) + \sum_{k=2}^{n} \sum_{j \in J_{n,k}} \tilde{j}(T, \Delta T).$$

*(vii) For every $(T, \Delta T) \in L(W, W) \times L(W, W)$, we have*

$$\|P_n(T + \Delta T) - P_n(T) - (dP_n)_T(\Delta T)\|_{op} \leq \sum_{k=2}^{n} \binom{n}{k} \|T\|_{op}^{n-k} \|\Delta T\|_{op}^{k} =$$

$$(\|T\|_{op} + \|\Delta T\|_{op})^n - \|T\|_{op}^n - n\|T\|_{op}^{n-1}\|\Delta T\|_{op}.$$

*Proof.* **(i)** This is proved by induction on $n$. The base case $n = 2$ is proved in Lemma 0.1. Thus let $n \geq 2$ be an integer, assume the result is true for $C_2, \dots, C_n$, and consider the result for $C_{n+1}$. Since $C_{n+1}(T_1, \dots, T_n, T_{n+1})$ is defined to be $C_2(C_n(T_1, \dots, T_n), T_{n+1})$, and since $C_2$ and $C_n$ are continuously differentiable, by the Chain Rule also $C_{n+1}$ is continuously differentiable and

$$(dC_{n+1})_{T_1,\dots,T_n,T_{n+1}}(\Delta T_1, \dots, \Delta T_n, \Delta T_{n+1}) =$$

$$(d(C_2)_{\bullet,T_n})_{C_n(T_1,\dots,T_n)} \circ (dC_n)_{(T_1,\dots,T_n)}(\Delta T_1, \dots, \Delta T_n) + (d(C_2)_{C_n(T_1,\dots,T_n),\bullet})_{T_{n+1}}(\Delta T_{n+1}) =$$

$$\sum_{k=1}^{n} C_n(T_1, \dots, T_{k-1}, \Delta T_k, T_{k+1}, \dots, T_n) \circ T_{n+1} + C_n(T_1, \dots, T_n) \circ \Delta T_{n+1} =$$

$$\sum_{k=1}^{n+1} C_{n+1}(T_1, \dots, T_{k-1}, \Delta T_k, T_{k+1}, \dots, T_n, T_{n+1}).$$

This proves (i) by induction on $n$. It is obvious that $\delta_n$ is a linear map. By the definition of the product norm on $L(W, W)^n$, it is clear that the operator norm is $\leq n$. In particular, $\delta_n$ is continuously differentiable with derivative equal to $\delta_n$. By the Chain Rule, $P_n = C_n \circ \delta_n$ is also continuously differentiable with the specified derivative.

**(ii)** It is obvious that this is a linear map. And by the triangle inequality

$$\|\widehat{j}(T_1, T_2)\|_{L(W,W)^n} = \sum_{k=1}^{n} \|T_{j(k)}\|_{L(W,W)} = |j^{-1}(\{1\})| \|T_1\|_{L(W,W)} + |j^{-1}(\{2\})| \|T_2\|_{L(W,W)}$$

$$\leq n(\|T_1\|_{L(W,W)} + \|T_2\|_{L(W,W)}) = n\|(T_1, T_2)\|_{L(W,W)\times L(W,W)}.$$

So $\widehat{j}$ is a bounded linear map with $\|\widehat{j}\|_{\text{op}} \leq n$.

**(iii)** By (i), $C_n$ is continuously differentiable. And by (ii), $\widehat{j}$ is continuously differentiable. By the Chain Rule, the composition $\tilde{j} = C_n \circ \widehat{j}$ is continuously differentiable with derivative

$$d\tilde{j}_{(T_1,T_2)}(\Delta T_1, \Delta T_2) = (dC_n)_{\widehat{j}(T_1,T_2)} \circ \widehat{j}(\Delta T_1, \Delta T_2) =$$

$$\sum_{k=1}^{n} C_n(T_{j(1)}, \dots, T_{j(k-1)}, \pi_k \circ \widehat{j}(\Delta T_1, \Delta T_2), T_{j(k+1)}, \dots, T_{j(n)}) =$$

$$\sum_{k=1}^{n} C_n(T_{j(1)}, \dots, T_{j(k-1)}, \Delta T_{j(k)}, T_{j(k+1)}, \dots, T_{j(n)}).$$

**(iv)** This is elementary combinatorics (and is proved in MAT 200, for instance).

**(v)** By (i) of Lemma 0.1 and induction,

$$\|\tilde{j}(T_1, T_2)\|_{\text{op}} = \|T_{j(1)} \circ \dots \circ T_{j(n)}\|_{\text{op}} \leq \prod_{1 \leq k \leq n} \|T_{j(k)}\|_{\text{op}} =$$

$$\left(\prod_{1\leq k\leq n,\ j(k)=1} \|T_1\|_{\mathrm{op}}\right)\cdot\left(\prod_{1\leq k\leq n,\ j(k)=2} \|T_2\|_{\mathrm{op}}\right) = \|T_1\|_{\mathrm{op}}^{n-k}\cdot\|T_2\|_{\mathrm{op}}^{k}.$$

Combined with the triangle inequality and (iv), this gives

$$\|d(P_n)_T(\Delta T)\|_{\mathrm{op}} \leq |J_{n,1}|\|T\|_{\mathrm{op}}^{n-1}\|\Delta T\|_{\mathrm{op}} = n\|T\|_{\mathrm{op}}\|^{n-1}\Delta T\|_{\mathrm{op}}.$$

Therefore we have

$$\|d(P_n)_T\|_{\mathrm{op}} \leq n\|T\|_{\mathrm{op}}^{n-1}.$$

**(vi)** Just as with the usual Binomial Theorem, this is proved by induction on $n$. For $n=1$ it is elementary; $J_1 = \{c_{1,1}, c_{1,2}\}$, $\tilde{c}_{1,1}(T,\Delta T) = T$ and $\tilde{c}_{1,2}(T,\Delta T) = \Delta T$. By way of induction, assume the result is proved for the integers $1,\dots,n$, and consider the result for $n+1$. For every $j\in J_n$ and for $i=1,2$, define $j_{(i)}\in J_{n+1}$ by

$$j_{(i)}(k) = \begin{cases} j(k), & k\in\{1,\dots,n\},\\ i, & k=n+1.\end{cases}$$

Then $J_{n+1}$ equals $\{j_{(1)}|j\in J_n\}\sqcup\{j_{(2)}|j\in J_n\}$. By construction of $P_{n+1}$ and by bilinearity of $C$ we have

$$P_{n+1}(T+\Delta T) = C(P_n(T+\Delta T),(T+\Delta T)) = C(P_n(T+\Delta T),T) + C(P_n(T+\Delta T),\Delta T).$$

By the induction hypothesis and by bilinearity of $C$ this is

$$C(\sum_{j\in J_n}\tilde{j}(T,\Delta T),T) + C(\sum_{j\in J_n}\tilde{j}(T,\Delta T),\Delta T) = \sum_{j\in J_n}\tilde{j}(T,\Delta T)\circ T + \sum_{j\in J_n}\tilde{j}(T,\Delta T)\circ\Delta T =$$

$$\sum_{j\in J_n}\tilde{j}_{(1)}(T,\Delta T) + \sum_{j\in J_n}\tilde{f}_{(2)}(T,\Delta T) = \sum_{j'\in J_{n+1}}\tilde{j}'(T,\Delta T).$$

Thus the result holds for $n+1$. So the result is proved by induction on $n$.

**(vii)** By (vi) and (i) we have

$$P_n(T+\Delta T) - P_n(T) - (dP_n)_T(\Delta T) = \sum_{k=2}^{n}\sum_{j\in J_{n,k}}\tilde{f}(T,\Delta T).$$

By the triangle inequality, this gives,

$$\|P_n(T+\Delta T) - P_n(T) - (dP_n)_T(\Delta T)\|_{\mathrm{op}} \leq \sum_{k=2}^{n}\sum_{j\in J_{n,k}}\|\tilde{f}(T,\Delta T)\|_{\mathrm{op}}.$$

And by (v), for $j\in J_{n,k}$ we have $\|\tilde{j}(T,\Delta T)\|_{\mathrm{op}} \leq \|T_1\|_{\mathrm{op}}^{n-k}\cdot\|T_2\|_{\mathrm{op}}^{k}$. Thus the inequality above gives,

$$\|P_n(T+\Delta T) - P_n(T) - (dP_n)_T(\Delta T)\|_{\mathrm{op}} \leq \sum_{k=2}^{n}|J_{n,k}|\cdot\|T_1\|_{\mathrm{op}}^{n-k}\cdot\|T_2\|_{\mathrm{op}}^{k}.$$

Combined with (iv), this gives,

$$\|P_n(T+\Delta T) - P_n(T) - (dP_n)_T(\Delta T)\|_{\text{op}} \leq \sum_{k=2}^{n} \binom{n}{k} \|T_1\|_{\text{op}}^{n-k} \cdot \|T_2\|_{\text{op}}^{k}.$$

□

Let $r > 0$ be a real number. Let $B_{\leq r}(0)$ denote the closed unit ball in $L(W,W)$. And for every integer $n \geq 0$, denote by $P_{n,r} : B_{\leq r}(0) \to L(W,W)$ the restriction of $P_n$ to $B_{\leq r}(0)$. This is continuous by Lemma 0.4(i), and it is bounded with $\|P_{n,r}\|_{\text{un}} \leq r^n$ by (v). For every integer $n \geq 1$, denote by $dP_n : L(W,W) \to L(L(W,W), L(W,W))$ the function $T \mapsto (dP_n)_T$, and denote by $dP_{n,r} : B_{\leq r}(0) \to L(L(W,W), L(W,W))$ the restriction to $B_{\leq r}(0)$. By Lemma 0.4(i), $dP_n$ is a continuous function. And by Lemma 0.4(v), $dP_{n,r}$ is bounded with $\|dP_{n,r}\|_{\text{un}} \leq nr^{n-1}$.

**Proposition 0.5.** *Let $R > 0$ be a real number. Let $(c_m)_{m=0,1,2,\dots}$ be a sequence of real numbers such that the power series $f_\infty(t) = \sum_{m=0}^\infty c_m t^m$ has radius of convergence $R$. Denote by $|f_\infty|(t)$ the associated power series $\sum_{m=0}^\infty |c_m| t^m$, which also has radius of convergence $R$. And denote by $g_\infty(t)$, $h_\infty(t)$ and $|h_\infty|(t)$ the associated power series as defined in Lemma 0.3. All of the following hold.*

*(i) For every real number $0 < r < R$, the sequence $(c_m P_{m,r})_{m=0,1,2,\dots}$ gives an absolutely convergent series $\sum_{m=0}^\infty c_m P_{m,r}$ in $BC(B_{\leq r}(0), L(W,W))$. Moreover, $\sum_{m=0}^\infty \|c_m P_{m,r}\|_{un} \leq \sum_{m=0}^\infty |c_m| r^m = |f_\infty|(r)$.*

*(ii) The absolutely convergent series $\sum_{m=0}^\infty c_m P_{m,r}$ is convergent; denote the limit by $f_{\infty,\leq r} : B_{\leq r}(0) \to L(W,W)$ and denote by $f_{\infty,R} : B_R(0) \to L(W,W)$ the unique function whose restriction to every $B_{\leq r}(0)$ equals $f_{\infty,\leq r}$. The function $f_{\infty,R}$ is continuous.*

*(iii) For every real number $0 < r < R$, the sequence $(c_m dP_{m,R})_{m=0,1,2,\dots}$ gives an absolutely convergent series $\sum_{m=0}^\infty c_m dP_{m,r}$ in $BC(B_{\leq r}(0), L(L(W,W), L(W,W)))$ Moreover, $\sum_{m=0}^\infty \|c_m dP_{m,r}\|_{un} \leq \sum_{m=0}^\infty m|c_m| r^{m-1} = |g_\infty|(r)$.*

*(iv) The absolutely convergent series $\sum_{m=0}^\infty c_m dP_{m,r}$ is convergent; denote the limit by*

$$Df_{\infty,\leq r} : B_{\leq r}(0) \to L(L(W,W), L(W,W)), \quad T \mapsto (Df_{\infty,\leq r})_T,$$

*and denote by $Df_{\infty,R} : B_R(0) \to L(L(W,W), L(W,W))$ the unique function whose restriction to every $B_{\leq r}(0)$ equals $Df_{\infty,\leq r}$. The function $Df_{\infty,R}$ is continuous.*

*(v) For every real number $0 < r < R$, for every $T \in B_r(0) \subset L(W,W)$, and for every $\Delta T \in L(W,W)$ with $\|\Delta T\|_{op} < r - \|T\|_{op}$, we have*

$$\|f_{\infty,r}(T+\Delta T) - f_{\infty,r}(T) - (Df_{\infty,R})_T(\Delta T)\|_{un} \leq \frac{1}{2}|h_\infty|(r)\|\Delta T\|_{op}^2.$$

*In particular, $f_{\infty,R}$ is differentiable at $T$ with derivative equal to $(Df_{\infty,R})_T$. As this is continuous, $f_{\infty,R}$ is continuously differentiable on $B_R(0)$.*

*Proof.* **(i)** As discussed above, $\|P_{m,r}\|_{\text{un}} \leq r^m$. Thus we have for every integer $n \geq 0$,

$$\sum_{m=0}^{n} \|c_m P_{m,r}\|_{\text{un}} \leq \sum_{m=0}^{n} |c_m| r^m \leq |f_\infty|(r).$$

So the series is absolutely convergent.

**(ii)** Since $L(W,W)$ is a Banach space, $BC(B_{\leq r}(0), L(W,W))$ is also a Banach space with respect to the uniform norm. Every series in a Banach space which is absolutely convergent is convergent. Finally, for every $T \in B_R(0)$, there exists a real number $r$ with $\|T\|_{\text{op}} < r < R$ so that $T \in B_r(0) \subset B_{\leq r}(0)$. The restriction of $f_{\infty,R}$ to the open set $B_r(0)$ equals the restriction of $f_{\infty,\leq r}$ by definition. And by construction, $f_{\infty,\leq r}$ is in $BC(B_{\leq r}(0), L(W,W))$, i.e., it is bounded and continuous. Thus the restriction of $f_{\infty,R}$ to $B_r(0)$ is continuous (and bounded), so $f_{\infty,R}$ is continuous at $T$. Since this holds for every $T \in B_R(0)$, $f_{\infty,R}$ is continuous.

**(iii)** This is similar to (i) using the estimate

$$|DP_{m,r}| \leq m r^{m-1}$$

from above.

**(iv)** This is similar to (ii).

**(v)** Since $\|\bullet\|_{\text{un}}$ is continuous (with respect to the uniform norm), and since $(f_{n,r})_{n=0,1,\dots}$ converges uniformly to $f_{\infty,r}$ and $(Df_{n,r})_{n=1,2,\dots}$ converges uniformly to $Df_{\infty,r}$, it suffices to prove for every integer $n \geq 0$ that

$$\|f_{n,r}(T+\Delta T) - f_{n,r}(T) - (Df_{n,r})_T(\Delta T)\|_{\text{op}} \leq \frac{1}{2}|h_\infty|(r)\|\Delta T\|_{\text{op}}^2.$$

And we have

$$f_{n,r}(T+\Delta T) - f_{n,r}(T) - (Df_{n,r})_T(\Delta T) = \sum_{m=0}^{n} c_m \left[P_{m,r}(T+\Delta T) - P_{m,r}(T) - (dP_{m,r})_T(\Delta T)\right].$$

Thus we have

$$\|f_{n,r}(T+\Delta T) - f_{n,r}(T) - (Df_{n,r})_T(\Delta T)\|_{\text{op}} \leq \sum_{m=0}^{n} |c_m| \|P_{m,r}(T+\Delta T) - P_{m,r}(T) - (dP_{m,r})_T(\Delta T)\|_{\text{op}}.$$

By Lemma 0.4(vii), we have

$$\|P_{m,r}(T+\Delta T) - P_{m,r}(T) - (dP_{m,r})_T(\Delta T)\|_{\text{op}} \leq (t+\Delta t)^m - t^m - m t^{m-1}\Delta t,$$

where $t := \|T\|_{\text{op}}$ and $\Delta t : -\|\Delta T\|_{\text{op}}$. Substituting this in gives

$$\|f_{n,r}(T+\Delta T) - f_{n,r}(T) - (Df_{n,r})_T(\Delta T)\|_{\text{op}} \leq \sum_{m=0}^{n} |c_m|((t+\Delta t)^m - t^m - m t^{m-1}\Delta t) =$$

$$|f_n|(t+\Delta t) - |f_n|(t) - |g_n|(t)\Delta t.$$

And by Lemma 0.3(iv), we have

$$||f_n|(t+\Delta t) - |f_n|(t) - |g_n|(t)\Delta t| \leq \frac{1}{2}|h_\infty|(r)|\Delta t|^2,$$

for every integer $n \geq 0$. Taking the limit as $n \to \infty$, this gives the bound

$$\|f_{\infty,r}(T+\Delta T) - f_{\infty,r}(T) - (Df_{\infty,R})_T(\Delta T)\|_{\text{un}} \leq \frac{1}{2}|h_\infty|(r)\|\Delta T\|_{\text{op}}^2.$$

$\square$

Of particular importance is the case when $T$ and $\Delta T$ commute. In this case $(dP_m)_T(\Delta T)$ equals $mT^{m-1} \circ \Delta T$. So in this special case, $(df_{\infty,R})_T(\Delta T)$ equals $g_{\infty,R}(T) \circ \Delta T$. In particular, when $f_\infty(t)$ is the power series about 0 giving $e^t$, then for $T$ and $\Delta T$ commuting, this gives $d\exp_T(\Delta T) = \exp(T) \circ \Delta T = \Delta T \circ \exp(T)$. This was the crucial step in proving that

$$A(t,t_0) := \exp\left(\int_{t_0}^t F_s ds\right)$$

solves the initial value problem,

$$\frac{d}{dt}A(t) = F_t \circ A(t), \quad A(t_0) = \text{Id}_W$$

in the special case that $F_t$ and $F_t$ commute for every $s, t \in I$.

**Solution to (c)** By induction on $n \geq 2$, it is straightforward to compute that $C_n(UT_1U^{-1}, \dots, UT_nU^{-1}) = UC_n(T_1, \dots, T_n)U^{-1}$: the base case $n = 2$ is

$$C(UT_1U^{-1}, UT_2U^{-1}) = (UT_1U^{-1})(UT_2U^{-1}) = UT_1(U^{-1}U)T_2U^{-1} = U(T_1T_2)U^{-1} = UC(T_1,T_2)U^{-1}.$$

And the induction step is

$$C_{n+1}(UT_1U^{-1}, \dots, UT_nU^{-1}, UT_{n+1}U^{-1}) = C(C_n(UT_1U^{-1}, \dots, UT_nU^{-1}), UT_{n+1}U^{-1}),$$

which by the induction hypothesis equals

$$C(UC_n(T_1, \dots, T_n)U^{-1}, UT_{n+1}U^{-1})$$

which applying the base case once more gives

$$UC(C_n(T_1, \dots, T_n), T_{n+1})U^{-1} = UC_{n+1}(T_1, \dots, T_n, T_{n+1})U^{-1}.$$

Thus also $P_n(UTU^{-1}) = UP_n(T)U^{-1}$. So for each of the partial sum polynomials $f_n(z)$, $f_n(UTU^{-1})$ is a finite linear combination of expressions $P_m(UTU^{-1})$, which by the last sentence equals $UP_m(T)U^{-1}$, and hence $f_n(UTU^{-1}) = Uf_n(T)U^{-1}$. Taking the limit as $n \to \infty$ gives $f_\infty(UTU^{-1}) = Uf_\infty(T)U^{-1}$.

**Solution to** (4)

**Solution to** (a) Define $x_0(t) = x(t)$ and $x_1(t) = x'(t)$. Then $x''(t) = x_1'(t)$. Thus the second order differential equation

$$\frac{d^2x}{dt^2}(t) = -9x(t) + 6\frac{dx}{dt}(t)$$

becomes a first order differential equation,

$$\frac{dx_1}{dt}(t) = -9x_0(t) + 6x_1(t).$$

Altogether, the differential system reads,

$$\left\{ \begin{array}{rcrcr} \frac{dx_0}{dt}(t) & = & 0x_0(t) & + & 1x_1(t) \\ \frac{dx_1}{dt}(t) & = & -9x_0(t) & + & 6x_1(t) \end{array} \right.$$

In other words, this is the vector-valued IVP,

$$\frac{d\vec{x}}{dt}(t) = A\vec{x}(t), \quad \vec{x}(t_0) = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix},$$

where $A$ and $\vec{x}(t)$ are

$$A = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix}, \quad \vec{x}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \end{bmatrix}.$$

**Solution to** (b) The first step in finding a Jordan normal form is to compute the characteristic polynomial $c_A(x) := \mathrm{Det}(xI_{2\times 2} - A)$,

$$c_A(x) = \mathrm{Det}\begin{bmatrix} x & -1 \\ 9 & x-6 \end{bmatrix} = x(x-6) - (-1)(9) = x^2 - 6x + 9.$$

Notice that for every $n \times n$ matrix, $c_A(x) := \mathrm{Det}(xI_{n\times n} - A)$ has the form

$$c_A(x) = x^n - \mathrm{Tr}(A)x^{n-2} + \cdots + (-1)^n\mathrm{Det}(A)x^0.$$

Since clearly $\mathrm{Tr}(A) = 6$ and $\mathrm{Det}(A) = 0(6) - (1)(-9) = 9$, this also gives $c_A(x) = x^2 - 6x + 9$.

The next step in computing the Jordan normal form is to compute the factorization of $c_A(x)$. By the quadratic formula, the only root is 3, so $c_A(x)$ equals $(x-3)^2$. So there is only one eigenvalue, $\lambda = 3$. So the semisimple part of the Jordan canonical form must be

$$\tilde{S} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = 3\mathrm{Id}_{2\times 2}.$$

Notice that for every invertible matrix $U$, $U(3\mathrm{Id}_{2\times 2})U^{-1}$ equals $3(UU^{-1}) = 3\mathrm{Id}_{2\times 2}$. So the semisimple part of $A$, $S = U\tilde{S}U^{-1}$, is still $3\mathrm{Id}_{2\times 2}$. Thus $A$ is diagonalizable if and only if $A$ equals

$S = 3\mathrm{Id}_{2\times 2}$, which it clearly does not. Thus $A$ is not diagonalizable, and the nilpotent part of $A$ is $N = A - S = A - 3\mathrm{Id}_{2\times 2}$,

$$N = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ -9 & 3 \end{bmatrix}.$$

As a double-check, we have

$$N^2 = N \cdot N = \begin{bmatrix} -3 & 1 \\ -9 & 3 \end{bmatrix} \cdot \begin{bmatrix} -3 & 1 \\ -9 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0_{2\times 2}.$$

By the Gauss-Jordan row reduction algorithm, or any other means, the kernel of $N$ and $N^2$ are

$$E_N = \mathrm{Ker}(N) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right), \quad E_{N^2} = \mathrm{Ker}(N^2) = \mathbb{R}^2.$$

Thus the algorithm for finding $U$ and $N$ is to first find a *primitive* subspace, i.e., a 1-dimensional subspace $G_2 \subset E_{N^2}$ such that $E_{N^2} = G_2 + E_{N^1}$. The span of any vector not in $E_{N^1}$ will do, say $G_2 = \mathbf{span}(\mathbf{e}_2)$. This subspace has as basis just the vector $\vec{v}_2 := \mathbf{e}_2$. Then for a second basis vector, we choose

$$\vec{v}_1 = N\vec{v}_2 = \begin{bmatrix} -3 & 1 \\ -9 & 3 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

So the change-of-basis matrix is

$$U = [\vec{v}_1|\vec{v}_2] = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}.$$

And $\tilde{N}$ is determined by

$$AU = A[\vec{v}_1|\vec{v}_2] = [A\vec{v}_1|A\vec{v}_2] = [\vec{0}|\vec{v}_1] = [0\vec{v}_1+0\vec{v}_2|1\vec{v}_1+0\vec{v}_2] = [\vec{v}_1|\vec{v}_2]\cdot\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = U\tilde{N}, \ \ \tilde{N} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

**Solution to (c)** Since $\tilde{S}$ equals $3\mathrm{Id}_{2\times 2}$, we have

$$\tilde{S}(t-t_0) = 3(t-t_0)\mathrm{Id}_{2\times 2} = \begin{bmatrix} 3(t-t_0) & 0 \\ 0 & 3(t-t_0) \end{bmatrix}.$$

Since this is diagonal,

$$\exp(\tilde{S}(t-t_0)) = e^{3(t-t_0)}\mathrm{Id}_{2\times 2} = \begin{bmatrix} e^{3(t-t_0)} & 0 \\ 0 & e^{3(t-t_0)} \end{bmatrix}.$$

Similarly, we have

$$\tilde{N}(t-t_0) = \begin{bmatrix} 0 & t-t_0 \\ 0 & 0 \end{bmatrix}.$$

Since $(\tilde{N}(t-t_0))^2$ equals $0_{2\times 2}$, we have

$$\exp(\tilde{N}(t-t_0)) = \mathrm{Id}_{2\times 2} + \tilde{N}(t-t_0) + \frac{1}{2!}(\tilde{N}(t-t_0))^2 + \cdots = \mathrm{Id}_{2\times 2} + \tilde{N}(t-t_0) =$$

$$\left[\begin{array}{cc} 1 & t-t_0 \\ 0 & 1 \end{array}\right].$$

Thus we have

$$\exp(\tilde{A}(t-t_0)) = \exp(\tilde{S}(t-t_0))\exp(\tilde{N}(t-t_0)) = \left[\begin{array}{cc} e^{3(t-t_0)} & 0 \\ 0 & e^{3(t-t_0)} \end{array}\right] \cdot \left[\begin{array}{cc} 1 & t-t_0 \\ 0 & 1 \end{array}\right] =$$

$$\left[\begin{array}{cc} e^{3(t-t_0)} & (t-t_0)e^{3(t-t_0)} \\ 0 & e^{3(t-t_0)} \end{array}\right] = e^{3(t-t_0)}\exp(N(t-t_0)).$$

By **Problem 2(c)**, we have

$$\exp(A(t-t_0))U = U\exp(\tilde{A}(t-t_0)) = U \cdot (e^{3(t-t_0)}\exp(\tilde{N}(t-t_0))) = e^{3(t-t_0)}U \cdot \exp(\tilde{N}(t-t_0)) =$$

$$e^{3(t-t_0)}\left[\begin{array}{cc} 1 & 0 \\ 3 & 1 \end{array}\right]\left[\begin{array}{cc} 1 & (t-t_0) \\ 0 & 1 \end{array}\right] = e^{3(t-t_0)}\left[\begin{array}{cc} 1 & (t-t_0) \\ 3 & 3(t-t_0)+1 \end{array}\right].$$

This gives,

$$\exp(A(t-t_0)) = (U\exp(\tilde{A}(t-t_0)))U^{-1} = e^{3(t-t_0)}\left[\begin{array}{cc} 1 & (t-t_0) \\ 3 & 3(t-t_0)+1 \end{array}\right]\left[\begin{array}{cc} 1 & 0 \\ -3 & 1 \end{array}\right] =$$

$$e^{3(t-t_0)}\left[\begin{array}{ll} -3(t-t_0)+1 & 1(t-t_0) \\ -9(t-t_0) & 3(t-t_0)+1 \end{array}\right].$$

So the Green's function is

$$A(t,t_0) = A(t-t_0) = e^{3(t-t_0)}\left[\begin{array}{ll} -3(t-t_0)+1 & 1(t-t_0) \\ -9(t-t_0) & 3(t-t_0)+1 \end{array}\right].$$

**Solution to (d)** The general solution of the initial value problem

$$\frac{d\vec{x}}{dt}(t) = A\vec{x}(t), \quad \vec{x}(t_0) = \vec{b}$$

is given by

$$\vec{x}(t) = \exp(A(t-t_0))\vec{b}.$$

If we write $\vec{b} = U\vec{c}$, this becomes,

$$\vec{x}(t) = \exp(A(t-t_0))U\vec{c} = U\exp(\tilde{A}(t-t_0))\vec{c}.$$

In our case, we have

$$\exp(\tilde{A}(t-t_0))\vec{c} = e^{3(t-t_0)}\exp(\tilde{N}(t-t_0))\vec{c} = e^{3(t-t_0)}\left[\begin{array}{cc} 1 & t-t_0 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} c_0 \\ c_1 \end{array}\right] = e^{3(t-t_0)}\left[\begin{array}{c} c_0 + c_1(t-t_0) \\ c_1 \end{array}\right].$$

Thus we have

$$\vec{x}(t) = U\exp(\tilde{A}(t-t_0))\vec{c} = e^{3(t-t_0)}U\exp(\tilde{N}(t-t_0))\vec{c} =$$
$$e^{3(t-t_0)}\left[\begin{array}{cc} 1 & 0 \\ 3 & 1 \end{array}\right]\left[\begin{array}{c} c_0 + c_1(t-t_0) \\ c_1 \end{array}\right] = e^{3(t-t_0)}\left[\begin{array}{c} c_0 + c_1(t-t_0) \\ (3c_0 + c_1) + 3c_1(t-t_0) \end{array}\right].$$

In particular, the solution of the original second order differential equation is

$$x(t) = x_0(t) = [c_0 + c_1(t-t_0)]e^{3(t-t_0)}.$$

**Solution to (5)** The associated inhomogeneous first order linear IVP is

$$\frac{d\vec{x}}{dt}(t) = A\vec{x}(t) + \vec{g}(t), \quad \vec{x}(t_0) = \vec{0},$$

where the inhomogeneous term is

$$\vec{g}(t) = \left[\begin{array}{c} 0 \\ e^{3t} \end{array}\right].$$

The Green's function is

$$A(t,s) = A(t-s) = e^{3(t-s)}\left[\begin{array}{ll} -3(t-s)+1 & 1(t-s) \\ -9(t-s) & 3(t-s)+1 \end{array}\right].$$

This gives

$$A(t,s)\vec{g}(s) = e^{3(t-s)}\left[\begin{array}{ll} -3(t-s)+1 & 1(t-s) \\ -9(t-s) & 3(t-s)+1 \end{array}\right]\left[\begin{array}{c} 0 \\ e^{3s} \end{array}\right] = \left[\begin{array}{c} (t-s)e^{3t} \\ (3(t-s)+1)e^{3t} \end{array}\right].$$

So by , the solution is

$$\vec{x}(t) = \int_{t_0}^{t} A(t,s)\vec{g}(s)ds = \int_{t_0}^{t}\left[\begin{array}{c} (t-s)e^{3t} \\ (3(t-s)+1)e^{3t} \end{array}\right]ds =$$
$$e^{3t}\int_{t_0}^{t}\left[\begin{array}{c} (t-s) \\ 3(t-s)+1 \end{array}\right]ds =$$
$$e^{3t}\left[\begin{array}{c} (1/2)(t^2-t_0^2) \\ (3/2)(t^2-t_0^2)+(t-t_0) \end{array}\right].$$

So the solution of the IVP is

$$\vec{x}(t) = \left[\begin{array}{c} x_0(t) \\ x_1(t) \end{array}\right] = e^{3t}\left[\begin{array}{c} (1/2)(t^2-t_0^2) \\ (3/2)(t^2-t_0^2)+(t-t_0) \end{array}\right].$$

So the solution of the original second order, inhomogeneous, linear IVP is

$$x(t) = x_0(t) = \frac{1}{2}(t^2-t_0^2)e^{3t}.$$

Direct computation confirms this is the solution.