Let $g : \mathbb{R}^k \longrightarrow \mathbb{R}, h : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be continuous functions. Then,

$$\begin{aligned}
Df(x) &= Dg(h(x))Dh(x) = \nabla g(h(x))^T Dh(x) \\
&= \begin{bmatrix} -\nabla g(h(x))^T- \end{bmatrix} \begin{bmatrix} -(\nabla h_1(x))^T- \\ \vdots \\ -(\nabla h_k(x))^T- \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathrm{H}(f)(x) &= [D_i D_j f] = [D_j D_i f] \ for \ i,j = 1,...,n \\
&= \nabla(Df(x)) = D(\nabla f(x)) = D((Df(x))^T) \\
&= D\begin{bmatrix}\begin{bmatrix} | & & | \\ \nabla h_1(x) & \cdots & \nabla h_k(x) \\ | & & | \end{bmatrix}\begin{bmatrix} | \\ \nabla g(h(x)) \\ | \end{bmatrix}\end{bmatrix}
\end{aligned}$$

Let $D_j h(x) = \begin{bmatrix} D_j h_1(x) & \cdots & D_j h_k(x) \end{bmatrix}$. Then,

$$\mathrm{H}(f)(x) = D\begin{bmatrix}\begin{bmatrix} -D_1 h(x)- \\ \vdots \\ -D_n h(x)- \end{bmatrix}\begin{bmatrix} | \\ \nabla g(h(x)) \\ | \end{bmatrix}\end{bmatrix} = D\begin{bmatrix} D_1 h(x)\nabla g(h(x)) \\ \vdots \\ D_n h(x)\nabla g(h(x)) \end{bmatrix}$$

The following holds true for general vector functions, $a(x)$ and $b(x)$.

$$\begin{aligned}
D(a(x))^T b(x) &= D\textstyle\sum_{j=1}^k (a_j(x)b_j(x)) = \sum_{j=1}^k D(a_j(x)b_j(x)) \\
&= \textstyle\sum_{j=1}^k (b_j(x)(\nabla a_j(x))^T + a_j(x)(\nabla b_j(x))^T) \\
&= (b(x))^T \begin{bmatrix} -(\nabla a_1(x))^T- \\ \vdots \\ -(\nabla a_n(x))^T- \end{bmatrix} + (a(x))^T \begin{bmatrix} -(\nabla b_1(x))^T- \\ \vdots \\ -(\nabla b_n(x))^T- \end{bmatrix} \\
&= (b(x))^T Da(x) + (a(x))^T Db(x)
\end{aligned}$$

Let $a(x)^T = D_j h(x)$ and $b(x) = \nabla g(h(x))$. Then,

$$\begin{aligned}
D(D_j h(x)\nabla g(h(x))) &= (\nabla g(h(x)))^T D((D_j h(x))^T) + D_j h(x) D(\nabla g(h(x))) \\
&= Dg(h(x))^T D((D_j h(x))^T) + D_j h(x)\mathrm{H}(g(h(x)))Dh(x)
\end{aligned}$$

Hence,

$$\mathrm{H}(f)(x) = \begin{bmatrix} Dg(h(x))D((D_1 h(x))^T) \\ \vdots \\ Dg(h(x)))D((D_n h(x))^T) \end{bmatrix} + (Dh(x))^T \mathrm{H}(g(h(x)))Dh(x)$$