A Note on the Identities in Section 5.2 Here's a derivation that hopefully clarifies some of the steps in section 5.2 of the note on EM.

Say that we could efficiently calculate the following quantities for any x of length n, for any $j \in 1 \ldots n$, and for any $p \in 1 \ldots (N-1)$ and $q \in 1 \ldots N$:

$$P(y_j = p, y_{j+1} = q | x, \Theta) = \sum_{y: y_j = p, y_{j+1} = q} P(y|x, \Theta) \tag{1}$$

The inner sum can now be re-written using terms such as that in Eq. 1, as

$$\sum_y P(y|x^i, \Theta^{t-1}) Count(x^i, y, p \rightarrow q) = \sum_{j=1}^{n_i} P(y_j = p, y_{j+1} = q | x^i, \Theta^{t-1})$$

This identity was stated in section 5.2 on the note on EM; this note gives a justification for the identity.

To see why this is true, define

$$g(y, j, p, q)$$

to be 1 if $y_j = p$ and $y_{j+1} = q$, and 0 otherwise. It then follows that

$$Count(x^i, y, p \rightarrow q) = \sum_{j=1}^{n_i} g(y, j, p, q)$$

We can then write

$$\begin{aligned}
\sum_y P(y|x^i, \Theta^{t-1}) Count(x^i, y, p \rightarrow q) &= \sum_y P(y|x^i, \Theta^{t-1}) \sum_{j=1}^{n_i} g(y, j, p, q) \\
&= \sum_{j=1}^{n_i} \sum_y P(y|x^i, \Theta^{t-1}) g(y, j, p, q) \\
&= \sum_{j=1}^{n_i} P(y_j = p, y_{j+1} = q | x^i, \Theta^{t-1})
\end{aligned}$$

where the last line follows because

$$\sum_y P(y|x^i, \Theta^{t-1}) g(y, j, p, q) = P(y_j = p, y_{j+1} = q | x^i, \Theta^{t-1})$$

A similar argument can be used to derive the other identities, for example

$$\sum_y P(y|x^i, \Theta^{t-1}) Count(x^i, y, p \uparrow o) = \sum_{j: x_j = o} P(y_j = p | x^i, \Theta^{t-1})$$