where e_i denotes the ith canonical basis vector in $\mathbb{R}^n$. If J is differentiable, necessary conditions for a minimum, which are also sufficient if J is convex, is that the directional derivatives $dJ_v(e_i)$ be all zero, that is,

$$\langle \nabla J_v, e_i \rangle = 0 \quad i = 0, \ldots, n.$$

The following result regarding the convergence of the mehod of relaxation is proven in Ciarlet [41] (Chapter 8, Theorem 8.4.2).

Proposition 49.12. *If the functional $J\colon \mathbb{R}^n \to \mathbb{R}$ is elliptic, then the relaxation method converges.*

Remarks: The proof of Proposition 49.12 uses Theorem 49.8. The finite dimensionality of $\mathbb{R}^n$ also plays a crucial role. The differentiability of the function J is also crucial. Examples where the method loops forever if J is not differentiable can be given; see Ciarlet [41] (Chapter 8, Section 8.4). The proof of Proposition 49.12 yields an *a priori* bound on the error $\|u - u_k\|$. If J is a quadratic functional

$$J(v) = \frac{1}{2} v^\top A v - b^\top v,$$

where A is a symmetric positive definite matrix, then $\nabla J_v = Av - b$, so the above method for solving for u_{k+1} in terms of u_k becomes the *Gauss–Seidel method* for solving a linear system; see Section 10.3.

We now discuss gradient methods.

49.6 Gradient Descent Methods for Unconstrained Problems

The intuition behind these methods is that the convergence of an iterative method ought to be better if the difference $J(u_k) - J(u_{k+1})$ is as large as possible during every iteration step. To achieve this, it is natural to pick the descent direction to be the one *in the opposite direction of the gradient vector* ∇J_{u_k}. This choice is justified by the fact that we can write

$$J(u_k + w) = J(u_k) + \langle \nabla J_{u_k}, w \rangle + \epsilon(w) \|w\|, \quad \text{with } \lim_{w \mapsto 0} \epsilon(w) = 0.$$

If $\nabla J_{u_k} \neq 0$, the first-order part of the variation of the function J is bounded in absolute value by $\|\nabla J_{u_k}\| \|w\|$ (by the Cauchy–Schwarz inequality), with equality if ∇J_{u_k} and w are collinear.

Gradient descent methods pick the direction of descent to be $d_k = -\nabla J_{u_k}$, so that we have

$$u_{k+1} = u_k - \rho_k \nabla J_{u_k},$$