Variation of Parameters

Here we will explore how to solve linear differential equations using variation of parameters. This will be developed by considering an equivalent first order linear system of differential equations.

Recall that any system of differential equations can be converted into a first order system by giving names to all but the highest order derivatives and then replacing those derivatives with the first derivative of the new name for the next to highest order derivative. For example: Given $z''' + y'' = 0$ and $y'z = y^3$, let $x_1 = y$, $x_2 = y'$, $x_3 = z$, $x_4 = z'$, and $x_5 = z''$. Then our equivalent system is $x_1' = x_2$, $x_3' = x_4$, $x_4' = x_5$ (to encode the relations between variables) and $x_5' + x_2' = 0$, $x_1'x_3 = (x_1)^3$ (encoding the original equations).

Using this trick on an n-th order (homogeneous) linear differential equation will yield a system of n first order (homogeneous) linear differential equations. In particular,

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_2(t)y'' + a_1(t)y' + a_0(t)y = g(t) \tag{1}$$

becomes

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_{n-2}'(t) \\ x_{n-1}'(t) \\ x_n'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_{n-1}(t) & -a_{n-2}(t) & -a_{n-3}(t) & \cdots & -a_1(t) & -a_0(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-2}(t) \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ g(t) \end{bmatrix} \tag{2}$$

where $x_1 = y$, $x_2 = x_1'$ $(= y')$, $x_3 = x_2'$ $(= y'')$, $\ldots$, $x_n = x_{n-1}'$ $(= y^{(n-1)})$. Briefly, we write $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$.

Now let $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$ be an arbitrary first order linear system (of n equations). First, we need to solve the (homogeneous) companion equation $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$. Unfortunately, this is impossible in general. However, in some cases we can write down potentially useful formula using the matrix exponential.

Homogeneous Systems: The Matrix Exponential

First, consider two matrices A and B whose entries are functions in t and whose product AB is defined (i.e., the number of columns of A matches the number of rows of B). The (i,j)-entry of AB is $\displaystyle\sum_{k=1}^{n} a_{ik}b_{kj}$. Differentiating this an applying the product rule, we get $\displaystyle\sum_{k=1}^{n} a_{ik}'b_{kj} + \sum_{k=1}^{n} a_{ik}b_{kj}'$. In other words, the (i,j)-entry of $(AB)'$ is the (i,j)-entry of $A'B + AB'$. Thus $(AB)' = A'B + AB'$ so the product rule still holds for matrices!

Next, let A be an $n \times n$ matrix. In addition suppose that A commutes with its derivative A' (we differentiate matrices one entry at a time just like vector valued functions in multivariable calculus). Notice that the derivative of $A^0 = I$ is 0 (since all of the entries are constant). Suppose for some positive integer k, we have $(A^k)' = kA^{k-1}A'$ (which equals $kA'A^{k-1}$ since A and A' commute). Then $(A^{k+1})' = (A^kA)' = (A^k)'A + A^kA' = kA^{k-1}A'A + A^kA' = kA^{k-1}AA' + A^kA' = (k+1)A^kA'$. Therefore, as long as A and A' commute, we have $(A^n)' = nA^{n-1}A' = nA'A^{n-1}$ for all positive integers n (i.e., the power rule for non-negative integer powers).

Define $\exp(A) = \displaystyle\sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots$. While we will skip over convergence concerns, just know that this matrix exponential converges for all complex matrices A and does so in essentially the best possible way. Thus we can freely interchange limits and do all kinds of things that normally make analysts cringe. Suppose that A is a matrix of functions and A commutes with A'. Then $(\exp(A))' = \left(\displaystyle\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)' = \sum_{k=0}^{\infty} \frac{(A^k)'}{k!} = \sum_{k=0}^{\infty} \frac{kA^{k-1}A'}{k!} = \sum_{k=1}^{\infty} \frac{A^{k-1}}{(k-1)!}A' = \sum_{\ell=0}^{\infty} \frac{A^\ell}{\ell!}A' = \exp(A)A'$ (and likewise also $= A'\exp(A)$).

The above calculation shows that, given $A(t)$ and $\int_{t_0}^{t} A(u)\,du$ commute, if we let $M(t) = \exp\left(\int_{t_0}^{t} A(u)\,du\right)$, then $M'(t) = A(t)M(t)$. In other words, each column of $M(t)$ is a solution of our companion equation $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$. Moreover, $M(t_0) = \exp\left(\int_{t_0}^{t_0} A(u)\,du\right) = \exp(0) = I$ so that the j-th column of $M(t)$ solves the initial value problem $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ where $\mathbf{x}(t_0) = \mathbf{e}_j$ (the j-th standard unit vector). Also, if $\mathbf{x}(t) = M(t)\mathbf{x}_0$, then we solve the initial value problem with initial condition $\mathbf{x}(t_0) = M(t_0)\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0$.

We note here that a (square) matrix whose columns are linearly independent solutions of $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ is called a *fundamental matrix* for that homogeneous system. It can be shown that when A and B commute, we have $\exp(A+B) = \exp(A)\exp(B)$. Thus since A and $-A$ commute, we have $\exp(A)\exp(-A) = \exp(A-A) = \exp(0) = I$ and so $(\exp(A))^{-1} =$

$\exp(-A)$. In particular, matrix exponentials are always invertible. Therefore, $M(t) = \exp\left(\int_{t_0}^t A(u)\,du\right)$ is a fundamental matrix for $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ (if we know that $A(t)$ and $\int_{t_0}^t A(u)\,du$ commute).

A note of interest, if $M(t)$ is any fundamental matrix for $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$, then $B(t) = M(t)M(t_0)^{-1}$ is also a fundamental matrix but then $B(t_0) = M(t_0)M(t_0)^{-1} = I$. Thus, assuming $A(t)$ commutes with its antiderivative, $B(t)$ solves the same initial value problem as $\exp\left(\int_{t_0}^t A(u)\,du\right)$. Therefore, by the uniqueness part of the existence uniqueness theorem (of solutions of linear systems of differential equations), we must have that $M(t)M(t_0)^{-1} = \exp\left(\int_{t_0}^t A(u)\,du\right)$.

As a special case, when A is constant, $\int_{t_0}^t A\,du = Au\Big|_{t_0}^t = A(t-t_0)$, so $M(t) = \exp(A(t-t_0))$ is a fundamental matrix for $\mathbf{x}'(t) = A\mathbf{x}(t)$. In this case, we can effectively compute the matrix exponential using the Jordan form (say $P^{-1}AP = J = D + N$, then $\exp(A(t-t_0)) = P\exp(D(t-t_0))\exp(N(t-t_0))P^{-1}$ and exponentiating diagonal and nilpotent matrices is easy).

As an even more special case, still supposing A is constant, notice that if $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ where $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}(t)$ solves our homogeneous system. This means that if we have n linearly independent eigenvectors for A, we'll get n linearly independent solutions of $\mathbf{x}'(t) = A\mathbf{x}(t)$. In other words, when A is diagonalizable (this is the case considered in most undergraduate differential equations courses), then we can construct a fundamental matrix $M(t)$ whose columns are of the form $e^{\lambda t}\mathbf{v}$ where $\mathbf{v}$ is an eigenvector with eigenvalue λ. From our note of interest above, if $M(t)$ is such a fundamental matrix, then $M(t)M(t_0)^{-1} = \exp(A(t-t_0))$. So in some sense, the matrix exponential is sort of a normalized fundamental matrix.

Non-Homogeneous Systems: Variation of Parameters

Suppose we begin with a fundamental matrix $M(t)$ for the companion equation $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ related to our system $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$. We know that $\mathbf{x}(t) = M(t)\mathbf{c}$ (where $\mathbf{c}$ is an arbitrary constant vector) is a general solution of our companion equation. Thus if we replace $\mathbf{c}$ with a vector of non-constant functions $\mathbf{v}(t)$, we should get solutions for some non-homogeneous system. To that end suppose $\mathbf{x}(t) = M(t)\mathbf{v}(t)$ solves $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$. Then $\mathbf{x}'(t) = (M(t)\mathbf{v}(t))' = M'(t)\mathbf{v}(t) + M(t)\mathbf{v}'(t)$, but $(M(t)\mathbf{v}(t))' = \mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t) = A(t)M(t)\mathbf{v}(t) + \mathbf{g}(t)$. Taking into consideration that $M(t)$ is a fundamental matrix for the companion equation (so that $M'(t) = A(t)M(t)$), we have $A(t)M(t)\mathbf{v}(t) + M(t)\mathbf{v}'(t) = A(t)M(t)\mathbf{v}(t) + \mathbf{g}(t)$. Thus $M(t)\mathbf{v}'(t) = \mathbf{g}(t)$ and since $M(t)$ is invertible: $\mathbf{v}'(t) = M(t)^{-1}\mathbf{g}(t)$.

Therefore, letting $\mathbf{v}(t) = \int_{t_0}^t M(u)^{-1}\mathbf{g}(u)\,du$, we have that $\mathbf{x}(t) = M(t)\mathbf{v}(t) = M(t)\int_{t_0}^t M(u)^{-1}\mathbf{g}(u)\,du$ is a solution of our non-homogeneous system. Moreover, letting $\mathbf{x}(t) = M(t)\mathbf{v}(t) = M(t)\left(\int_{t_0}^t M(u)^{-1}\mathbf{g}(u)\,du + \mathbf{c}\right)$ (i.e., adding in a general solution of the companion equation: $M(t)\mathbf{c}$), we get a general solution of our system. Also, noting that $\mathbf{x}(t_0) = M(t_0)\left(\int_{t_0}^{t_0} M(u)^{-1}\mathbf{g}(u)\,du + \mathbf{c}\right) = M(t_0)\mathbf{c}$, we have that $\mathbf{x}(t) = M(t)\left(\int_{t_0}^t M(u)^{-1}\mathbf{g}(u)\,du + M(t_0)^{-1}\mathbf{x}_0\right)$ is the solution of the initial value problem $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$ with $\mathbf{x}(t_0) = \mathbf{x}_0$. Thus if we can solve the companion problem, then we can effectively solve the non-homogeneous problem too!

As a special case, notice that when $A(t)$ commutes with its derivative, we can let $M(t) = \exp\left(\int_{t_0}^t A(u)\,du\right)$ and so $M(t)^{-1} = \exp\left(-\int_{t_0}^t A(u)\,du\right)$. Thus $\mathbf{x}(t) = \exp\left(\int_{t_0}^t A(u)\,du\right)\left(\int_{t_0}^t \exp\left(-\int_{t_0}^w A(u)\,du\right)\mathbf{g}(w)\,dw + \mathbf{x}_0\right)$ solves the initial value problem $\mathbf{x}'(t) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$ with $\mathbf{x}(t_0) = \mathbf{x}_0$. This is the exact same formula as our solution for a single first order linear equation (the exponential of the integral of $A(t)$ is our integrating factor)!

Cramer's Rule, Matrix Inverses, and the Wronskian

Before we get back our original problem of solving n-th order linear differential equations, we take care of a computational issue: How do we invert a matrix – especially if its entries are functions?!? One possible answer is Cramer's Rule. This formula allows one to explicitly solve linear systems when the coefficient matrix is invertible.

Consider a linear system $A\mathbf{x} = \mathbf{b}$ where A is a square matrix. Let $A = [\ \mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n\ ]$ where $\mathbf{a}_j$ is the j-th column of A. Also, let $A_j = [\ \mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \mathbf{b}\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\ ]$ be the matrix A with its j-th column replaced by $\mathbf{b}$. Also, let $\mathbf{x} = [x_1\ x_2\ \cdots\ x_n]^T$ be a solution of our linear system. Then $\mathbf{a}_1x_1 + \cdots + \mathbf{a}_nx_n = A\mathbf{x} = \mathbf{b}$.

Next, let's take a determinant: $\det A_j =$

$$\det[\,\mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \mathbf{b}\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\,] = \det\left[\,\mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \sum_{k=1}^{n} x_k\mathbf{a}_k\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\,\right] = \sum_{k=1}^{n} x_k \det[\,\mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \mathbf{a}_k\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\,]$$

where we can take the sum out because determinants are multilinear in their columns. Now determinants are not just multilinear but also alternating in their columns. This means that if a column is repeated, the determinant is zero. Therefore, the only determinant in our sum above that survives is the one where $k = j$ (so we don't have a repeated column). Therefore, $\det A_j = x_j \det[\,\mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \mathbf{a}_j\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\,] = x_j \det A$. Thus, if $\det A$ is invertible, we have that $\mathbf{x} = [x_1\ x_2\ \cdots\ x_n]^T$ solves $A\mathbf{x} = \mathbf{b}$ if we let $x_i = \dfrac{\det A_j}{\det A}$ for each $i = 1, \ldots, n$. This is Cramer's Rule!

Example: Suppose we have a system $ax + by = p$ and $cx + dy = q$ where this system has a unqiue solution. Then Cramer tells us that
$$x = \frac{\det\begin{bmatrix} p & b \\ q & d \end{bmatrix}}{\det\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{pd - bq}{ad - bc} \qquad \text{and} \qquad y = \frac{\det\begin{bmatrix} a & p \\ c & q \end{bmatrix}}{\det\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{aq - cp}{ad - bc}.$$

Notice that if $\mathbf{e}_i$ is the i-th standard unit vector (i.e., the i-th column of the identity matrix), then solving $A\mathbf{x} = \mathbf{e}_i$ we have $\mathbf{x}$'s j-th entry must satisfy the equation, $\det A_j = x_j \det A$. Notice here that $A_j = [\,\mathbf{a}_1\ \cdots\ \mathbf{a}_{j-1}\ \mathbf{e}_i\ \mathbf{a}_{j+1}\ \cdots\ \mathbf{a}_n\,]$. The determinant of this matrix could be computed by expanding down the j-th column. We would sum a bunch of zeros until we got to row i (where $\mathbf{e}_i$ has the non-zero entry 1) and get $(-1)^{i+j}$ times the subdeterminant obtained by deleting row i and column j. In other words, if we let A_{ij} be the matrix A with row i and column j struck out, then this determinant is nothing more than $(-1)^{i+j}\det A_{ij}$. Therefore, a solution of $A\mathbf{x} = \mathbf{e}_i$ will have an j-th entry satisfying $(-1)^{i+1}\det A_{ij} = x_j \det A$.

Create a square matrix C and let its (i,j)-entry be $c_{ij} = (-1)^{i+1}\det A_{ij}$. Then the i-th row of C is $\det A$ times a solution of $A\mathbf{x} = \mathbf{e}_i$. This implies that $AC^T = (\det A)I$. Similarly, $C^T A = (\det A)I$. The matrix C^T is called the *classical adjoint transpose*. If $\det A$ is invertible, we have $A^{-1} = \dfrac{1}{\det A}C^T$. This gives us an explicit formula for the inverse of a matrix! We can compute it using additions, subtractions, multiplications of the entries of A followed by a single division (by the determinant of A).

Heading back to the problem of solving an n-th order linear equation, let $y_1, \ldots, y_n$ be a fundamental solution set for the companion equation of (1) (i.e., $y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0 y = 0$). These solutions correspond to solutions of the companion equations of (2) (i.e., $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$). In fact, since $x_1 = y$, $x_2 = y'$, $\ldots x_n = y^{(n-1)}$, we have that our fundamental solution set corresponds to the fundamental matrix:

$$M(t) = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix} \tag{3}$$

Conversely, the top row of a fundamental matrix for $\mathbf{x}'(t) = A(t)\mathbf{x}(t)$ yields a fundamental solution set for the companion equation of our n-th order linear differential equation.

Consider actually implementing our variation of parameters formulas, we now know that the classical adjoint transpose (essential a bunch of determinants) can take a fundamental matrix and compute its inverse. However, let us focus on the case arising from an n-th order linear equation. Here we really aren't interested in the entire column vector $\mathbf{x}(t)$. Instead we really just want the first entry since $y = x_1$. To that end, we ask what is $\mathbf{v}' = M^{-1}\mathbf{g}$? Well, by Cramer's rule, the i-th entry of $\mathbf{v}'$ (i.e., of the solution of $M\mathbf{v}' = \mathbf{g}$) is nothing more than

$$v_i'(t) = \frac{\det\begin{bmatrix} y_1 & \cdots & y_{i-1} & 0 & y_{i+1} & \cdots & y_n \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ y_1^{(n-2)} & \cdots & y_{i-1}^{(n-2)} & 0 & y_{i+1}^{(n-2)} & \cdots & y_n^{(n-2)} \\ y_1^{(n-1)} & \cdots & y_{i-1}^{(n-1)} & g(t) & y_{i+1}^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix}}{\det M(t)} = \frac{(-1)^{i+n}g(t)\det M_{ni}(t)}{\det M(t)}$$

where $M_{ni}(t)$ is the matrix $M(t)$ with its n-th row and i-th column struck out. Notice $M_{ni}(t)$ is a matrix filled with

$y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n$ and their derivatives up to order $n-2$. This leads to a definition. Let $z_1, \ldots, z_k$ be functions define

$$W[z_1, z_2, \ldots, z_k] = \det \begin{bmatrix} z_1 & z_2 & \cdots & z_k \\ z_1' & z_2' & \cdots & z_k' \\ \vdots & \vdots & & \vdots \\ z_1^{(k-1)} & z_2^{(k-1)} & \cdots & z_k^{(k-1)} \end{bmatrix}$$

This is called the *Wronskian* of the functions $z_1, \ldots, z_k$. More precisely, Wronskian might refer to the matrix and Wronskian determinant refer to its determinant. This is the same abuse of language as is commonplace when dealing with Jacobians. Using this new notation, after integrating, we have that

$$v_i(t) = \int_{t_0}^{t} \frac{(-1)^{i+n} g(u)\, W[y_1(u), \ldots, y_{i-1}(u), y_{i+1}(u), \ldots, y_n(u)]}{W[y_1(u), \ldots, y_n(u)]}\, du$$

Therefore, $y = v_1(t)y_1(t) + \cdots + v_n(t)y_n(t)$ is a solution of (1). Adding in a general solution of the companion equation yields a general solution of (1).

Example: Suppose we solved the companion equation of $y'' + a_1(t)y' + a_0(t)y = g(t)$ and found linearly independent solutions y_1 and y_2. Then $W[y_1, y_2] = \det \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} = y_1y_2' - y_1'y_2$. Notice that $W[z] = z$. Therefore,

$$v_1(t) = \int_{t_0}^{t} \frac{(-1)^{1+2} g(u) W[y_2]}{W[y_1, y_2]}\, du = -\int_{t_0}^{t} \frac{g(u)y_2(u)}{y_1(u)y_2'(u) - y_1'(u)y_2(u)}\, du$$

and

$$v_2(t) = \int_{t_0}^{t} \frac{(-1)^{2+2} g(u) W[y_1]}{W[y_1, y_2]}\, du = \int_{t_0}^{t} \frac{g(u)y_1(u)}{y_1(u)y_2'(u) - y_1'(u)y_2(u)}\, du$$

We have $y(t) = v_1(t)y_1(t) + v_2(t)y_2(t) + C_1y_1(t) + C_2y_2(t)$ is a general solution of our equation.

Liouville's Formula

An interesting fact about the Wronskian of a fundamental solution set is that it is non-zero on its entire domain. First, we let $M(t)$ be a fundamental matrix for some homogeneous system $\mathbf{x}'(t) = A(t)\mathbf{x}$. Let $W = \det M(t)$. In the special case that our system came from an n-th order linear equation, we called W the Wronskian of our solution set. We still call it a Wronskian in this more general setting.

Let $y_{ij}(t)$ be the (i,j)-entry in $M(t)$. Then $W = \det M(t) = \sum_{\sigma \in S_n} (-1)^\sigma y_{1\sigma(1)} \cdots y_{n\sigma(n)}$ where S_n is the group of permutations on n things and $(-1)^\sigma$ is the sign of the permutation σ (i.e., it is 1 when σ is even and -1 when odd). Take the derivative of W and apply the product rule: $W' = \sum_{\sigma \in S_n} (-1)^\sigma (y_{1\sigma(1)} \cdots y_{n\sigma(n)})' = \sum_{\sigma \in S_n} (-1)^\sigma y_{1\sigma(1)}' \cdots y_{n\sigma(n)} + \cdots + \sum_{\sigma \in S_n} (-1)^\sigma y_{1\sigma(1)} \cdots y_{n\sigma(n)}'$ (we are successively differentiating each row). Let M_i denote M with the i-th row replaced by its determinants (so that $W' = \det M_1 + \cdots + \det M_n$).

Fix some i and focus on M_i. Notice that M_i's entries are y_{kj} $(k \neq i)$ and y_{ij}' for any j. Now $M(t)$ is a fundamental matrix so that $M'(t) = A(t)M(t)$. This implies $y_{ij}' = \sum_{k=1}^{n} a_{ik}(t)y_{kj}$ for any j. If we add $-a_{ik}$ times row k to row i for all $k \neq i$, the resulting matrix's (i,j)-entry would be $y_{ij}' - \sum_{k=1,\ldots,i-1,i+1,\ldots,n} a_{ik}(t)y_{kj} = a_{ii}y_{ij}$ (for any j). In other words, by adding multiples of various rows to row i, we obtain a matrix identical to M except its i-th row is scaled by a_{ii}. Therefore, since adding multiples of rows to other rows does not change determinants and scaling a row just scales the determinant, we have $\det M_i = a_{ii} \det M$.

We have now shown that $W' = \det M_1 + \cdots + \det M_n = a_{11} \det M + \cdots + a_{nn} \det M = (a_{11} + \cdots + a_{nn}) \det M = \mathrm{trace}(A)\, W$. This is a first order (and actually separable) linear differential equation! We solve and get $W = Ce^{\int \mathrm{trace}(A(t))}$. Or more precisely, $W(t) = W(t_0)\exp\left(\int_{t_0}^{t} \mathrm{trace}(A(u))\, du\right)$. Since $M(t)$ is a fundamental matrix, its columns are linearly independent. *Note:* $M(t_0)$ is invertible so that $W(t_0) = \det M(t_0)$ is non-zero.

When our system comes from an n-th order linear differential equation, this formula takes on a special form and is called *Abel's formula.* Notice that the trace of the coefficient matrix $A(t)$ in (2) is $0 + \cdots + 0 - a_0(t) = -a_0(t)$. Therefore, for a fundamental solution set $y_1, \ldots, y_n$, we have $W[y_1, \ldots, y_n] = W[y_1, \ldots, y_n](t_0) \cdot \exp\left(\int_{t_0}^{t} -a_0(u)\, du\right)$.