Appendix

Following [**georges1991expand**] lets define energy as:

$$E = -\sum_{ij} w_{ij} s_i s_j - \sum_i \theta_i s_i \tag{1}$$

and introduce the following operator:

$$U \equiv E - \mathbb{E}(E) - \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta}(s_i - m_i) \tag{2}$$

which poses useful property – $\mathbb{E}(U) = 0$. For any other operator O we then have:

$$\frac{\partial \mathbb{E}(O)}{\partial \beta} = \mathbb{E}\left(\frac{\partial O}{\partial \beta}\right) - \mathbb{E}(OU). \tag{3}$$

Now, the first derivative from the Taylor expansion is:

$$\begin{aligned}\frac{\partial(\beta F)}{\partial \beta} =& \frac{\sum_{\mathbf{s}} \exp\left(\beta \sum_{(ij)} w_{ij} s_i s_j + \beta \sum_i \theta_i s_i + \sum_i \lambda_i(\beta)(s_i - m_i)\right)\left(\sum_{(ij)} w_{ij} s_i s_j + \sum_i \theta_i s_i + \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta}(s_i - m_i)\right)}{\sum_{\mathbf{s}} \exp\left(\beta \sum_{(ij)} w_{ij} s_i s_j + \sum_i \theta_i s_i + \sum_i \lambda_i(\beta)(s_i - m_i)\right)} \\ =& \sum_{(ij)} w_{ij} \mathbb{E}(s_i s_j) + \sum_i a_i \mathbb{E}(s_i) + \frac{\partial \lambda_i(\beta)}{\partial \beta} \sum_i \mathbb{E}(s_i - m_i).\end{aligned}$$

In the case of $\beta = 0$ we have:

$$\left.\frac{\partial(\beta F)}{\partial \beta}\right|_{\beta=0} = \sum_{(ij)} w_{ij} m_i m_j + \sum_i \theta_i m_i.$$

Using 3 we obtain:

$$\frac{\partial m_i}{\partial \beta} = 0 = \frac{\partial \mathbb{E}(s_i)}{\partial \beta} = \mathbb{E}\left(\frac{\partial s_i}{\partial \beta}\right) - \mathbb{E}(s_i U) = -\mathbb{E}(s_i U) = -\mathbb{E}(U(s_i - m_i)). \tag{4}$$

The first derivative of the operator U has the form:

$$\begin{aligned}\frac{\partial U}{\partial \beta} &= \frac{\partial E}{\partial \beta} - \frac{\partial \mathbb{E}(E)}{\partial \beta} - \sum_i \frac{\partial^2 \lambda_i(\beta)}{\partial \beta^2}(s_i - m_i) \\ &= \mathbb{E}(U^2) - \sum_i \frac{\partial^2 \lambda_i(\beta)}{\partial \beta^2}(s_i - m_i)\end{aligned} \tag{5}$$

and the second derivative is:

$$\begin{aligned}\frac{\partial^2 U}{\partial \beta^2} &= 2\mathbb{E}\left(\frac{\partial U}{\partial \beta} U\right) - \mathbb{E}(U^3) - \sum_i \frac{\partial^3 \lambda_i(\beta)}{\partial \beta^3}(s_i - m_i) \\ &= -\mathbb{E}(U^3) - \sum_i \frac{\partial^3 \lambda_i(\beta)}{\partial \beta^3}(s_i - m_i)\end{aligned} \tag{6}$$

The expansion of free energy using formulas derived above can now be reformulated in terms of the operator U:

$$\frac{\partial(\beta F)}{\partial \beta} = \mathbb{E}(E) - \sum_i \frac{\partial \lambda_i(\beta)}{\partial \beta} \mathbb{E}(s_i - m_i) = \mathbb{E}(E) \tag{7}$$

and the higher orders:

$$\begin{aligned}\frac{\partial^2(\beta F)}{\partial \beta^2} &= \mathbb{E}\left(\frac{\partial E}{\partial \beta}\right) - \mathbb{E}(EU) = -\mathbb{E}(U^2), \\ \frac{\partial^3(\beta F)}{\partial \beta^3} &= -2\mathbb{E}\left(U \frac{\partial U}{\partial \beta}\right) + \mathbb{E}(U^3) = \mathbb{E}(U^3).\end{aligned} \tag{8}$$

Taylor expansion was considered around point $\beta = 0$. Using derivations from 7 we obtain again a 'naive' term:

$$\left.\frac{\partial(\beta F)}{\partial\beta}\right|_{\beta=0} = \mathbb{E}_{\beta=0}(E) = -\sum_{(ij)} w_{ij} m_i m_j - \sum_i \theta_i m_i = -\frac{1}{2}\sum_i\sum_j w_{ij} m_i m_j - \sum_i \theta_i m_i. \tag{9}$$

Consider now:

$$\left.\frac{\partial(\beta F)}{\partial m_i \partial\beta}\right|_{\beta=0} = -\sum_{j\neq i} w_{ij} m_j - \theta_i. \tag{10}$$

On the other hand:

$$\left.\frac{\partial(\beta F)}{\partial m_i \partial\beta}\right|_{\beta=0} = \left.\frac{\partial(\beta F)}{\partial\beta\partial m_i}\right|_{\beta=0} = \frac{\partial}{\partial\beta}\mathbb{E}(\lambda_i(\beta)) = \left.\frac{\partial\lambda_i(\beta)}{\partial\beta}\right|_{\beta=0} \tag{11}$$

Substituting 11 into 2 gives us:

$$\begin{aligned} U_{\beta=0} &= -\sum_{(ij)} w_{ij} s_i s_j - \sum_i \theta_i s_i + \frac{1}{2}\sum_i\sum_j w_{ij} m_i m_j + \sum_i \theta_i m_i + \sum_i \left(\sum_{j\neq i} w_{ij} m_j + \theta_i\right)(s_i - m_i) \\ &= -\sum_{(ij)} w_{ij} s_i s_j - \frac{1}{2}\sum_{(ij)} w_{ij} m_i m_j + \sum_i\sum_{j\neq i} w_{ij} s_i m_j \\ &= -\sum_{(ij)} w_{ij}(s_i - m_i)(s_j - m_j) = -\sum_l w_l y_l \end{aligned} \tag{12}$$

where $w_l = w_{ij}$ and $y_l = (s_i - m_i)(s_j - m_j)$ stands for the 'link' operator which poses useful properties:

$$\begin{aligned} \mathbb{E}(y_l)_{\beta=0} &= \mathbb{E}(s_i s_j) - m_j\mathbb{E}(s_i) - m_i\mathbb{E}(s_j) + m_i m_j = 0 \\ \mathbb{E}(y_l(s_i - m_i))_{\beta=0} &= m_j - m_j - m_i^2 m_j + m_i^2 m_j \\ &\quad - m_i^2 m_j + m_i^2 m_j + m_i^2 m_j - m_i^2 m_j \\ &= 0. \end{aligned} \tag{13}$$

Finally, if $k \neq l$ then:

$$\mathbb{E}(y_k y_l) = \mathbb{E}(y_k)\mathbb{E}(y_l) = 0$$

while for $k = l$ we have:

$$\begin{aligned} \mathbb{E}((s_i - m_1)^2(s_j - m_j)^2) &= m_i m_j - 2m_i m_j^2 + m_i m_j^2 - 2m_i^2 m_j + 4m_1^2 m_j^2 \\ &\quad - 2m_i^2 m_j^2 + m_i^2 m_j - 2m_i^2 m_j^2 + m_i^2 m_j^2 \\ &= (m_i - m_i^2)(m_j - m_j^2). \end{aligned} \tag{14}$$

Using properties from y_l in equations 8 we can derive:

$$\begin{aligned} \left.\frac{\partial^2(\beta F)}{\partial\beta^2}\right|_{\beta=0} &= -\,\mathbb{E}(U^2)_{\beta=0} \\ &= -\sum_{l_i l_2} w_{l_i} w_{l_2}\mathbb{E}_{\beta=0}(y_{l_1} y_{l_2}) \\ &= -\sum_{(i,j)} w_{ij}^2 (m_i - m_i^2)(m_j - m_j^2) \end{aligned}$$

which yields the TAP-Onsager term. To obtain the next term in the Taylor expansion we need to compute $\mathbb{E}(y_{l_1} y_{l_2} y_{l_3})$ term and by definition the structure of the RBM model doesn't admit triangles in its corresponding factor graphs. Thus, we need to consider only the case when $l_1 = l_2 = l_3$:

$$\begin{aligned} \mathbb{E}((s_i - m_1)^3(s_j - m_j)^3) &= m_i m_j - 3m_i m_j^2 + 2m_i m_j^2 + 2m_i m_j^3 - 3m_i^2 m_j + 2m_i^3 m_j \\ &\quad + 9m_i^2 m_j^2 - 6m_i^3 m_j^2 - 6m_i^2 m_j^3 + 4m_i^3 m_j^3 \\ &= 4(m_i - m_i^2)(\frac{1}{2} - m_i)(m_j - m_j^2)(\frac{1}{2} - m_j) \end{aligned} \tag{15}$$

and the third-order term in the case of the RBM structure:

$$\left.\frac{\partial^3(\beta F)}{\partial\beta^3}\right|_{\beta=0} = \frac{2\beta^3}{3}\sum_{(ij)} w_{ij}^3 (m_i - m_i^2)(\frac{1}{2} - m_i)(m_j - m_j^2)(\frac{1}{2} - m_j).$$