example derivation of a neural mass model

Tina Wentz

May 22, 2015

Abstract

Remember this is not the exact model we discussed in class, this is the derivation of the Jansen and Rit model (there are a few more equations in the David et al. model as well as a different sigmoid function - the derivation though is equivalent)

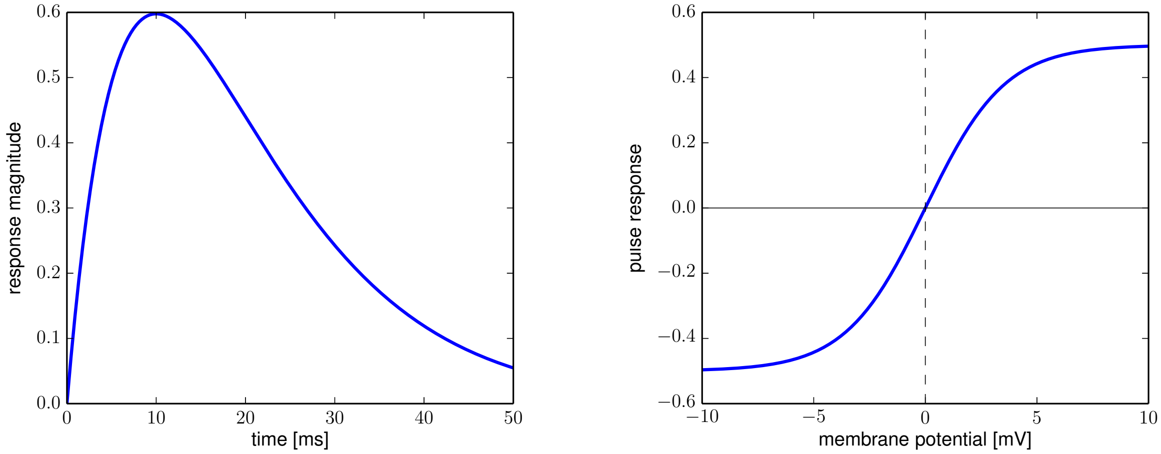

Figure 1: left panel: convolution kernel, right panel: sigmoid

1 one - neural mass

We know that different neurons work in concert to fulfil specific functions, for example an excitatory pyramidal cell that is influenced by an excitatory and an inhibitory interneuron, and that this behaviour could be described by a set of coupled differential equations. Solving (not to speak of inferring on) these equations for every single neuron becomes intractable even for small brain regions. Experiments have shown that neurons in close proximity behave the same when excited due to connections (the strength of these can change over time = plasticity (the synapses and differences in concentration of neurotransmitters in these are included in that term.)), find reference, which allows for a so called mean field assumption to be made in certain cases: we can describe a neuronal population by its average behaviour, specifically the average membrane potential and the mean firing rate. This has been used in the earliest models based on early physiological measurements, see for example **?** and references therein.

To simulate the behaviour of a cortical region with different populations we will follow the derivation in **?** for our slightly simpler example using local inhibitory and excitatory feedback and excitatory input which is represented by a pulse density $p(t)$. Let the average activity in an excitatory subpopulation be denoted as $E(t)$, $I(t)$ for inhibitory interneurons respectively. For each population the mean incoming firing rate m - depending on activity in connected cells and possibly external input $p(t)$ - is transformed into an average postsynaptic membrane potential v by convolving it with a kernel h:

$$v(t) = (m \otimes h)(t) \tag{1}$$

where $\otimes$ denotes the convolution operator and

$$h_e(t) = \begin{cases} \eta_e \frac{t}{\tau_e} exp[-t/\tau_e] & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \qquad h_i(t) = \begin{cases} \eta_i \frac{t}{\tau_i} exp[-t/\tau_i] & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{2}$$

η_e and η_i determine the maximum postsynaptic potential amplitude for excitatory and inhibitory cells, while τ_e and τ_i are average time constants for excitatory and inhibitory populations that represent time delays and decay rates for dendritic signal transport.

The average membrane potential can then be transformed into a firing rate using a sigmoid function such as:

$$S_k(v) = \frac{{c_k}^1 e_0}{1 + exp[r(v_0 - c_k^2 v)]}, \tag{3}$$

in which case r is... , the c_k^i are average numbers of synaptic connections, and v_0 provides a firing threshold.

For additional clarity: the convolution equations are

$$\begin{aligned} v_1(t) &= \int_0^t [E_2(t) + p(t)] h_e(t' - t) dt \\ v_2(t) &= \int_0^t I(t) h_i(t' - t) dt \\ v_3(t) &= \int_0^t E_1(t) h_e(t' - t) dt \end{aligned} \tag{4}$$

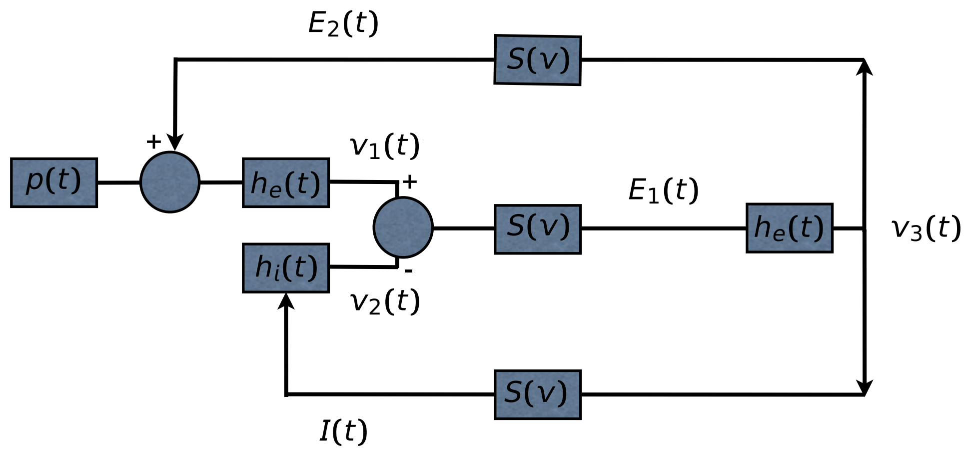

Figure 2: from Jansen&Rit - this is how they did it - remember that these are usually written in transfer functions so already in the s-domain.

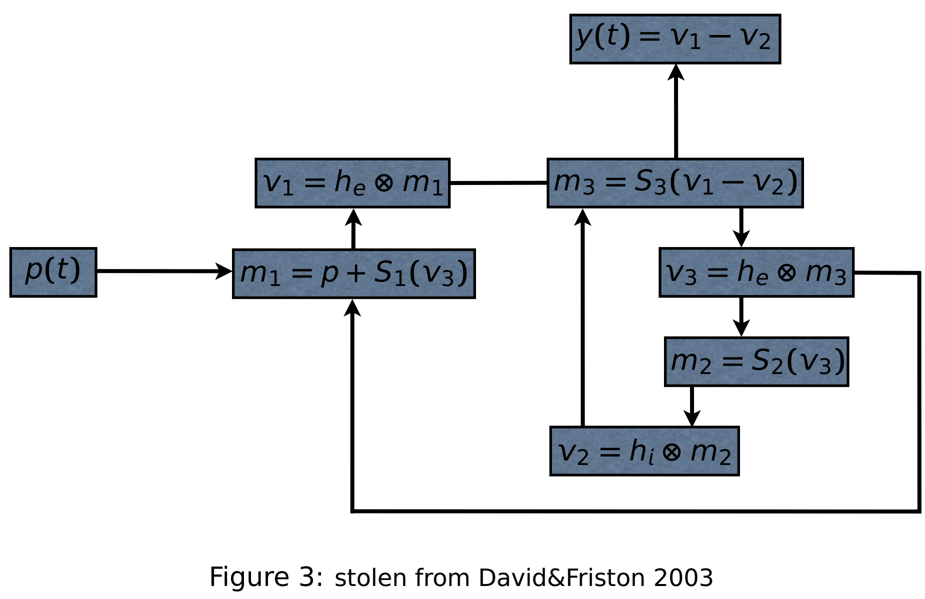

Figure 3: stolen from David&Friston 2003

using (17) - Laplace transformed (A, A.2):

$$\begin{aligned} V_1(s) &= [E_2(s) + P(s)]H_e(s) \\ V_2(s) &= I(s)H_i(s) \\ V_3(s) &= E_1(s)H_e(s) \end{aligned} \tag{5}$$

To find the transfer function for the pulse to the change in membrane potential, the kernel $h(t)$ that we chose above is transformed to the s-domain in a bilateral Laplace transform:

$$\begin{aligned} \mathcal{L}(h(t)) &:= \eta\kappa \int_0^\infty t e^{-\kappa t} e^{-st} dt \\ &= \eta\kappa \int_0^\infty t e^{-t(\kappa+s)} dt \\ &= \lim_{a\to\infty} \left[\eta\kappa \int_0^a t e^{-t(\kappa+s)} dt \right] \\ &= \lim_{a\to\infty} \left[\eta\kappa \Big[\frac{t}{-(\kappa+s)} - \frac{1}{(\kappa+s)^2} \Big] e^{-t(\kappa+s)} \Big|_0^a \right] \\ &= \lim_{a\to\infty} \left[\eta\kappa \frac{-t(\kappa+s)-1}{(\kappa+s)^2} e^{-t(\kappa+s)} \Big|_0^a \right] \\ &= \lim_{a\to\infty} \left[\eta\kappa \underbrace{\frac{-a(\kappa+s)-1}{(\kappa+s)^2} e^{-a(\kappa+s)}}_{0 \text{ for } a\to\infty} - \frac{-\eta\kappa}{(\kappa+s)^2} \right] \\ &= \frac{\eta\kappa}{(\kappa+s)^2} \end{aligned} \tag{6}$$

now that we got that: inverse Laplace to get differential equations:

$$\begin{aligned} V_1(s) &= [E_2(s) + P(s)]H_e(s) \\ &= \frac{\eta_e\kappa_e[E_2(s) + P(s)]}{(\kappa_e + s)^s} \\ &\Rightarrow \\ (\kappa_e^2 + 2\kappa_e s + s^2)V_1(s) &= \eta_e\kappa_e[E_2(s) + P(s)] \\ s^2V_1(s) + 2\kappa_e sV_1(s) + \kappa_e^2 V_1(s) &= \eta_e\kappa_e[E_2(s) + P(s)] \\ &\text{inverse Laplace:} \\ \mathcal{L}^{-1}\left(s^2V_1(s) + 2\kappa_e sV_1(s) + \kappa_e^2 V_1(s)\right) &= \mathcal{L}^{-1}\left(\eta_e\kappa_e[E_2(s) + P(s)]\right) \\ &\overset{(14)}{\Rightarrow} \\ \ddot{v}_1(t) + 2\kappa_e\dot{v}_1(t) + \kappa_e^2 v_1(t) &= \eta_e\kappa_e[E_2(t) + p(t)] \\ &\Rightarrow \\ \ddot{v}_1(t) &= \eta_e\kappa_e[E_2(t) + p(t)] - 2\kappa_e\dot{v}_1(t) - \kappa_e^2 v_1(t) \\ &\text{using the convention } \dot{v} = x \text{ and thus } \ddot{v} = \dot{x} \\ \dot{x}_1 &= \eta_e\kappa_e[E_2(t) + p(t)] - 2\kappa_e x_1(t) - \kappa_e^2 v_1(t) \end{aligned} \tag{7}$$

see Appendix section B for the other two, giving us exactly the set of differential equations of the **?** neural mass model, that can be expanded to include more populations and coupled to other such sets to model multiple connected regions, see for example **?**.

A the Laplace transform

A Laplace transform changes the representation of a signal from the time-domain to the so called s-domain which is equivalent to the frequency domain of Fourier transforms. Let $f(t)$, $t \in \mathbb{R}$ be defined for $t \geq 0$ and $s \in \mathbb{C}$, then the Laplace transform is defined as:

$$\mathcal{L}(f, t \to s) = \int_0^\infty f(t)e^{-st} dt, \tag{8}$$

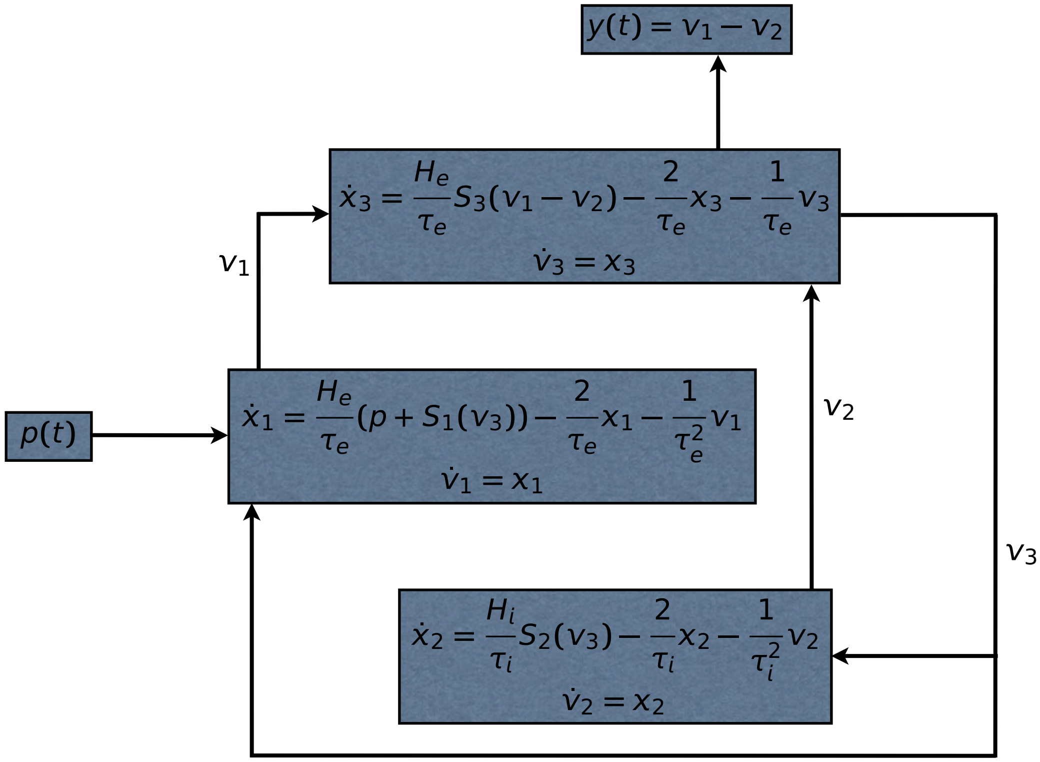

Figure 4: stolen from David&Friston 2003

provided that the integral exists. Sufficient, but not necessary conditions are:
- $f(t)$ is piecewise continuous on every finite closed interval $0 \leq t < a$, $\forall a > 0$
- $f(t)$ is of exponential order: $\exists b, M > 0$ and $t_0 > 0$ such that $|f(t)| \leq Me^{bt}$ for $t > t_0$
note: this is always satisfied if f(t) is a bounded function.
one can then prove that the integral converges for $s > b$.

So, for example:
let:

$$\begin{aligned}\mathcal{L}(x(t)) &:= X(s) \\ &= \int_0^\infty x(t) e^{-st} dt\end{aligned} \tag{9}$$

then:

$$\begin{aligned}\mathcal{L}\left(\frac{dx}{dt}\right) &= \int_0^\infty \frac{dx}{dt} e^{-st} dt \\ &= \lim_{a\to\infty} \left[\underbrace{x(t) e^{-st} \Big|_0^a}_{0 \text{ for } a\to\infty} \right] + s \int_0^\infty x(t) e^{-st} dt \\ &= x(0) + sX(s) \\ &= x(0) + s\,\mathcal{L}(x(t))\end{aligned} \tag{10}$$

which is the part of the beauty of the Laplace transform: Taking a derivative reduces to multiplication with s in the frequency domain. Using this last result we can derive the Laplace transform for a second derivative:

$$\begin{aligned}
\mathcal{L}\left(\frac{d^2x}{dt^2}\right) &= \int_0^\infty \frac{d^2x}{dt^2} e^{-st} dt \\
&= s\,\mathcal{L}\left(\frac{dx}{dt}\right) - \frac{dx}{dt}\Big|_{t=0} \\
&= s\Big[s\,\mathcal{L}(x(t)) - x(0)\Big] - \frac{dx}{dt}\Big|_{t=0} \\
&= s^2 X(s) - sx(0) - \frac{dx}{dt}\Big|_{t=0} \\
&= s^2\,\mathcal{L}(x(t)) - sx(0) - \frac{dx}{dt}\Big|_{t=0}
\end{aligned} \tag{11}$$

A.1 transfer functions

Transfer functions are used to describe the relation between input and output of a time-invariant, linear system with zero initial conditions and with zero-point equilibrium. They are defined using the bilateral Laplace transformation:

$$\begin{aligned}
\mathcal{L}(x(t)) &:= X(s) \\
&= \int_{-\infty}^{\infty} x(t) e^{-st} dt
\end{aligned} \tag{12}$$

which is equivalent to assuming null initial conditions in a unilateral Laplace transformation.
Let $X(s)$ be the s-domain representation of the input to a time-invariant linear system, and similarly $Y(s)$ the output. The transfer function is then defined as:

$$G(s) = \frac{Y(s)}{X(s)} \tag{13}$$

It can easily be shown that the bilateral Laplace transforms of time derivatives are:

$$\begin{aligned}
\mathcal{L}\left(\frac{dx}{dt}\right) &= s\,\mathcal{L}(x(t)) = sX(s) \\
\mathcal{L}\left(\frac{d^2x}{dt^2}\right) &= s^2\,\mathcal{L}(x(t)) = s^2 X(s)
\end{aligned} \tag{14}$$

Note that this makes the conversion from transfer functions to differential equations extremely simple.

A.2 convolution theorem

We will now prove the convolution theorem for Laplace transforms:
rewriting the simplest possible version of the equations (4) we realise that this is just the definition of a convolution:

$$v(t) := f \otimes h(t) = \int_{x=0}^{t} f(x) h(t-x) dx \tag{15}$$

we can now show that the convolution is just a product in frequency space (and thus commutable: $f \otimes h(t) = h \otimes f(t)$).
Let:

$$\begin{aligned}
F(s) &= \mathcal{L}(f(x))\Big|_s = \int_0^\infty f(x) e^{-sx} dx \\
H(s) &= \mathcal{L}(h(t))\Big|_s = \int_0^\infty h(t) e^{-st} dt
\end{aligned} \tag{16}$$

then:

$$\begin{aligned}
\mathcal{L}(v(t)) &:= V(s) \\
&= \int_{t=0}^{\infty} V(t)e^{-st}dt \\
&= \int_{t=0}^{\infty} \Big[\int_{x=0}^{t} f(x)h(t-x)dx\Big]e^{-st}dt \\
&= \int_{t=0}^{\infty}\int_{x=0}^{t} e^{-st}f(x)h(t-x)dxdt \\
\text{using: } e^{-st}=e^{-st+sx-sx}=e^{-s(t-x)}e^{-sx}& \\
&= \int_{t=0}^{\infty}\int_{x=0}^{t} e^{-s(t-x)}e^{-sx}f(x)h(t-x)dxdt \\
\text{changing order of integration:}& \\
&= \int_{x=0}^{\infty}\int_{t=x}^{\infty} e^{-s(t-x)}e^{-sx}f(x)h(t-x)dtdx \\
&= \int_{x=0}^{\infty} e^{-sx}f(x)\Big[\int_{t=x}^{\infty} e^{-s(t-x)}h(t-x)dt\Big]dx \\
\text{Let: } t'=t-x& \\
&= \int_{x=0}^{\infty} e^{-sx}f(x)\Big[\int_{t'=x-x}^{\infty-x} e^{-st'}h(t')dt\Big]dx \\
\text{now that integrals are independent:}& \\
&= \int_{x=0}^{\infty} e^{-sx}f(x)dx\int_{t'=0}^{\infty} e^{-st'}h(t')dt \\
\text{(16):}& \\
&= F(s)H(s)
\end{aligned} \tag{17}$$

B inversion for v_2 and v_3

for v2:

$$\begin{gathered}
V_2(s) = I(s)H_i(s) = \frac{\eta_i\kappa_i I(s)}{(\kappa_i+s)^s} \\
\Rightarrow \\
s^2V_2(s)+2\kappa_i sV_2(s)+\kappa_i^2V_2(s) = \eta_i\kappa_i I(s) \\
\text{inverse Laplace:} \\
\mathcal{L}^{-1}\left(s^2V_2(s)+2\kappa_i sV_2(s)+\kappa_i^2V_2(s)\right) = \mathcal{L}^{-1}\left(\eta_i\kappa_i I(s)\right) \\
\ddot{v}_2(t)+2\kappa_i\dot{v}_2(t)+\kappa_i^2v_2(t) = \eta_i\kappa_i I(t) \\
\Rightarrow \\
\dot{x}_2 = \eta_i\kappa_i I(t) - 2\kappa_i x_2(t) - \kappa_i^2 v_2(t)
\end{gathered} \tag{18}$$

and v3:

$$\begin{gathered}
V_3(s) = E_1(s)H_e(s) = \frac{\eta_e\kappa_e E_1(s)}{(\kappa_e+s)^s} \\
\Rightarrow \\
s^2V_3(s)+2\kappa_e sV_3(s)+\kappa_e^2V_3(s) = \eta_e\kappa_e E_1(s) \\
\text{inverse Laplace:} \\
\mathcal{L}^{-1}\left(s^2V_3(s)+2\kappa_e sV_3(s)+\kappa_e^2V_3(s)\right) = \mathcal{L}^{-1}\left(\eta_e\kappa_e E_1(s)\right) \\
\ddot{v}_3(t)+2\kappa_e\dot{v}_3(t)+\kappa_e^2v_3(t) = \eta_e\kappa_e E_1(t) \\
\Rightarrow \\
\dot{x}_3 = \eta_e\kappa_e E_1(t) - 2\kappa_e x_3(t) - \kappa_e^2 v_3(t)
\end{gathered} \tag{19}$$

C table of variables - following DavidFriston nomenclature

$v(t)$	(average) postsynaptic membrane potential
$p(t)$	pulse density (input)
m	mean incoming firing rate
$h_{e,i}(t)$	pulse-to-wave kernel
$S_k(v)$	wave-to-pulse transform
$\tau_{e,i}$	average time constants representing l.....
$\eta_{e,i}$	maximum postsynaptic potential amplitude
$c_k^{1,2}$	
e_0	
r	
v_0	