Here we give the full gory details of the proof that our generating function gives rise to a canonical transformation. As usual let \( x = (q, p) \) and \( y = (Q, P) \). Let \( S = S(q, P) \) be given, where \( P = P(q, p) \). Set

\[
\begin{align*}
P &= \frac{\partial S}{\partial q}, & Q &= \frac{\partial S}{\partial P}.
\end{align*}
\]

Note that \( Q = Q(q, P) \). We will show this gives rise to a canonical transformation. The function \( S \) is known as the generating function. The Jacobian matrix for \( x \to y(x) \) is

\[
\begin{align*}
Dy &= \begin{pmatrix}
\frac{\partial Q_i / \partial q_j}{\partial p_j} |_p & \frac{\partial Q_i / \partial p_j}{\partial q_j} |_p \\
\frac{\partial P_i / \partial q_j}{\partial p_j} |_q & \frac{\partial P_i / \partial p_j}{\partial q_j} |_q
\end{pmatrix}
\end{align*}
\]

where each entry is itself an \( n \times n \) matrix. We have to be very careful which variables are fixed in the partial derivatives. For instance

\[
\frac{\partial Q_i}{\partial q_k} |_p = \frac{\partial Q_i}{\partial q_k} |_p + \frac{\partial Q_i}{\partial P_l} |_q \frac{\partial P_l}{\partial q_k} |_p, \quad \frac{\partial Q_i}{\partial p_k} |_q = \frac{\partial Q_i}{\partial P_l} |_q \frac{\partial P_l}{\partial p_k} |_q.
\]

We want to compute

\[
(Dy) J(Dy)^\dagger = \begin{pmatrix}
\frac{\partial Q_i / \partial q_j}{\partial p_j} |_p & \frac{\partial Q_i / \partial p_j}{\partial q_j} |_p \\
\frac{\partial P_i / \partial q_j}{\partial p_j} |_q & \frac{\partial P_i / \partial p_j}{\partial q_j} |_q
\end{pmatrix} \begin{pmatrix} 0 & I_n \end{pmatrix} \begin{pmatrix} 0 & I_n \end{pmatrix} = \begin{pmatrix} \frac{\partial Q_i / \partial q_j}{\partial p_j} |_p & \frac{\partial P_i / \partial p_j}{\partial q_j} |_q \end{pmatrix}.
\]

Let us examine the first block. It is (assuming summation convention)

\[
\{Q_i, Q_j\}_{q,p} = \frac{\partial Q_i}{\partial q_k} |_p \frac{\partial Q_j}{\partial p_k} |_q - (i \leftrightarrow j)
\]

\[
= \left( \frac{\partial Q_i}{\partial q_k} |_p + \frac{\partial Q_i}{\partial P_l} |_q \frac{\partial P_l}{\partial q_k} |_p \right) \frac{\partial Q_j}{\partial p_k} |_q - (i \leftrightarrow j).
\]

First we note

\[
\frac{\partial Q_i}{\partial q_k} |_p = \frac{\partial^2 S}{\partial q_k \partial P_i} |_q = \frac{\partial p_k}{\partial P_i} |_q,
\]

so that

\[
\frac{\partial Q_i}{\partial q_k} |_p \frac{\partial Q_j}{\partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q \frac{\partial P_m}{\partial p_k} |_q = \frac{\partial p_k}{\partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q = \frac{\partial p_k}{\partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q = \frac{\partial^2 S}{\partial P_i \partial P_j},
\]

which is symmetric in \( i \) and \( j \). So

\[
\{Q_i, Q_j\}_{q,p} = \frac{\partial Q_i}{\partial P_l} |_q \frac{\partial Q_j}{\partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q \frac{\partial P_m}{\partial p_k} |_q - (i \leftrightarrow j)
\]

\[
= \frac{\partial^2 S}{\partial P_l \partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q \frac{\partial^2 S}{\partial P_j \partial P_m} \quad (i \leftrightarrow j)
\]

\[
= \frac{\partial^2 S}{\partial q_k \partial P_i} |_q \frac{\partial P_m}{\partial p_k} |_q \frac{\partial^2 S}{\partial P_j \partial P_m} \quad (i \leftrightarrow j)
\]

\[
= \frac{\partial^2 S}{\partial P_i \partial P_j} \quad (i \leftrightarrow j)
\]

\[
= \frac{\partial^2 S}{\partial P_i \partial P_j} \quad (i \leftrightarrow j)
\]
which vanishes for the same reason. Now look at the block

$$\{P_i, P_j\}_{q,p} = \frac{\partial P_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - (i \leftrightarrow j).$$

Since \( p = \partial_q S(q, P(q, p)) \) holds identically we can differentiate it with respect to \( q \) or \( p \), while holding the other constant. The chain rule gives us

$$0 = \frac{\partial^2 S}{\partial q_i \partial q_j} + \frac{\partial^2 S}{\partial q_i \partial P_k} \frac{\partial P_k}{\partial q_j} + \delta_{ij} \frac{\partial^2 S}{\partial q_i \partial p_k} \frac{\partial P_k}{\partial p_j}.$$ 

So \( C = -B^{-1}A \) and \( D = B^{-1} \) so that \( CD^t = -DAD^t \). Hence

$$\frac{\partial P_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} = C_{ik}D_{jk} = C_{ik}D_{kj} = (CD^t)_{ij} = -(DAD^t)_{ij}.$$

But since \( A \) is symmetric we have \((DAD^t)^t = DA^tD^t = DAD^t\), i.e. it is symmetric! So

$$\{P_i, P_j\}_{q,p} = \frac{\partial P_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - (i \leftrightarrow j) = 0.$$

Now we look at the second block. It is

$$\{Q_i, P_j\}_{q,p} = \frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - \frac{\partial P_i}{\partial q_k} \frac{\partial Q_j}{\partial p_k} = \frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} + \frac{\partial Q_i}{\partial P_k} \frac{\partial P_j}{\partial q_k} - \frac{\partial P_i}{\partial q_k} \frac{\partial Q_j}{\partial P_k} - \frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial q_k}.$$

The first term is just \( \delta_{ij} \) and the term in brackets is exactly \( \{P_i, P_j\}_{q,p} \), which we know vanishes. Since the third block is the same as the second but with a sign change, we conclude

$$DyJ(Dy)^t = \begin{pmatrix} 0 & \delta_{ij} \\ -\delta_{ij} & 0 \end{pmatrix} = J.$$

So the derivative of the map is symplectic and the transformation is necessarily canonical.