Summer 2020
Prof. Thomas Slawig
Dr. Joscha Reimer

Kiel University
Department of Computer Science

Optimization and Data Science
Exam exercise 2

We have proved the following theorem: Let

- f be differentiable and bounded from below,
- the sequence of directions $(d_k)_{k\in\mathbb{N}}$ be gradient-related,
- the sequence of step-sizes $(\rho_k)_{k\in\mathbb{N}}$ be efficient.

Then the descent method generates a sequence $(x_k)_{k\in\mathbb{N}}$ that satisfies

- either $\nabla f(x_k) = 0$ for some $k \in \mathbb{N}$
- or $\lim_{k\to\infty} \nabla f(x_k) = 0$.

Why do the following steps in the proof hold?

$$f(x_\ell) - f(x_0) \leq -c_S \sum_{k=0}^{\ell-1} \left(\frac{\nabla f(x_k)^\top d_k}{\|d_k\|} \right)^2, \quad \ell \in \mathbb{N}, \tag{1}$$

$$\sum_{k=0}^{\infty} \left(\frac{\nabla f(x_k)^\top d_k}{\|d_k\|} \right)^2 < \infty, \tag{2}$$

$$\lim_{k\to\infty} \frac{\nabla f(x_k)^\top d_k}{\|d_k\|} = 0, \tag{3}$$

$$\lim_{k\to\infty} \nabla f(x_k) = 0. \tag{4}$$