Refactoring the ODE system... a proposal

ODEs solved in Stan have the form

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{f}(t, \mathbf{y}, \boldsymbol{\theta})$$

with the inital condition

$$\mathbf{y}(t = t_0) = \mathbf{y}_0,$$

which fully defines the solution $\mathbf{y}(t)$ at all time points $t > t_0$. Here the state $\mathbf{y}$ has N components and the parameter vector $\boldsymbol{\theta}$ is of size M. However, depending on what is requested, we also need the gradients of $\mathbf{y}(t)$ wrt to the initials $\mathbf{y}_0$ and/or the parameters $\boldsymbol{\theta}$, which we construct with what is referred to usually as *forward sensitivity analysis.*

The sensitivity of the solution wrt to a parameter p_i is $\frac{d\mathbf{y}(t)}{dp_i}$, which is the gradient of $\mathbf{y}(t)$ wrt to p_i. It can be shown that the N components of this gradient are related to the base ODE system as

$$\frac{d\mathbf{s}_i(t)}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \mathbf{s}_i + \frac{\partial \mathbf{f}}{\partial p_i} \text{ with } \mathbf{s}_i(t = t_0) = \frac{\partial \mathbf{y}_0}{\partial p_i}.$$

First we recognize that we always need the Jacobian wrt to the states, $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}$. Next, let's consider these terms for the two cases sensitivities for either the initials or the parameters:

- Sensitivities for the parameters: Then the term $\frac{\partial \mathbf{f}}{\partial p_i}$ is the ith column of the Jacobian wrt to the parameters, $\frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}$. The term $\frac{\partial \mathbf{y}_0}{\partial \theta_i}$ is 0 since the initial state does not depend on any parameter.
- Sensitivities for the initials: The term $\frac{\partial \mathbf{f}}{\partial p_i}$ vanishes, since the solution at time $t > t_0$ does not depend on the initial (in fact it does, but only implicitly). The term $\frac{\partial \mathbf{y}_0}{\partial p_i}$ is easiest written with the N terms collapsed to a matrix $(\frac{\partial \mathbf{y}_0}{\partial y_1}, ..., \frac{\partial \mathbf{y}_0}{\partial y_N})$ which are equal to the $\mathbf{1}_{NxN}$ identity matrix.

Collpasing now the two types of sensitivities in matrix notation with the definitions

$$\mathbf{S}_{\mathbf{y_0}} = \left(\frac{\partial \mathbf{y}}{\partial y_{0,1}}, ..., \frac{\partial \mathbf{y}}{\partial y_{0,N}} \right)$$

$$\mathbf{S}_{\boldsymbol{\theta}} = \left(\frac{\partial \mathbf{y}}{\partial \theta_1}, ..., \frac{\partial \mathbf{y}}{\partial \theta_M} \right)$$

let's us write the coupled ODE system in a compact matrix form, i.e.

$$\frac{d\mathbf{S}_{\mathbf{y_0}}}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \mathbf{S}_{\mathbf{y_0}}$$

$$\frac{d\mathbf{S}_\theta}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \mathbf{S}_\theta + \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}.$$

Hence, for Stan we need in addition to the ODE RHS, which is $\mathbf{f}$ in the notation above, also always the objects

- Jacobian wrt to states, $\frac{\partial \mathbf{f}}{\partial \mathbf{y}}$
- Jacobian wrt to parameters, $\frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}$

which we usually generate using autodiff. As any ODE solver used in Stan which uses the forward sensitivity approach will need the above objects, I am proposing to introduce the ode_model object. The ode_model object has as template parameter the ODE RHS system functor. As its default implementation it uses AD to generate the needed Jacobians. With template specialisations we are then able to inject analytic Jacobians into the ODE code. This will be of particular importance whenever AD becomes a performance bottleneck, i.e. for very stiff systems which need frequent evaluation of the Jacobian OR whenever the solution is requested with a high precision.