Vamos a hallar $p(x_b | x_a)$ con $p(x) = p(x_a, x_b)$

Término del exponente de $p(x_b|x_a)$:

$$= -\frac{1}{2}(x_b - \mu_{b|a})^T \Sigma_{b|a}^{-1}(x_b - \mu_{b|a})$$

$$= \boxed{-\frac{1}{2} x_b^T \Sigma_{b|a}^{-1} x_b} + \underbrace{x_b^T \Sigma_{b|a}^{-1} \mu_{b|a}} - \frac{1}{2}\mu_{b|a}^T \Sigma_{b|a}^{-1} \mu_{b|a}$$

Por otro lado,

$$-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)$$

$$= -\frac{1}{2}([x_a, x_b] - [\mu_a, \mu_b])^T \begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix} ([x_a, x_b] - [\mu_a, \mu_b])$$

$$= -\frac{1}{2}\left[(x_a - \mu_a)^T \Lambda_{aa} + (x_b - \mu_b)^T \Lambda_{ba}, \; (x_a - \mu_a)^T \Lambda_{ab} + (x_b - \mu_b)^T \Lambda_{bb}\right]([x_a, x_b] - [\mu_a, \mu_b])$$

$$= -\frac{1}{2}\Big((x_a - \mu_a)^T \Lambda_{aa}(x_a - \mu_a) + (x_b - \mu_b)^T \Lambda_{ba}(x_a - \mu_a)$$
$$+ (x_a - \mu_a)^T \Lambda_{ab}(x_b - \mu_b) + (x_b - \mu_b)^T \Lambda_{bb}(x_b - \mu_b)\Big)$$

Teniendo en cuenta que Λ es simétrica.

$$\Lambda_{ab}^T = \Lambda_{ba}$$

$$= -\frac{1}{2} x_a^T \Lambda_{aa} x_a + x_a^T \Lambda_{aa} \mu_a - \frac{1}{2} \mu_a^T \Lambda_{aa} \mu_a$$

$$\underbrace{-\frac{1}{2} x_b^T \Lambda_{ba} x_a} + \underbrace{\frac{1}{2} x_b^T \Lambda_{ba} \mu_a} + \frac{1}{2} \mu_b^T \Lambda_{ba} x_a$$

$$-\frac{1}{2} \mu_b^T \Lambda_{ba} \mu_a \underbrace{-\frac{1}{2} x_a^T \Lambda_{ab} x_b} + \frac{1}{2} x_a^T \Lambda_{ab} \mu_b$$

$$\underbrace{+\frac{1}{2} \mu_a^T \Lambda_{ab} x_b} - \frac{1}{2} \mu_a^T \Lambda_{ab} \mu_b \boxed{-\frac{1}{2} x_b^T \Lambda_{bb} x_b}$$

$$\underbrace{+ x_b^T \Lambda_{bb} \mu_b} - \frac{1}{2} \mu_b^T \Lambda_{bb} \mu_b$$

Asumiendo x_a constante (ya que estamos hallando $p(x_b | x_a)$)

$\boxed{}$ Términos cuadrático

$\underbrace{}$ Términos lineales (el resto es constante)

- Igualando los términos cuadrados:

$$-\frac{1}{2} x_b^T \Sigma_{b|a}^{-1} x_b = -\frac{1}{2} x_b \Lambda_{bb} x_b$$

$$\therefore \quad \Sigma_{b|a}^{-1} = \Lambda_{bb}$$

- Factorizamos los términos lineales

$$-\frac{1}{2} x_b^T \Lambda_{ba} x_a + \frac{1}{2} x_b^T \Lambda_{ba} \mu_a - \frac{1}{2} x_a^T \Lambda_{ab} x_b$$

$$+\frac{1}{2} \mu_a^T \Lambda_{ab} x_b + x_b^T \Lambda_{bb} \mu_b$$

$$= -\frac{1}{2} x_b^T \Lambda_{ba} x_a + \frac{1}{2} x_b^T \Lambda_{ba} \mu_a - \frac{1}{2} x_b^T \Lambda_{ab} x_a + \frac{1}{2} x_b^T \Lambda_{ba} \mu_a + x_b^T \Lambda_{bb} \mu_b$$

$$= - x_b^T \Lambda_{ba} x_a + x_b^T \Lambda_{ba} \mu_a + x_b^T \Lambda_{bb} \mu_b$$

$$= x_b^T \left(\Lambda_{bb} \mu_b + \Lambda_{ba} \mu_a - \Lambda_{ba} x_a \right)$$

$$= x_b^T \left(\Lambda_{bb} \mu_b + \Lambda_{ba} (\mu_a - x_a) \right)$$

- Igualamos los términos lineales

$$x_b^T \Sigma_{b|a}^{-1} \mu_{b|a} = x_b^T \left(\Lambda_{bb} \mu_b + \Lambda_{ba} (\mu_a - x_a) \right)$$

$$\Rightarrow \quad \Sigma_{b|a}^{-1} \mu_{b|a} = \Lambda_{bb} \mu_b + \Lambda_{ba} (\mu_a - x_a)$$

$$\Rightarrow \quad \Sigma_{b|a} \Sigma_{b|a}^{-1} \mu_{b|a} = \Sigma_{b|a} \Lambda_{bb} \mu_b + \Sigma_{b|a} \Lambda_{ba} (\mu_a - x_a)$$

Recordemos que $\Sigma_{b|a}^{-1} = \Lambda_{bb}$

$$\Rightarrow \quad \mu_{b|a} = \Sigma_{b|a} \Sigma_{b|a}^{-1} \mu_b + \Sigma_{b|a} \Lambda_{ba} (\mu_a - x_a)$$

$$\mu_{b|a} = \mu_b + \Sigma_{b|a} \Lambda_{ba} (\mu_a - x_a)$$

Ahora, la idea es dejar $\mu_{b|a}$ y $\Sigma_{b|a}$ en términos conocidos ($\Sigma_{aa}, \Sigma_{ab}, \Sigma_{bb}$)

Para ello, vamos a utilizar la identidad de la matriz inversa por partes

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{bmatrix}$$

Siendo $M = \underbrace{(A - BD^{-1}C)^{-1}}_{\text{complemento de Schur}}$

Aplicándolo a nuestro caso

$$\begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix} = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} M & -M\Sigma_{ab}\Sigma_{bb}^{-1} \\ -\Sigma_{bb}^{-1}\Sigma_{ba}M & \Sigma_{bb}^{-1} + \Sigma_{bb}^{-1}\Sigma_{ba}M\Sigma_{ab}\Sigma_{bb}^{-1} \end{bmatrix}$$

donde $M = \left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}$

Así,

- $\Sigma_{b|a} = \Lambda_{bb}^{-1}$

$$= \left(\Sigma_{bb}^{-1} + \Sigma_{bb}^{-1}\Sigma_{ba}M\Sigma_{ab}\Sigma_{bb}^{-1}\right)^{-1}$$

$$= \left(\Sigma_{bb}^{-1} + \Sigma_{bb}^{-1} \Sigma_{ba} \left(\Sigma_{aa} - \Sigma_{ab} \Sigma_{bb}^{-1} \Sigma_{ba} \right)^{-1} \Sigma_{ab} \Sigma_{bb}^{-1} \right)^{-1}$$

- $$\mu_{b|a} = \mu_b + \Sigma_{b|a} \Lambda_{ba} (\mu_a - x_a)$$

$$= \mu_b + \left(\Sigma_{bb}^{-1} + \Sigma_{bb}^{-1} \Sigma_{ba} M \Sigma_{ab} \Sigma_{bb}^{-1} \right)^{-1} \left(- \Sigma_{bb} \Sigma_{ba} M \right) \cdots$$

$$\cdots (\mu_a - x_a)$$

$$= \mu_b + \left(\Sigma_{bb}^{-1} + \Sigma_{bb}^{-1} \Sigma_{ba} M \Sigma_{ab} \Sigma_{bb}^{-1} \right)^{-1} \left(\Sigma_{bb} \Sigma_{ba} M \right) (x_a - \mu_a)$$

Por tanto, $p(x_b | x_a) = \mathcal{N}(x_b \mid \mu_{b|a}, \Sigma_{b|a})$

con $\mu_{b|a}$ y $\Sigma_{b|a}$ como se hallaron previamente.