and then $u_{k+1} = u_k - \rho_k d_k$. In fact, by $(*_2)$, since

$$\Delta_k = \sum_{i=0}^k \delta_i^k \nabla J_{u_i} = \delta_k^k \left(\sum_{i=0}^{k-1} \frac{\delta_i^k}{\delta_k^k} \nabla J_{u_i} + \nabla J_{u_k} \right),$$

we must have

$$\Delta_k = \delta_k^k d_k \quad \text{and} \quad \rho_k = -\delta_k^k.$$
 (*4)

Remarkably, the coefficients λ_i^k and the descent directions d_k can be computed easily using the following formulae.

Proposition 49.17. Assume that $\nabla J_{u_i} \neq 0$ for i = 0, ..., k. If we write

$$d_{\ell} = \sum_{i=0}^{\ell-1} \lambda_i^{\ell} \nabla J_{u_i} + \nabla J_{u_{\ell}}, \quad 0 \le \ell \le k,$$

then we have

$$\begin{cases} \lambda_i^k = \frac{\|\nabla J_{u_k}\|^2}{\|\nabla J_{u_i}\|^2}, & 0 \le i \le k - 1, \\ d_0 = \nabla J_{u_0} \\ d_\ell = \nabla J_{u_\ell} + \frac{\|\nabla J_{u_\ell}\|^2}{\|\nabla J_{u_{\ell-1}}\|^2} d_{\ell-1}, & 1 \le \ell \le k. \end{cases}$$

Proof. Since by $(*_4)$ we have $\Delta_k = \delta_k^k d_k$, $\delta_k^k \neq 0$, (by Proposition 49.16) we have

$$\langle Ad_{\ell}, \Delta_i \rangle = 0, \quad 0 \le i < \ell \le k.$$

By $(*_1)$ we have $\nabla J_{u_{\ell+1}} = \nabla J_{u_{\ell}} + A\Delta_{\ell}$, and since A is a symmetric matrix, we have

$$0 = \langle Ad_k, \Delta_\ell \rangle = \langle d_k, A\Delta_\ell \rangle = \langle d_k, \nabla J_{u_{\ell+1}} - \nabla J_{u_\ell} \rangle,$$

for $\ell = 0, \dots, k-1$. Since

$$d_k = \sum_{i=0}^{k-1} \lambda_i^k \nabla J_{u_i} + \nabla J_{u_k},$$

we have

$$\left\langle \sum_{i=0}^{k-1} \lambda_i^k \nabla J_{u_i} + \nabla J_{u_k}, \nabla J_{u_{\ell+1}} - \nabla J_{u_\ell} \right\rangle = 0, \quad 0 \le \ell \le k-1.$$

Since by Proposition 49.15 the gradients ∇J_{u_i} are pairwise orthogonal, the above equations yield

$$-\lambda_{k-1}^{k} \|\nabla J_{u_{k-1}}\|^{2} + \|\nabla J_{u_{k}}\|^{2} = 0$$

$$-\lambda_{\ell}^{k} \|\nabla J_{u_{\ell}}\|^{2} + \lambda_{\ell+1}^{k} \|\nabla J_{u_{\ell+1}}\|^{2} = 0, \quad 0 \le \ell \le k-2, \quad k \ge 2,$$