If we write the LR statistic (13.08) in partitioned form, we obtain

$$\begin{aligned} LR &= (\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}})^\top \boldsymbol{I} (\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) \\ &= \begin{bmatrix} \tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1 \\ \tilde{\boldsymbol{\theta}}_2 - \hat{\boldsymbol{\theta}}_2 \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{I}_{11} & \boldsymbol{I}_{12} \\ \boldsymbol{I}_{21} & \boldsymbol{I}_{22} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1 \\ \tilde{\boldsymbol{\theta}}_2 - \hat{\boldsymbol{\theta}}_2 \end{bmatrix} \\ &= (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1)^\top \boldsymbol{I}_{11} (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1) - 2(\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1)^\top \boldsymbol{I}_{12} \hat{\boldsymbol{\theta}}_2 + \hat{\boldsymbol{\theta}}_2^\top \boldsymbol{I}_{22} \hat{\boldsymbol{\theta}}_2. \end{aligned}$$

where the last line uses the fact that $\tilde{\boldsymbol{\theta}}_2 = \mathbf{0}$. Making use of the result (13.11), the LR statistic can then be rewritten as

$$\begin{aligned} LR &= (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1)^\top \boldsymbol{I}_{11} (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1) - 2(\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1)^\top \boldsymbol{I}_{11} (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1) + \hat{\boldsymbol{\theta}}_2^\top \boldsymbol{I}_{22} \hat{\boldsymbol{\theta}}_2 \\ &= \hat{\boldsymbol{\theta}}_2^\top \boldsymbol{I}_{22} \hat{\boldsymbol{\theta}}_2 - (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1)^\top \boldsymbol{I}_{11} (\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1). \end{aligned} \tag{13.12}$$

We now show that the Wald statistic is equal to (13.12). Since the restrictions take the form (13.09), we see that $\boldsymbol{r}(\boldsymbol{\theta}) = \boldsymbol{\theta}_2$ and $\hat{\boldsymbol{r}} = \hat{\boldsymbol{\theta}}_2$. This implies that the matrix $\boldsymbol{R}$ can be written as

$$\boldsymbol{R}(\boldsymbol{\theta}) = [\mathbf{0} \quad \mathbf{I}],$$

where the $\mathbf{0}$ matrix is $r \times (k - r)$, and the identity matrix $\mathbf{I}$ is $r \times r$. Then the expression $\hat{\boldsymbol{R}} \boldsymbol{I}^{-1} \hat{\boldsymbol{R}}^\top$ that appears in the Wald statistic (13.05) is just the $(2, 2)$ block of the inverse matrix $\boldsymbol{I}^{-1}$. By the results in Appendix A on partitioned matrices, we obtain

$$\big(\hat{\boldsymbol{R}} \boldsymbol{I}^{-1} \hat{\boldsymbol{R}}^\top\big)^{-1} = \big((\boldsymbol{I}^{-1})_{22}\big)^{-1} = \boldsymbol{I}_{22} - \boldsymbol{I}_{21} \boldsymbol{I}_{11}^{-1} \boldsymbol{I}_{12}. \tag{13.13}$$

This result allows us to put (13.05) in the form

$$W = \hat{\boldsymbol{\theta}}_2^\top \big(\boldsymbol{I}_{22} - \boldsymbol{I}_{21} \boldsymbol{I}_{11}^{-1} \boldsymbol{I}_{12}\big) \hat{\boldsymbol{\theta}}_2.$$

By (13.11), this last expression is equal to

$$\hat{\boldsymbol{\theta}}_2^\top \boldsymbol{I}_{22} \hat{\boldsymbol{\theta}}_2 - \big(\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1\big)^\top \boldsymbol{I}_{11} \big(\tilde{\boldsymbol{\theta}}_1 - \hat{\boldsymbol{\theta}}_1\big),$$

which is the same as (13.12). The proof of the equality of the three classical statistics for the quadratic loglikelihood function (13.07) is therefore complete.

It is of interest to see how the three classical test statistics are related geometrically. Figure 13.1 depicts the graph of a loglikelihood function $\ell(\boldsymbol{y}, \theta_1, \theta_2)$. It is drawn for a *given* sample vector $\boldsymbol{y}$ and consequently a given sample size n. For simplicity, the parameter space has been supposed to be two-dimensional. There is only one restriction, which is that the second component of the parameter vector, θ_2, is equal to zero. Therefore, the function ℓ can be treated as a function of the two variables θ_1 and θ_2 only, and its