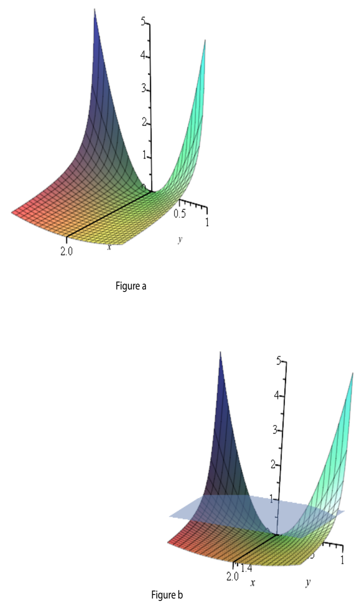

Figure 51.7: Figure (a) illustrates the proper convex function of Example 51.4. Figure (b) illustrates the approach to $(0,0)$ along the planar parabolic curve $(y^2/2, y)$. Then $f(y^2/2, y) = 1$ and Figure b shows the intersection of the surface with the plane $z = 1$.

51.2 Subgradients and Subdifferentials

We saw in the previous section that proper convex functions have "good" continuity properties. Remarkably, if f is a convex function, for any $x \in \mathbb{R}^n$ such that $f(x)$ is finite, the one-sided derivative $f'(x; u)$ exists for all $u \in \mathbb{R}^n$; This result has been shown at least since 1893, as noted by Stoltz (see Rockafellar [138], page 428). Directional derivatives will be discussed in Section 51.3. If f is differentitable at x, then of course

$$df_x(u) = \langle \nabla f_x, u \rangle \quad \text{for all } u \in \mathbb{R}^n,$$

where ∇f_x is the gradient of f at x.

But even if f is not differentiable at x, it turns out that for "most" $x \in \mathrm{dom}(f)$, in particular if $x \in \mathbf{relint}(\mathrm{dom}(f))$, there is a nonempty closed convex set $\partial f(x)$ which may be viewed as a generalization of the gradient ∇f_x. This convex set of $\mathbb{R}^n$, $\partial f(x)$, called the *subdifferential of f at x*, has some of the properties of the gradient ∇f_x. The vectors in $\partial f(x)$ are called *subgradients* at x. For example, if f is a proper convex function, then f achieves its minimum at $x \in \mathbb{R}^n$ iff $0 \in \partial f(x)$. Some of the theorems of Chapter 50 can be generalized to convex functions that are not necessarily differentiable by replacing conditions involving gradients by conditions involving subdifferentials. These generalizations are crucial for the justification that various iterative methods for solving optimization programs converge. For