Test functions for optimization needs

Marcin Molga, Czesław Smutnicki

3 kwietnia 2005

Streszczenie

This paper provides the review of literature benchmarks (test functions) commonly used in order to test optimization procedures dedicated for multidimensional, continuous optimization task. Special attention has been paid to multiple-extreme functions, treated as the quality test for "resistant" optimization methods (GA, SA, TS, etc.)

1 Introduction

Quality of optimization procedures (those already known and these newly proposed) are frequently evaluated by using common standard literature benchmarks. There are several classes of such test functions, all of them are continuous:

(a) unimodal, convex, multidimensional,

(b) multimodal, two-dimensional with a small number of local extremes,

(c) multimodal, two-dimensional with huge number of local extremes

(d) multimodal, multidimensional, with huge number of local extremes, .

Class (a) contains nice functions as well as malicious cases causing poor or slow convergence to single global extremum. Class (b) is mediate between (a) and (c)-(d), and is used to test quality of standard optimization procedures in the hostile environment, namely that having few local extremes with single global one. Classes (c)-(d) are recommended to test quality of intelligent "resistant" optimization methods, as an example GA, SA, TS, etc. These classes are considered as very hard test problems. Class (c) is "artificial" in some sense, since the behavior of optimization procedure is usually being justified, explain and supported by human intuitions on 2D surface. Moreover, two-dimensional optimization problems appear very rarely in practice. Unfortunately, practical discrete optimization problems provide instances with large number of dimensions, laying undoubtedly in class (d). For example, the smallest known currently benchmark ft10 for so called *job shop scheduling problem* has dimension 90, the biggest known - has dimension 1980. Therefore, in order to test real quality of proposed algorithms, we need to consider chiefly instances from class (d). As the shocking contrast, the proposed GA approaches for continuous optimization do not exceed dimension 10.

Notice, polarization (a constant added to function value) has no influence on the result of minimization. Therefore, definitions of functions can differ from these original ones by a constant. All tests are formulated hereinafter as *minimization* problems, nevertheless can be applied also for maximization problems by simple inverting sign of the function.

2 Test functions

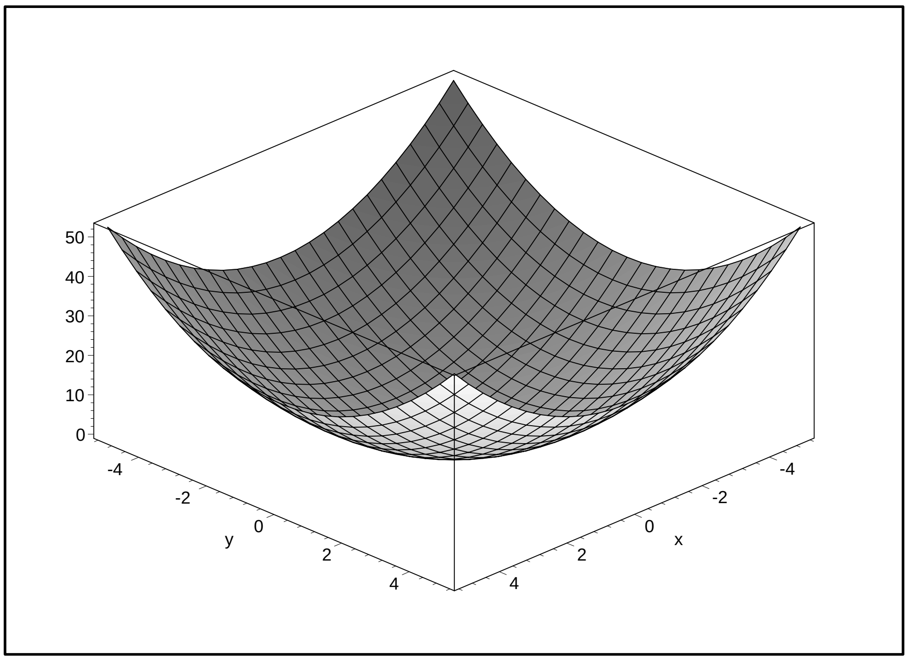
In this section we present benchmarks commonly known in the literature.

2.1 De Jong's function

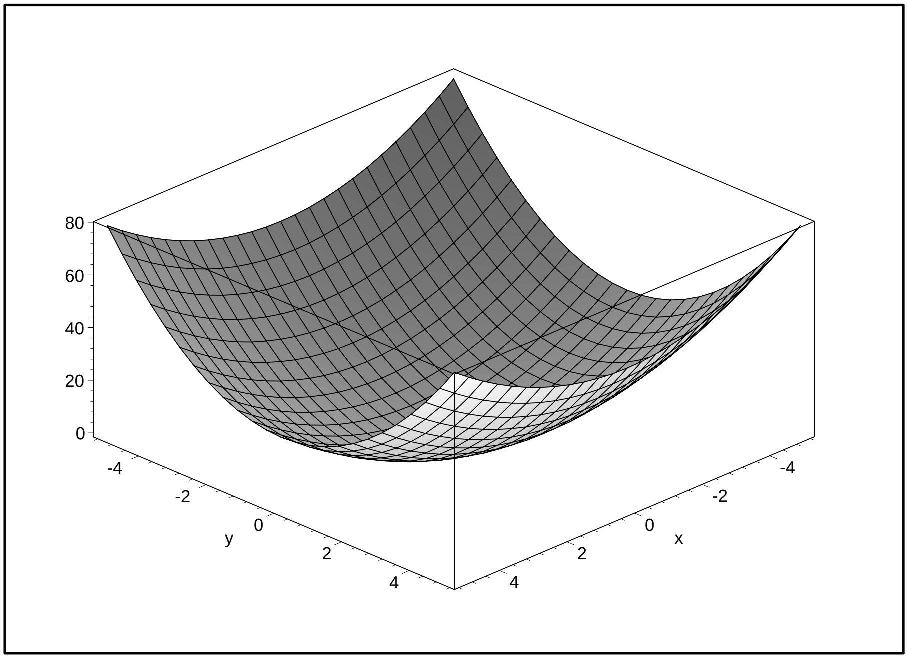
So called *first function of De Jong's* is one of the simplest test benchmark. Function is continuous, convex and unimodal. It has the following general definition

$$f(x) = \sum_{i=1}^{n} x_i^2. \tag{1}$$

Test area is usually restricted to hyphercube $-5.12 \leq x_i \leq 5.12$, $i = 1, \ldots, n$. Global minimum $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$.



Rysunek 1: De Jong's function in 2D, $f(x, y) = x^2 + y^2$



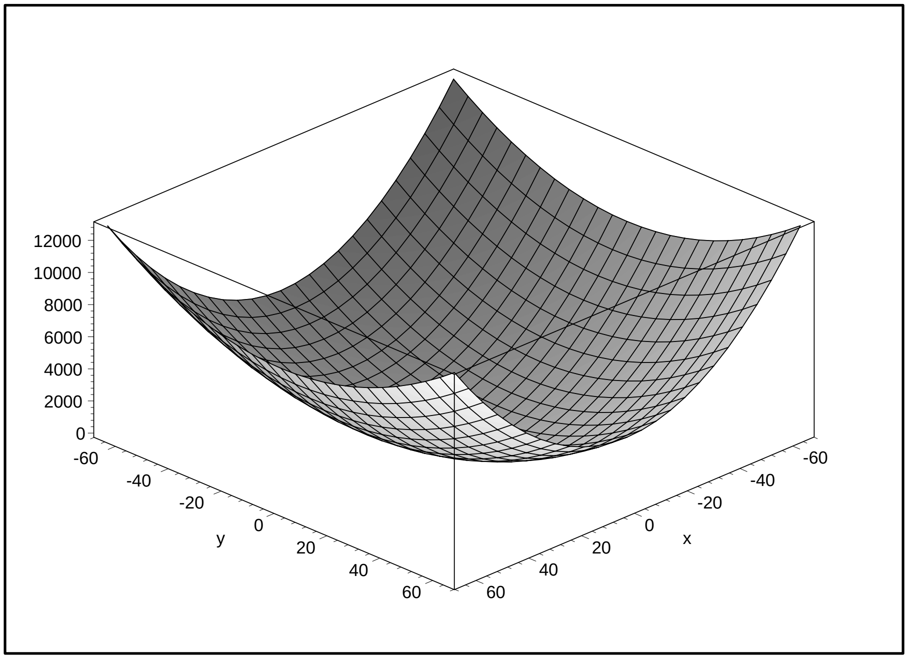
Rysunek 2: Axis parallel hyper-ellipsoid function in 2D, $f(x, y) = x^2 + 2y^2$

2.2 Axis parallel hyper-ellipsoid function

The axis parallel hyper-ellipsoid is similar to function of De Jong. It is also known as the *weighted sphere model*. Function is continuous, convex and unimodal. It has the following general definition

$$f(x) = \sum_{i=1}^{n} (i \cdot x_i^2). \tag{2}$$

Test area is usually restricted to hyphercube $-5.12 \leq x_i \leq 5.12$, $i = 1, \dots, n$. Global minimum equal $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \dots, n$.



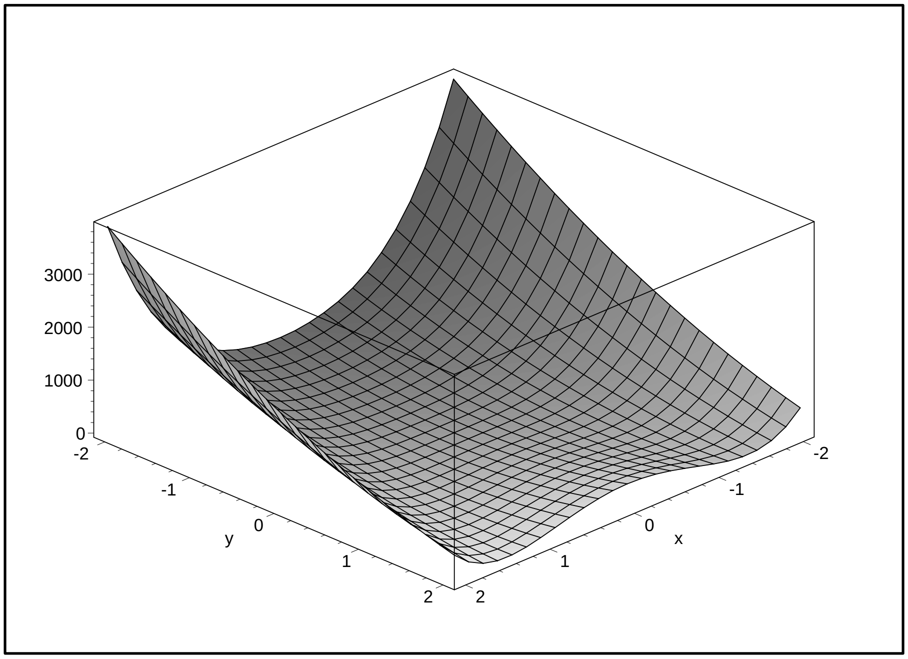
Rysunek 3: Rotated hyper-ellipsoid function in 2D, $f(x, y) = x^2 + (x^2 + y^2)$

2.3 Rotated hyper-ellipsoid function

An extension of the axis parallel hyper-ellipsoid is Schwefel's function. With respect to the coordinate axes, this function produces rotated hyper-ellipsoids. It is continuous, convex and unimodal. Function has the following general definition

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{i} x_j^2. \tag{3}$$

Test area is usually restricted to hyphercube $-65.536 \leq x_i \leq 65.536$, $i = 1, \ldots, n$. Its global minimum equal $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$.



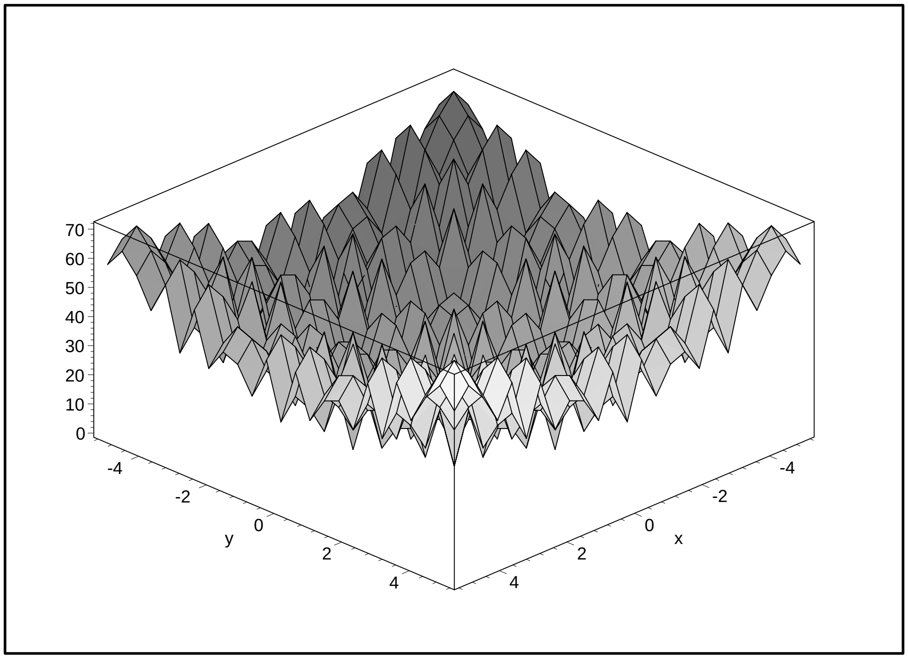
Rysunek 4: Rosenbrock's valley in 2D, $f(x, y) = 100(y - x^2)^2 + (1 - x)^2$

2.4 Rosenbrock's valley

Rosenbrock's valley is a classic optimization problem, also known as *banana function* or the *second function of De Jong*. The global optimum lays inside a long, narrow, parabolic shaped flat valley. To find the valley is trivial, however convergence to the global optimum is difficult and hence this problem has been frequently used to test the performance of optimization algorithms. Function has the following definition

$$f(x) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right]. \tag{4}$$

Test area is usually restricted to hyphercube $-2.048 \le x_i \le 2.048$, $i = 1, \ldots, n$. Its global minimum equal $f(x) = 0$ is obtainable for x_i, $i = 1, \ldots, n$.



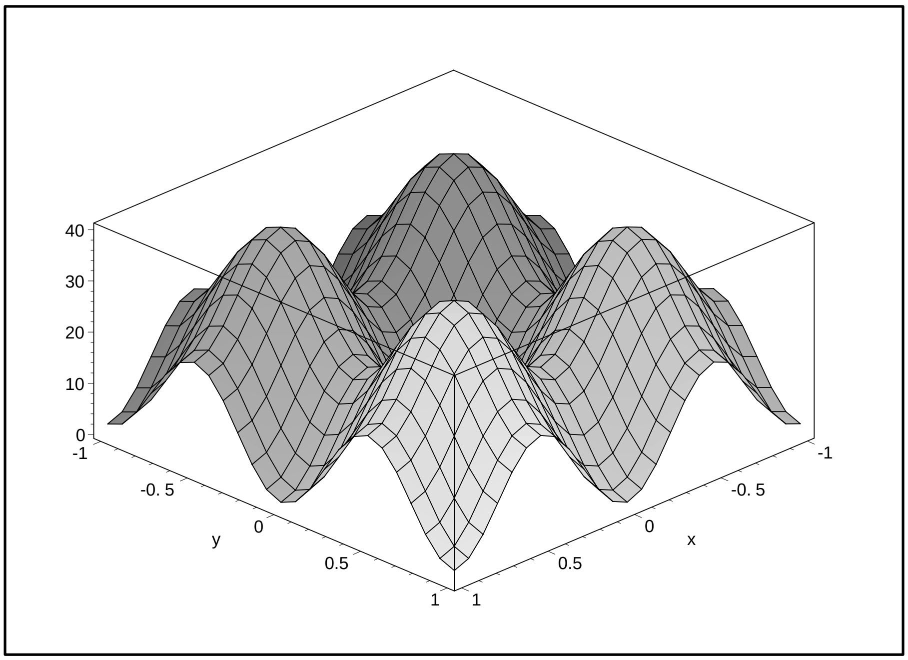
Rysunek 5: An overview of Rastrigin's function in 2D, $f(x, y) = 10 \cdot 2 + [x^2 - 10\cos(2\pi x)] + [y^2 - 10\cos(2\pi y)]$

2.5 Rastrigin's function

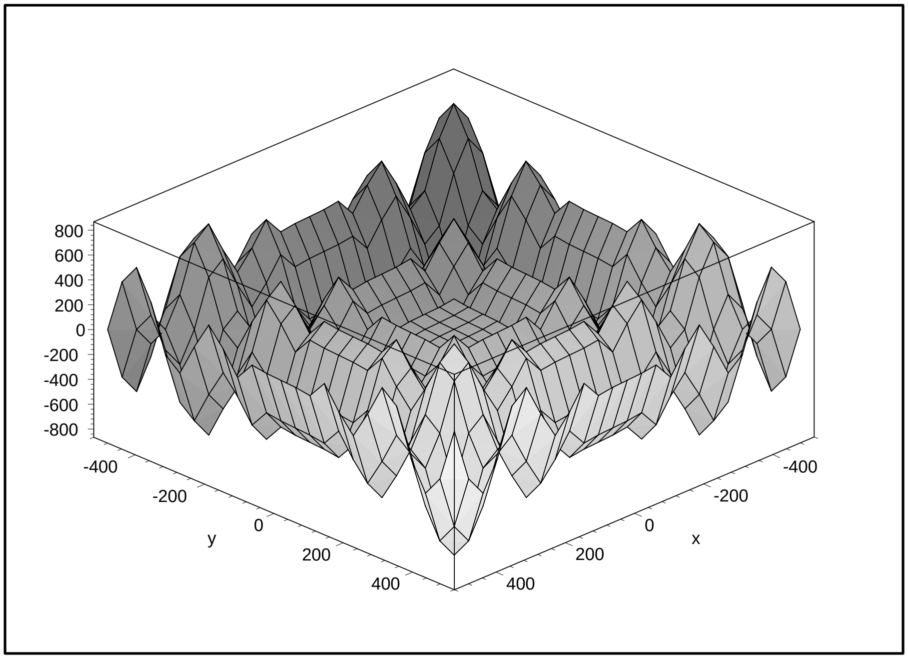
Rastrigin's function is based on the function of De Jong with the addition of cosine modulation in order to produce frequent local minima. Thus, the test function is highly multimodal. However, the location of the minima are regularly distributed. Function has the following definition

$$f(x) = 10n + \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i) \right]. \tag{5}$$

Test area is usually restricted to hyphercube $-5.12 \leq x_i \leq 5.12$, $i = 1, \ldots, n$. Its global minimum equal $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$.



Rysunek 6: Zoom on Rastrigin's function in 2D, $f(x, y) = 10 \cdot 2 + [x^2 - 10\cos(2\pi x)] + [y^2 - 10\cos(2\pi y)]$



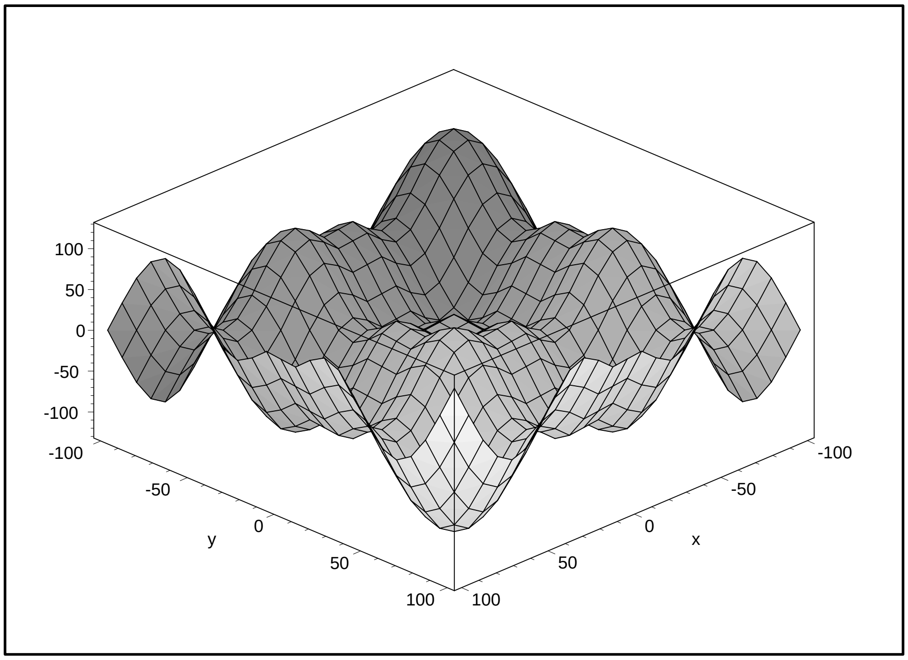
Rysunek 7: An overview of Schwefel's function in 2D, $f(x, y) = -x \sin(\sqrt{|x|} - y \sin(\sqrt{|y|}$

2.6 Schwefel's function

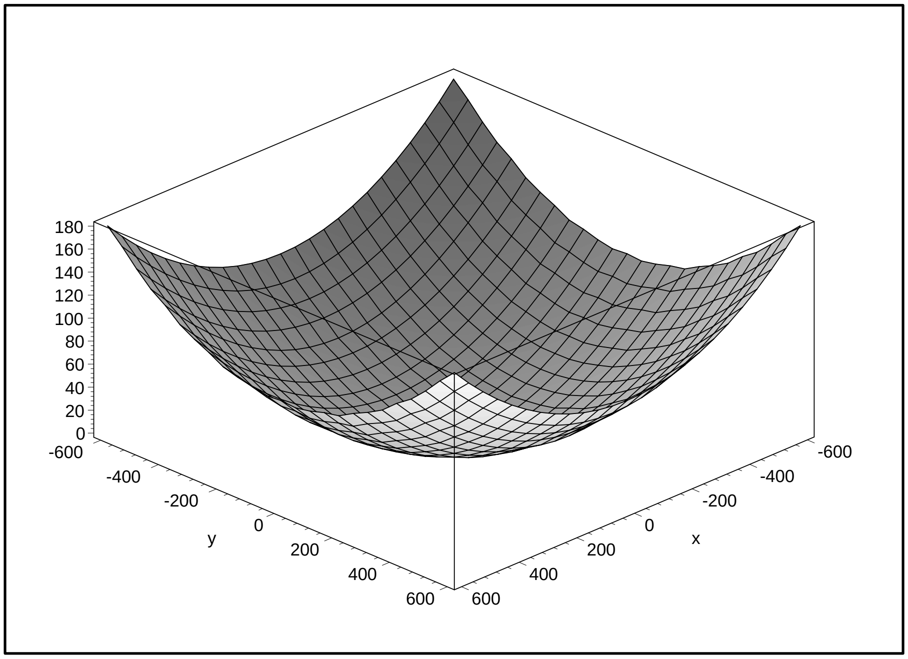
Schwefel's function is deceptive in that the global minimum is geometrically distant, over the parameter space, from the next best local minima. Therefore, the search algorithms are potentially prone to convergence in the wrong direction. Function has the following definition

$$f(x) = \sum_{i=1}^{n} \left[-x_i \sin(\sqrt{|x_i|}) \right]. \tag{6}$$

Test area is usually restricted to hyphercube $-500 \leq x_i \leq 500$, $i = 1, \ldots, n$. Its global minimum $f(x) = -418.9829n$ is obtainable for $x_i = 420.9687$, $i = 1, \ldots, n$.



Rysunek 8: Zoom on Schwefel's function in 2D, $f(x,y) = -x\sin(\sqrt{|x|} - y\sin(\sqrt{|y|}$



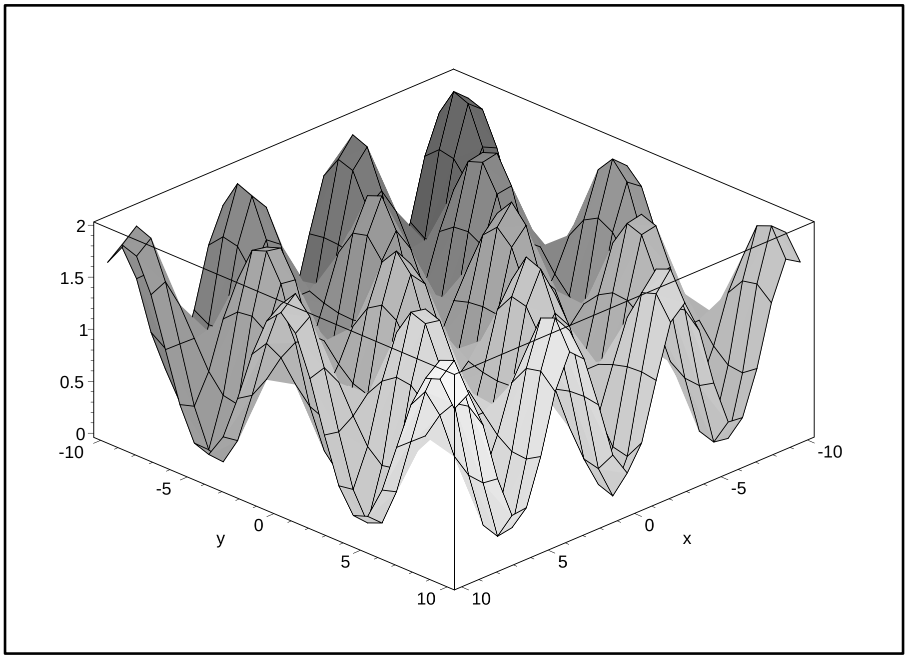
Rysunek 9: An overview of Griewangk's function in 2D, $f(x,y) = \frac{x^2+y^2}{4000} - \cos(x)\cos(\frac{y}{\sqrt{2}}) + 1$

2.7 Griewangk's function

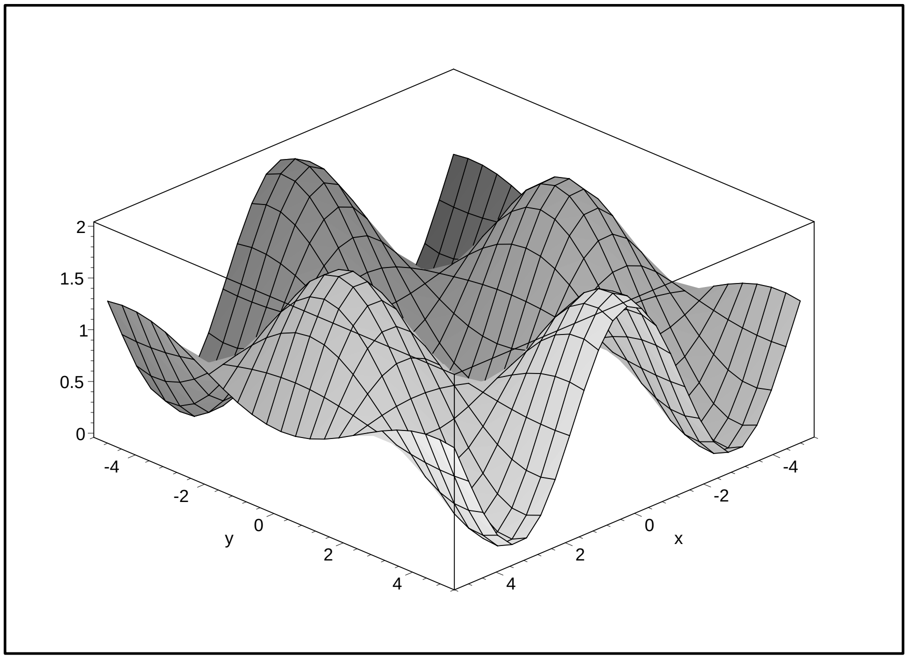
Griewangk's function is similar to the function of Rastrigin. It has many widespread local minima regularly distributed. Function has the following definition

$$f(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1. \quad (7)$$

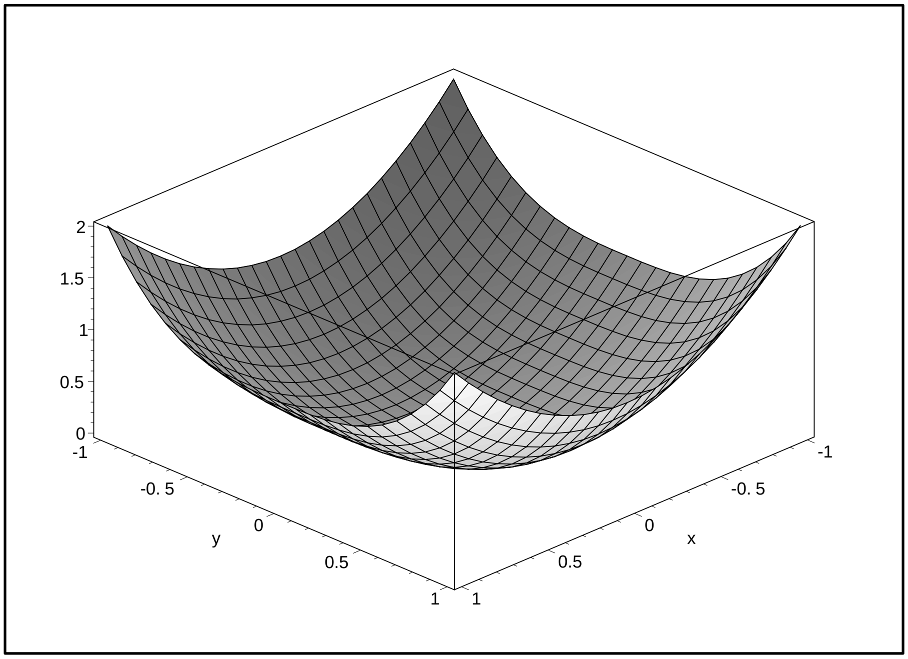
Test area is usually restricted to hyphercube $-600 \leq x_i \leq 600$, $i = 1, \ldots, n$. Its global minimum equal $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$. The function interpretation changes with the scale; the general overview suggests convex function, medium-scale view suggests existence of local extremum, and finally zoom on the details indicates complex structure of numerous local extremum.



Rysunek 10: Medium-scale view of Griewangk's function in 2D, $f(x, y) = \frac{x^2+y^2}{4000} - \cos(x)\cos(\frac{y}{\sqrt{2}}) + 1$



Rysunek 11: Zoom on Griewangk's function in 2D, $f(x, y) = \frac{x^2+y^2}{4000} - \cos(x)\cos(\frac{y}{\sqrt{2}}) + 1$



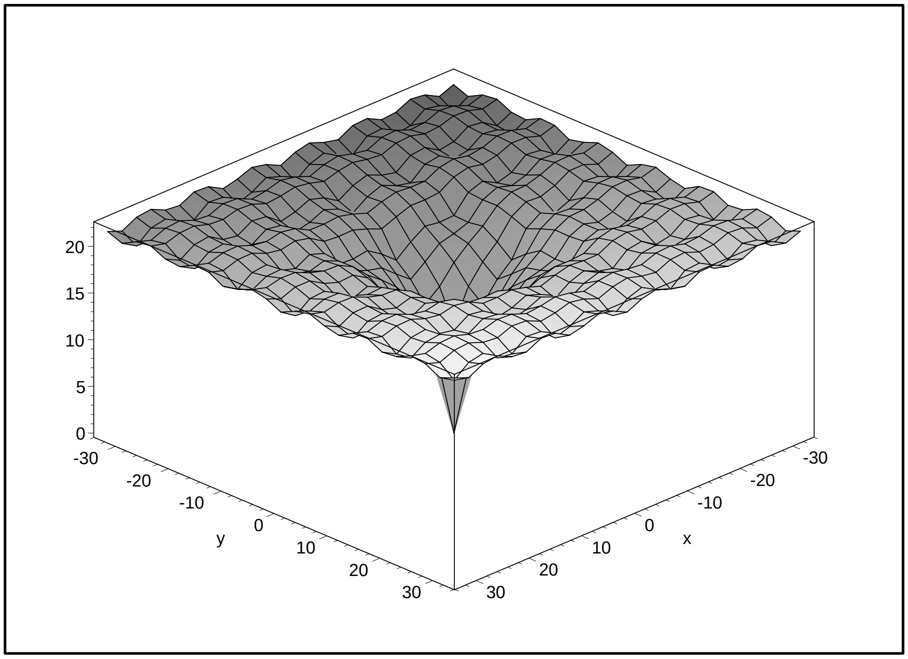
Rysunek 12: Sum of different power functions in 2D, $f(x,y) = |x|^2 + |y|^3$

2.8 Sum of different power functions

The sum of different powers is a commonly used unimodal test function. It has the following definition

$$f(x) = \sum_{i=1}^{n} |x_i|^{i+1} . \quad (8)$$

Test area is usually restricted to hyphercube $-1 \le x_i \le 1$, $i = 1, \ldots, n$. Its global minimum equal $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$.



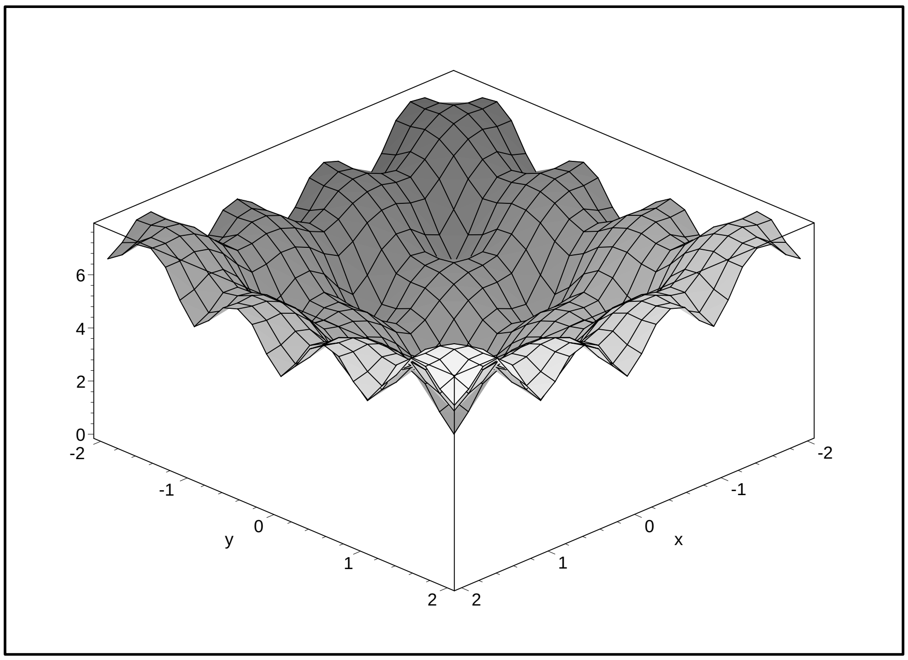
Rysunek 13: An overview of Ackley's function in 2D, $f(x, y) = -x\sin(\sqrt{|x|} - y\sin(\sqrt{|y|}$

2.9 Ackley's function

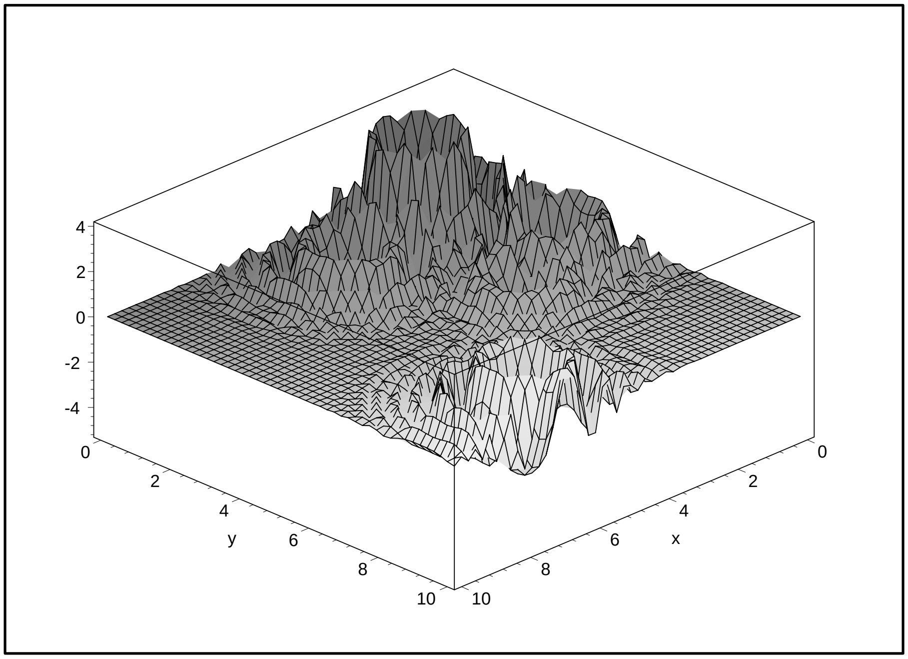
Ackley's is a widely used multimodal test function. It has the following definition

$$f(x) = -a \cdot \exp(-b \cdot \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(cx_i)) + a + \exp(1) \quad (9)$$

It is recommended to set $a = 20$, $b = 0.2$, $c = 2\pi$. Test area is usually restricted to hyphercube $-32.768 \leq x_i \leq 32.768$, $i = 1, \ldots, n$. Its global minimum $f(x) = 0$ is obtainable for $x_i = 0$, $i = 1, \ldots, n$.



Rysunek 14: Zoom on Ackley's function in 2D, $f(x,y) = -x\sin(\sqrt{|x|} - y\sin(\sqrt{|y|}$



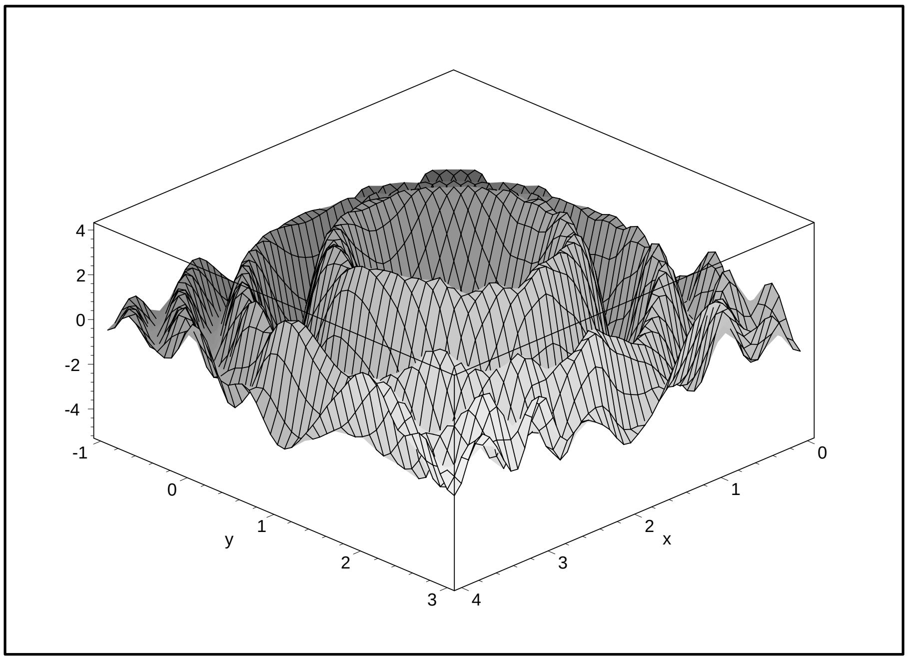
Rysunek 15: An overview of Langermann's function in 2D. $f(x,y) = \sum_{i=1}^{m} c_i \exp(-(x-a_j)^2/\pi - (y-b_j)^2/\pi)\cos(\pi(x-a_j)^2 + \pi(y-b_j)^2)$, $m = 5$, $a = [3,5,2,1,7]$, $b = [5,2,1,4,9]$, $c = [1,2,5,2,3]$

2.10 Langermann's function

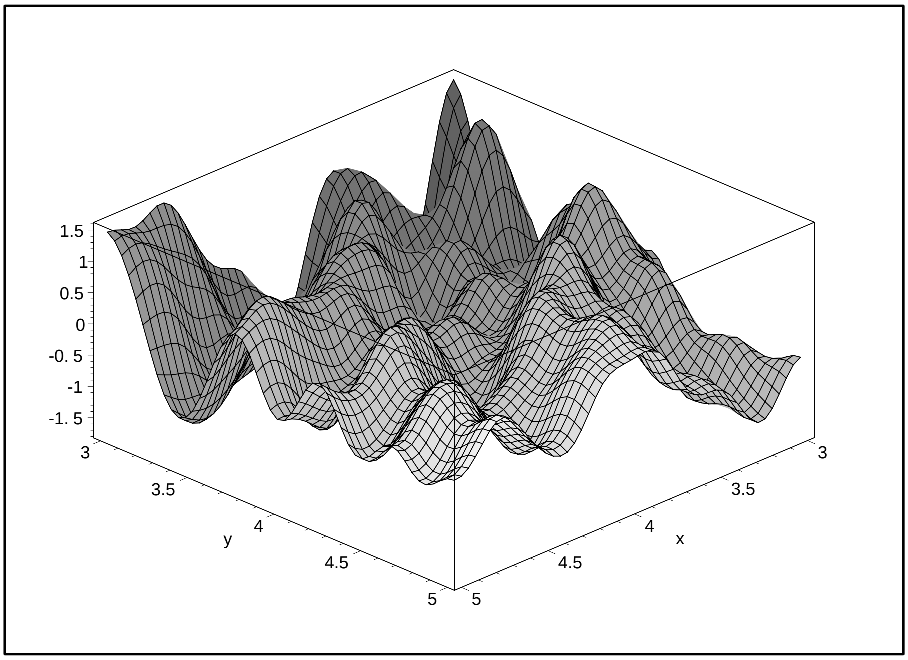
The Langermann function is a multimodal test function. The local minima are unevenly distributed. Function has the following definition

$$f(x) = \sum_{i=1}^{m} c_i \exp[-\frac{1}{\pi}\sum_{j=1}^{n}(x_j - a_{ij})^2]\cos[\pi\sum_{j=1}^{n}(x_j - a_{ij})^2] \qquad (10)$$

where (c_i, $i = 1, \ldots, m$), (a_{ij}, $j = 1, \ldots, n$, $i = 1, \ldots, m$) are constant numbers fixed in advance. It is recommended to set $m = 5$.



Rysunek 16: Medium-scale view on Langermann's function in 2D. $f(x, y) = \sum_{i=1}^{m} c_i \exp(-(x - a_j)^2/\pi - (y - b_j)^2/\pi) \cos(\pi(x - a_j)^2 + \pi(y - b_j)^2)$, $m = 5$, $a = [3, 5, 2, 1, 7]$, $b = [5, 2, 1, 4, 9]$, $c = [1, 2, 5, 2, 3]$



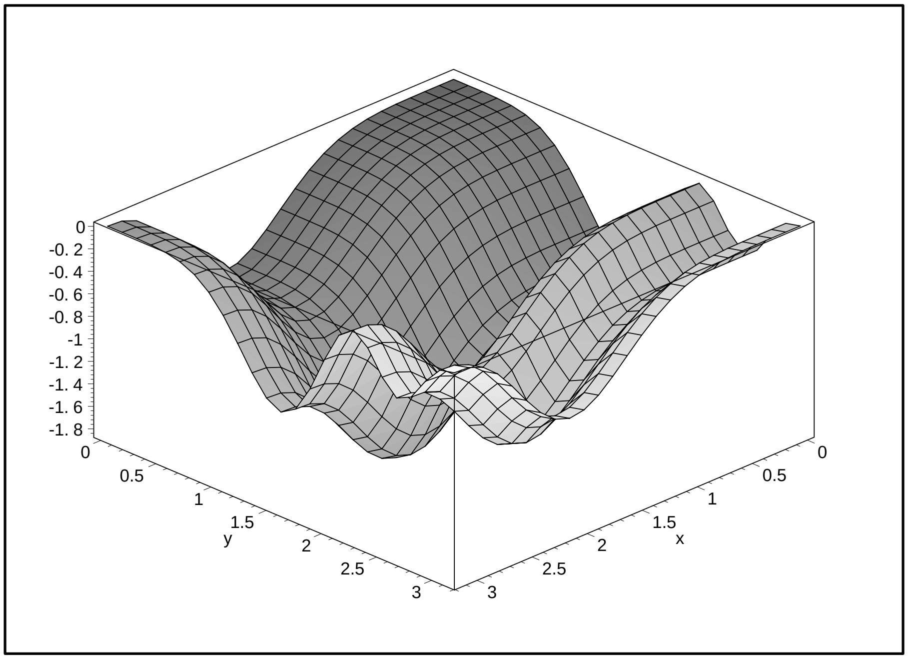
Rysunek 17: Zoom on Langermann's function in 2D. $f(x,y) = \sum_{i=1}^{m} c_i \exp(-(x-a_j)^2/\pi - (y-b_j)^2/\pi)\cos(\pi(x-a_j)^2 + \pi(y-b_j)^2)$, $m = 5$, $a = [3,5,2,1,7]$, $b = [5,2,1,4,9]$, $c = [1,2,5,2,3]$

2.11 Michalewicz's function

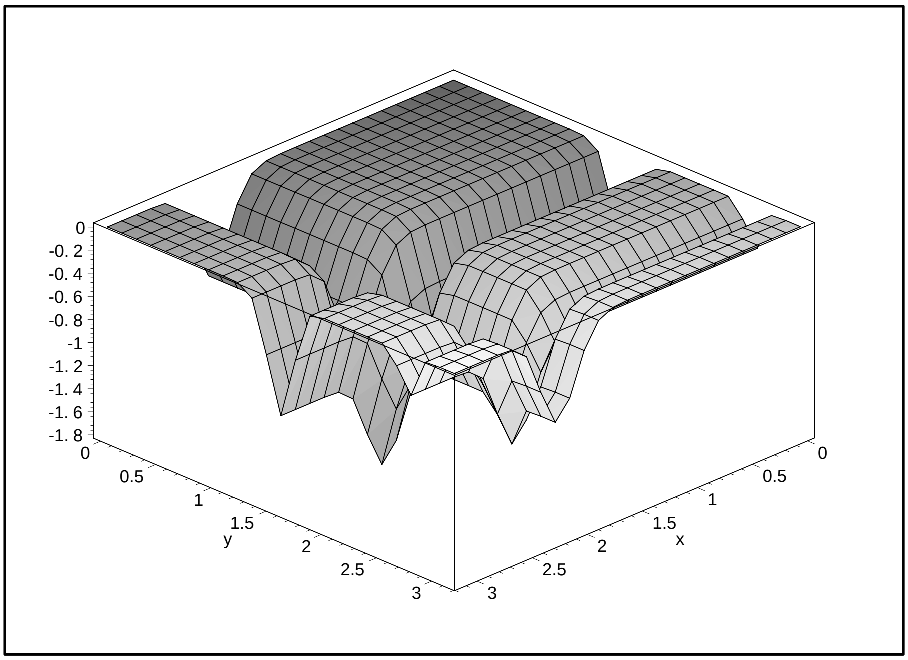
The Michalewicz function is a multimodal test function (owns $n!$ local optima). The parameter m defines the "steepness" of the valleys or edges. Larger m leads to more difficult search. For very large m the function behaves like a needle in the haystack (the function values for points in the space outside the narrow peaks give very little information on the location of the global optimum). Function has the following definition

$$f(x) = -\sum_{i=1}^{n} \sin(x_i) \left[\sin(\frac{ix_i^2}{\pi}) \right]^{2m} \tag{11}$$

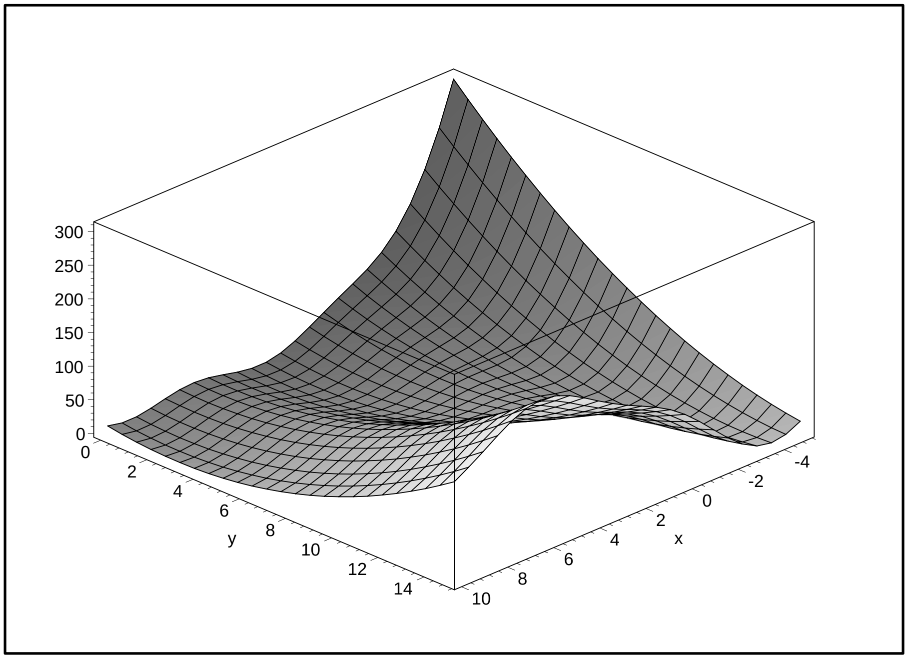
It is usually set $m = 10$. Test area is usually restricted to hyphercube $0 \leq x_i \leq \pi$, $i = 1, \ldots, n$. The global minimum value has been approximated by $f(x) = -4.687$ for $n = 5$ and by $f(x) = -9.66$ for $n = 10$. Respective optimal solutions are not given.



Rysunek 18: Michalewicz's function in 2D for $m = 1$, $f(x, y) = -\sin(x)(\sin(\frac{x^2}{\pi})^{2m} - \sin(y)(\sin(\frac{y^2}{\pi})^{2m}$



Rysunek 19: Michalewicz's function in 2D for $m = 10$, $f(x,y) = -\sin(x)(\sin(\frac{x^2}{\pi})^{2m} - \sin(y)(\sin(\frac{y^2}{\pi})^{2m}$



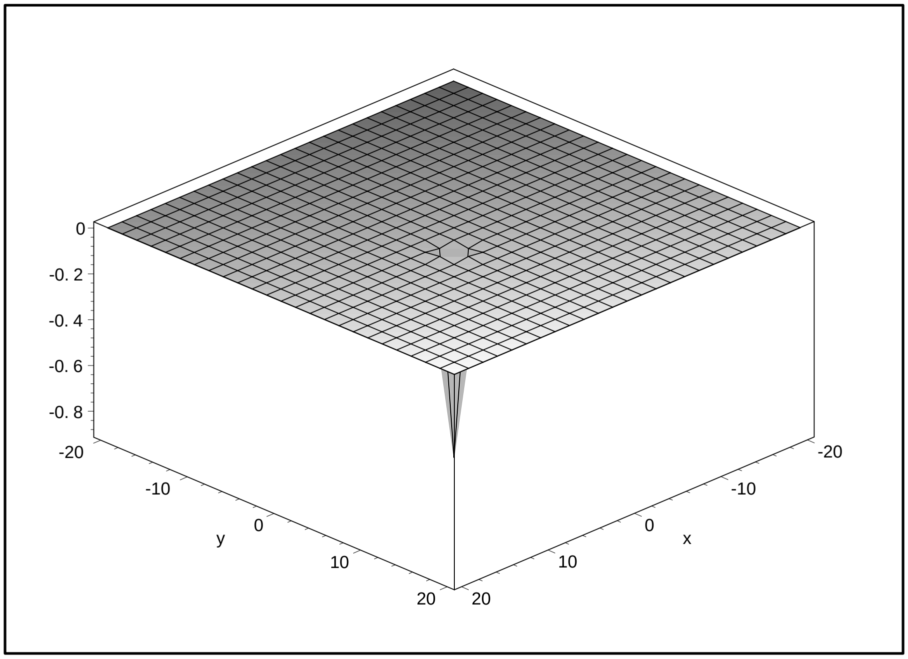
Rysunek 20: Branins's function

2.12 Branins's function

The Branin function is a global optimization test function having only two variables. The function has three equal-sized global optima, and has the following definition

$$f(x_1, x_2) = a(x_2 - bx_1^2 + cx_1 - d)^2 + e(1 - f)\cos(x_1) + e. \tag{12}$$

It is recommended to set the following values of parameters: $a = 1$, $b = \frac{5.1}{4\pi^2}$, $c = \frac{5}{\pi}$, $d = 6$, $e = 10$, $f = \frac{1}{8\pi}$. Three global optima equal $f(x_1, x_2) = 0.397887$ are located as follows: $(x_1, x_2) = (-\pi, 12.275),\ (\pi, 2.275),\ (9.42478, 2.475)$.



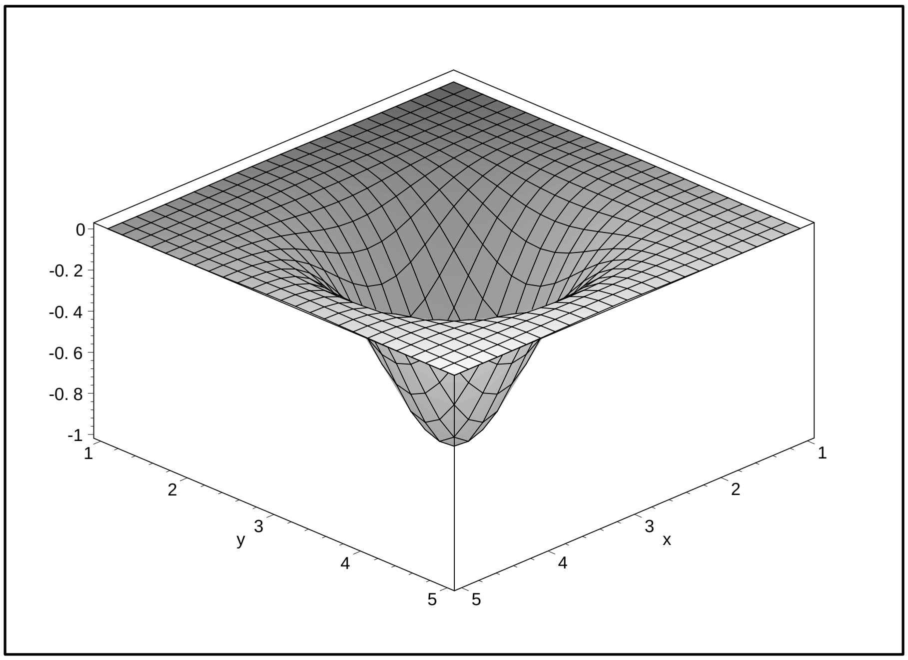
Rysunek 21: Easom's function

2.13 Easom's function

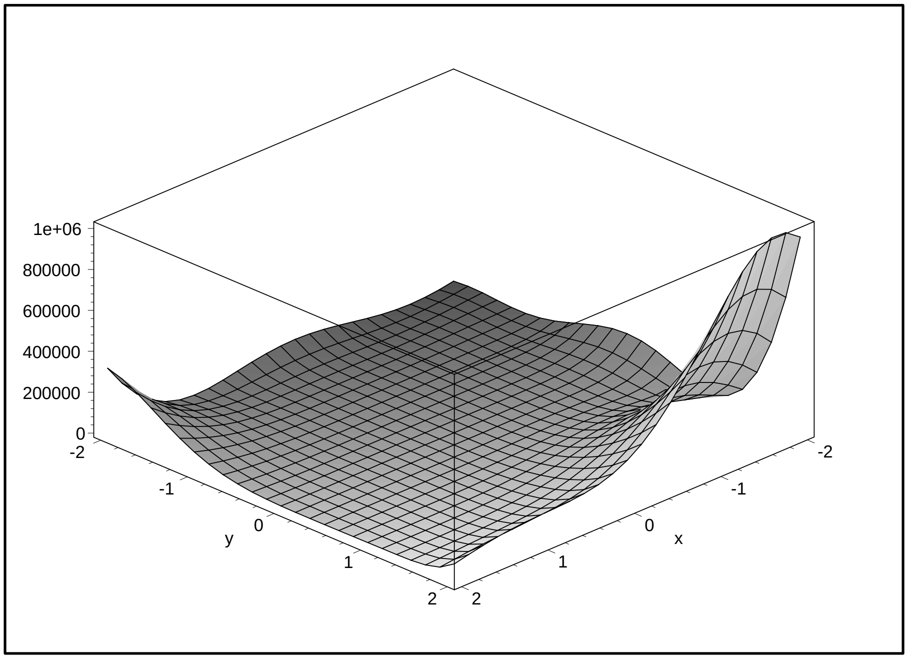
The Easom function is a unimodal test function, where the global minimum has a small area relative to the search space. The function was inverted for minimization. It has only two variables and the following definition

$$f(x_1, x_2) = -\cos(x_1)\cos(x_2)\exp(-(x_1-\pi)^2-(x_2-\pi)^2) \qquad (13)$$

Test area is usually restricted to square $-100 \leq x_1 \leq 100$, $-100 \leq x_2 \leq 100$. Its global minimum equal $f(x) = -1$ is obtainable for $(x_1, x_2) = (\pi, \pi)$.



Rysunek 22: Zoom on Easom's function



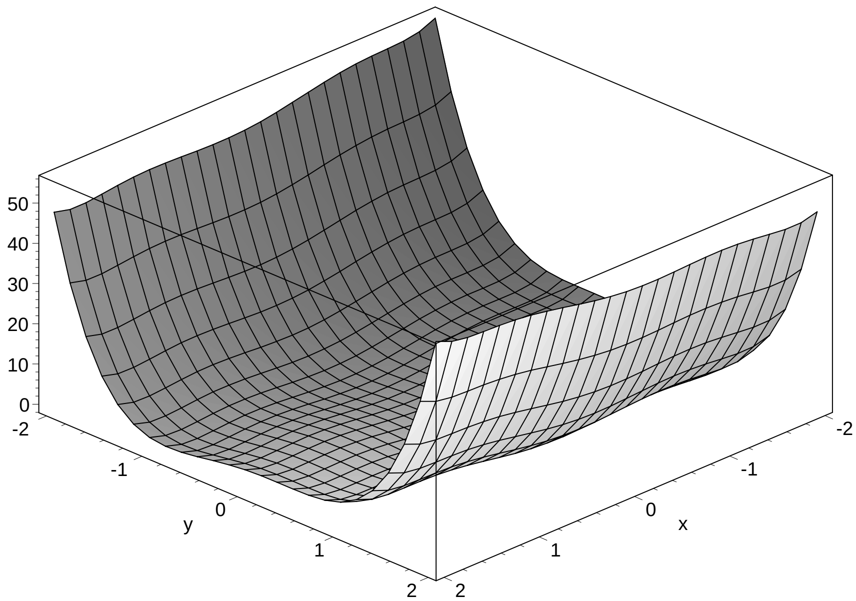
Rysunek 23: Goldstein-Price's function

2.14 Goldstein-Price's function

The Goldstein-Price function is a global optimization test function. It has only two variables and the following definition

$$f(x_1, x_2) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2]\cdot$$
$$[30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2]. \quad (14)$$

Test area is usually restricted to the square $-2 \leq x_1 \leq 2$, $-2 \leq x_2 \leq 2$. Its global minimum equal $f(x) = 3$ is obtainable for $(x_1, x_2) = (0, -1)$.



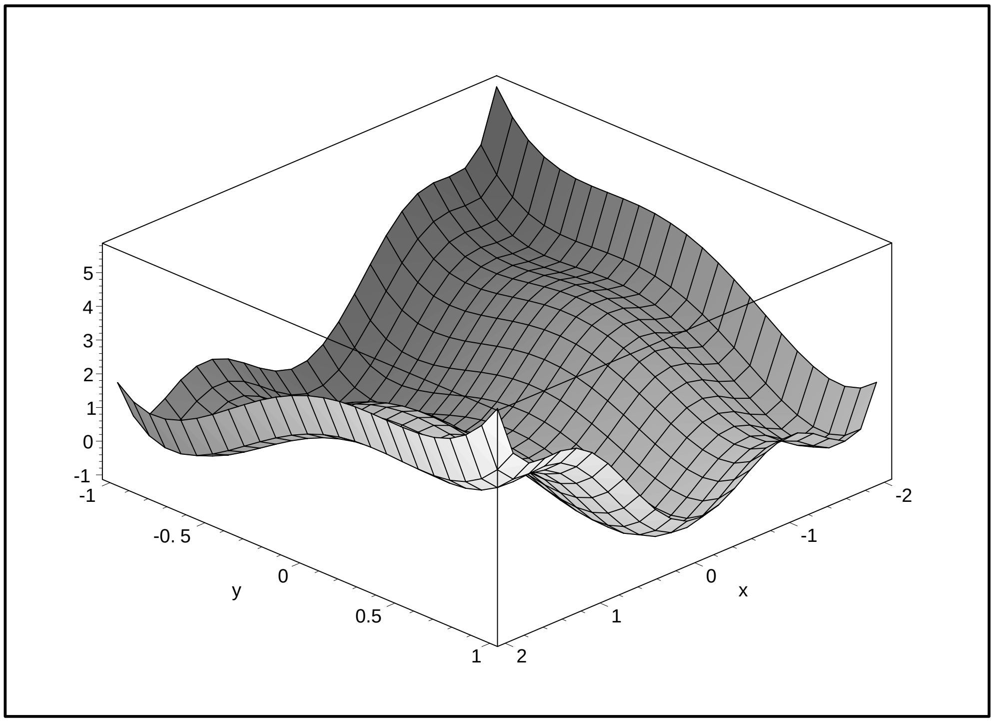
Rysunek 24: Six-hump camel back function

2.15 Six-hump camel back function

The Six-hump camel back function is a global optimization test function. Within the bounded region it owns six local minima, two of them are global ones. Function has only two variables and the following definition

$$f(x_1, x_2) = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1 x_2 + (-4 + 4x_2^2)x_2^2. \tag{15}$$

Test area is usually restricted to the rectangle $-3 \le x_1 \le 3$, $-2 \le x_2 \le 2$. Two global minima equal $f(x) = -1.0316$ are located at $(x_1, x_2) = (-0.0898, 0.7126)$ and $(0.0898, -0.7126)$.



Rysunek 25: Zoom on six-hump camel back function

2.16 Fifth function of De Jong

This is a multimodal test function. The given form of function has only two variables and the following definition

$$f(x_1, x_2) = \{0.002 + \sum_{j=1}^{25} [j + (x_1 - a_{1j})^6 + (x_2 - a_{2j})^6]^{-1}\}^{-1}, \qquad (16)$$

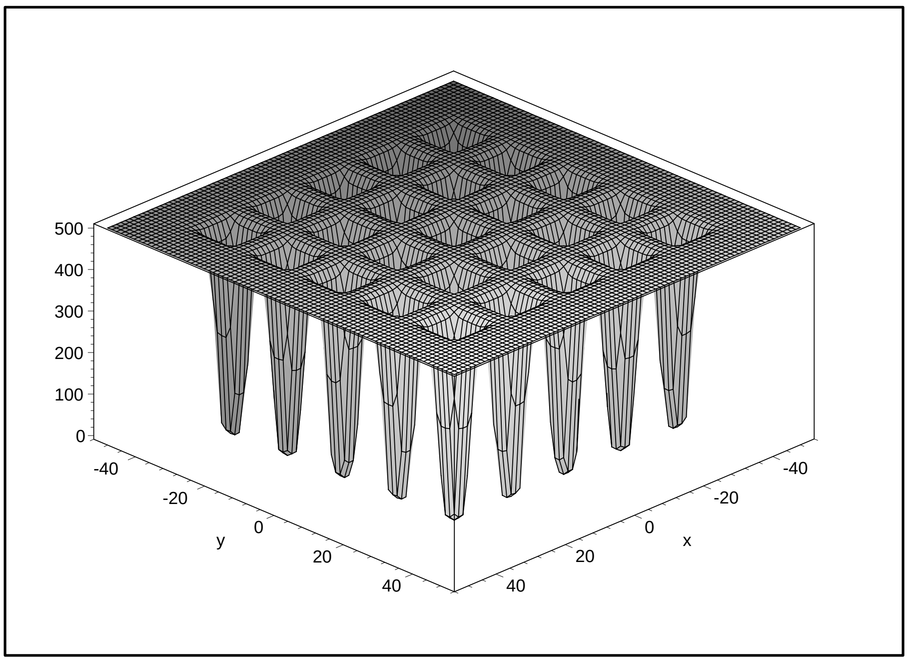
where

$$(a_{ij}) = \begin{pmatrix} -32 & -16 & 0 & 16 & 32 & -32 & \dots & 0 & 16 & 32 \\ -32 & -32 & -32 & -32 & -32 & -16 & \dots & 32 & 32 & 32 \end{pmatrix}$$

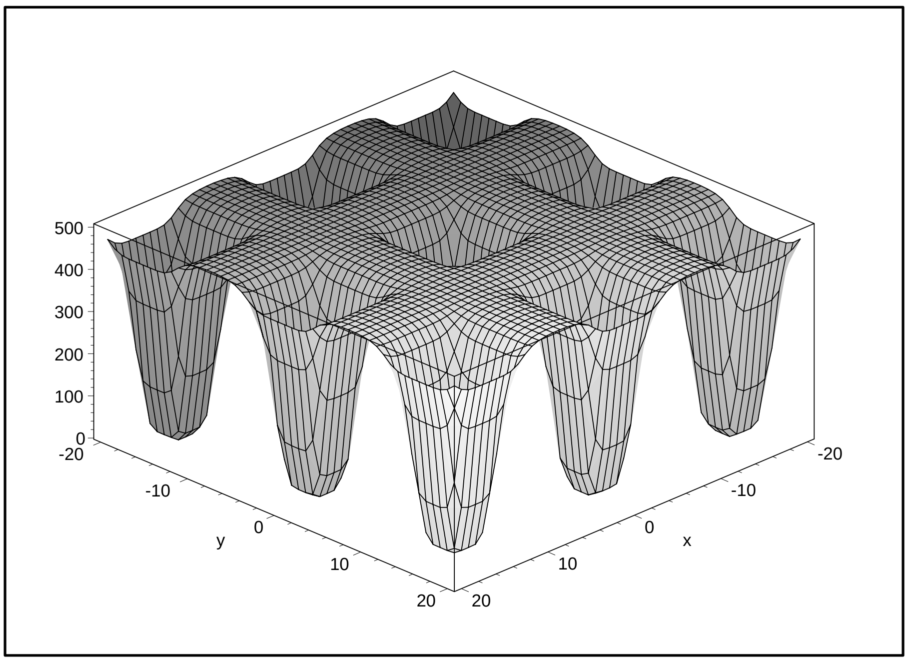
The function can also be rewritten as follows

$$f(x_1, x_2) = \{0.002 + \sum_{i=-2}^{2} \sum_{j=-2}^{2} [5(i+2) + j + 3 + (x_1 - 16j)^6 + (x_2 - 16i)^6]^{-1}\}^{-1}, \qquad (17)$$

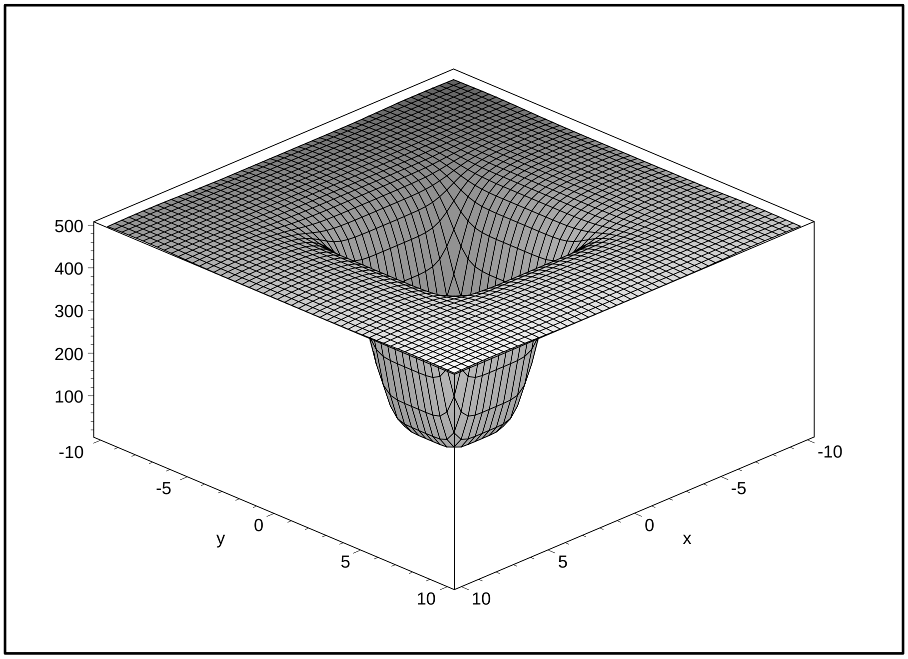
Test area is usually restricted to the square $-65.536 \leq x_1 \leq 65.536$, $-65.536 \leq x_2 \leq 65.536$.



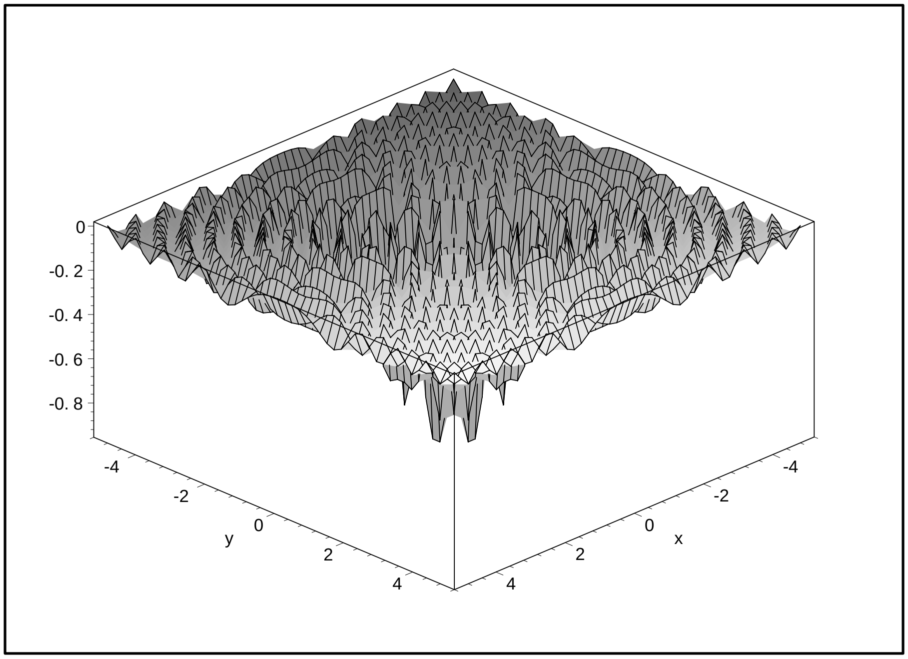
Rysunek 26: An overview of fifth function of De Jong



Rysunek 27: Medium-scale view on the fifth function of De Jong



Rysunek 28: Zoom on fifth function of De Jong



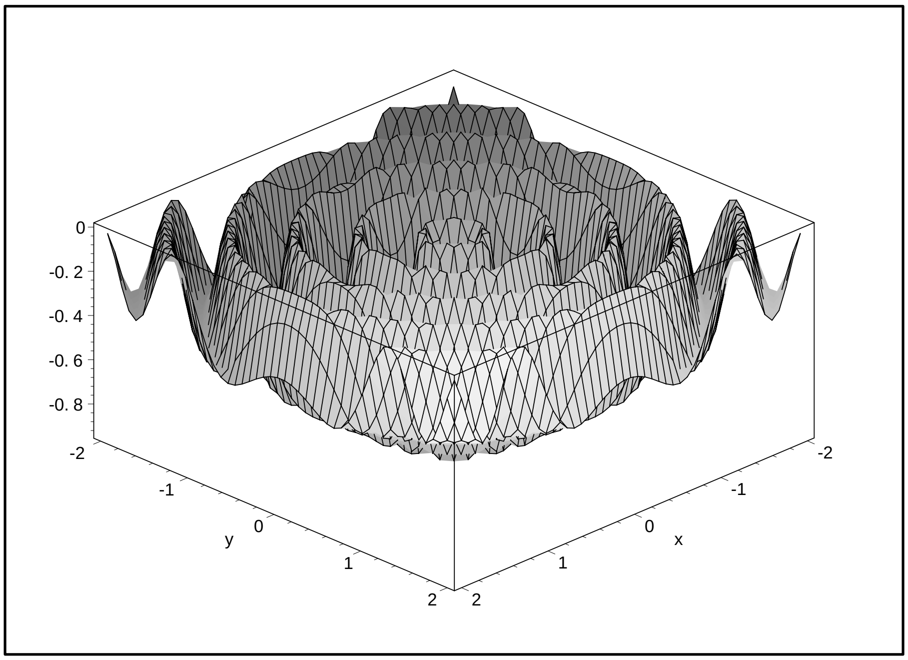
Rysunek 29: An overview of "drop wave" function

2.17 "Drop wave" function

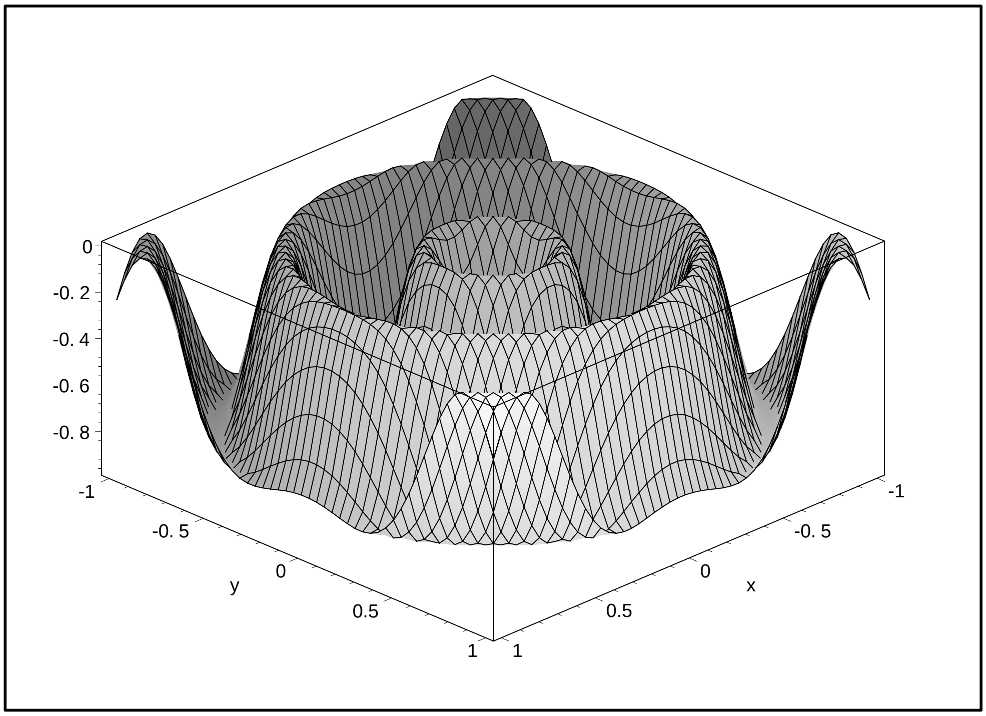
This is a multimodal test function. The given form of function has only two variables and the following definition

$$f(x_1, x_2) = -\frac{1 + \cos(12\sqrt{x^2 + x_2^2})}{\frac{1}{2}(x_1^2 + x_2^2) + 2} \tag{18}$$

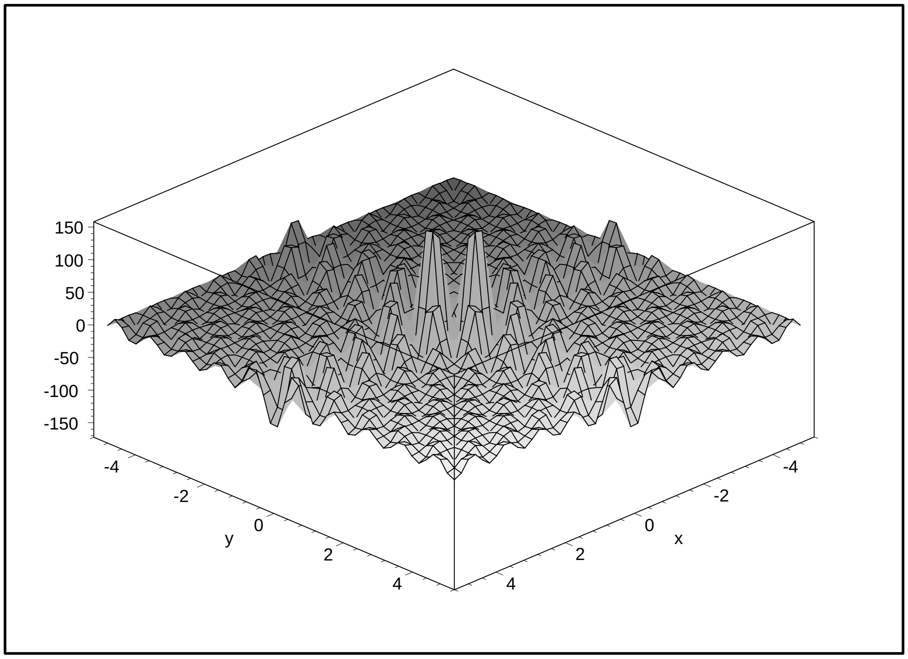
Test area is usually restricted to the square $-5.12 \le x_1 \le 5.12$, $-5.12 \le x_2 \le 5.12$.



Rysunek 30: Medium-scale view on “drop wave” function



Rysunek 31: Zoom on “drop wave” function



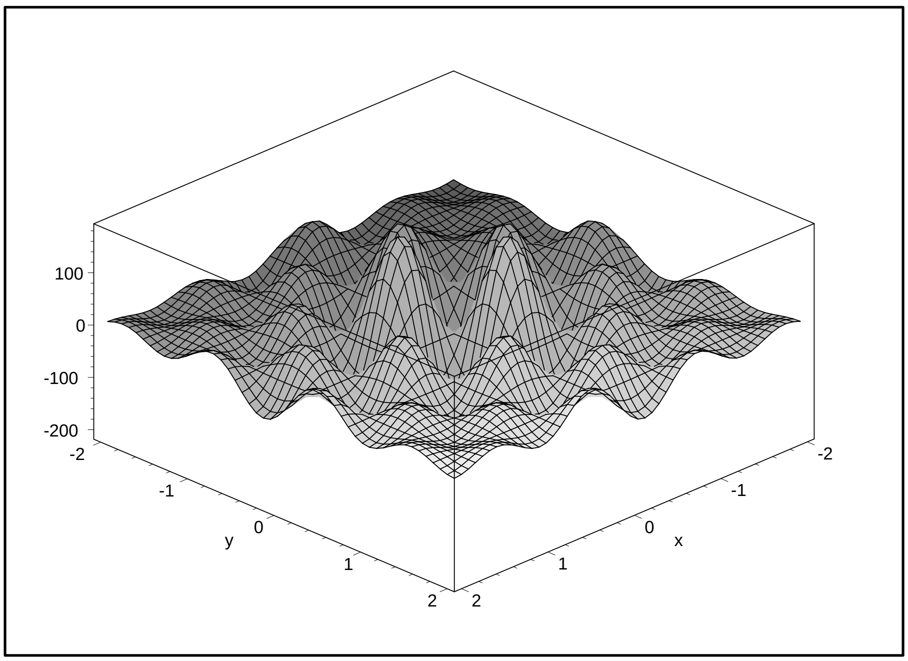
Rysunek 32: An overview of Shubert's function

2.18 Shubert's function

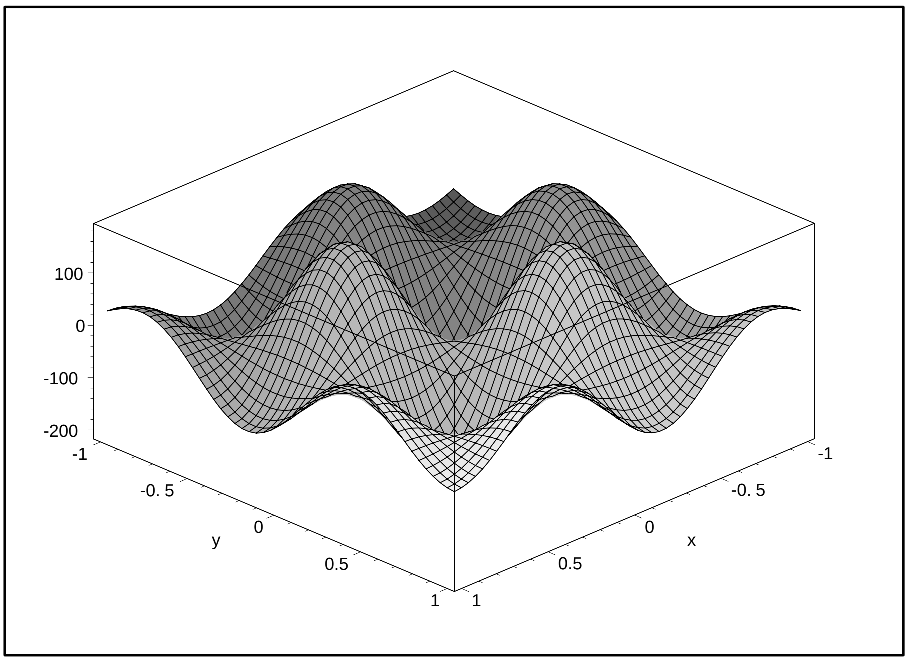
This is a multimodal test function. The given form of function has only two variables and the following definition

$$f(x_1, x_2) = -\sum_{i=1}^{5} i\cos((i+1)x_1 + 1)\sum_{i=1}^{5} i\cos((i+1)x_2 + 1), \quad (19)$$

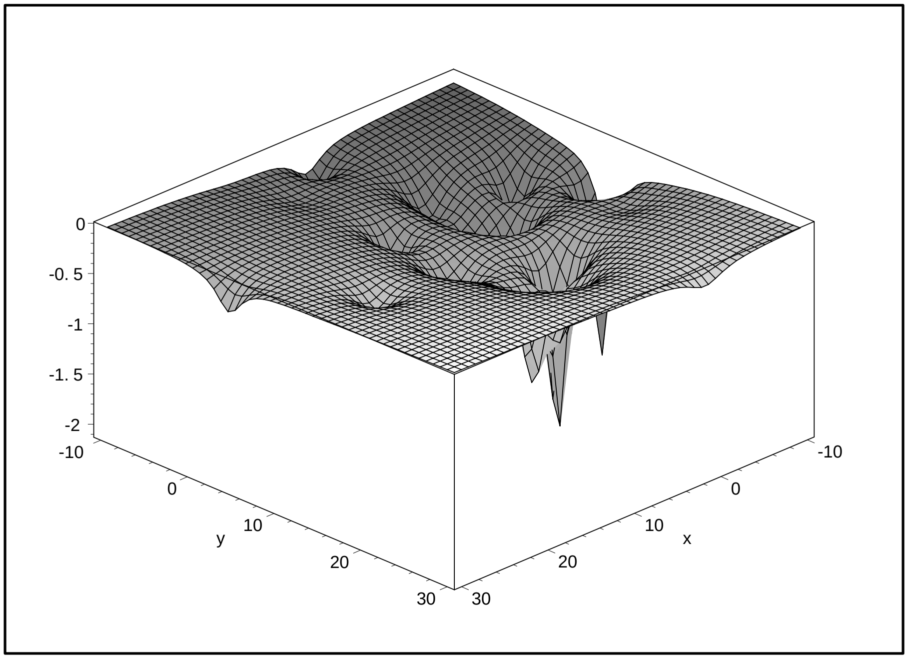
Test area is usually restricted to the square $-5.12 \leq x_1 \leq 5.12$, $-5.12 \leq x_2 \leq 5.12$.



Rysunek 33: Medium-scale view on Shubert's function



Rysunek 34: Zoom on Shubert’s function



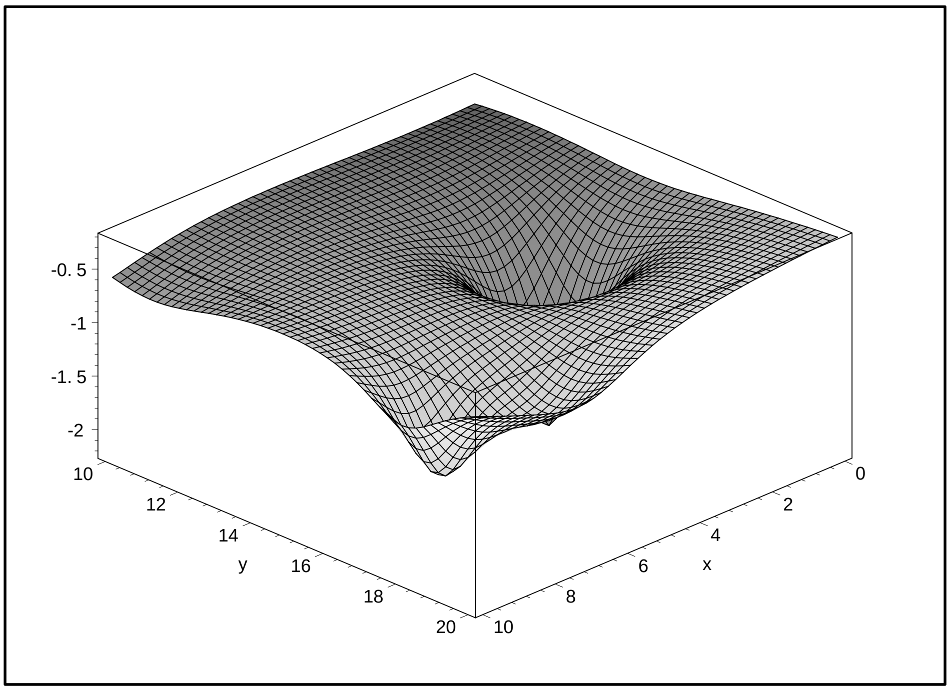
Rysunek 35: An overview of Shekel's foxholes in 2D. $f(x,y) = -\sum_{i=1}^{m}[(x - a_j)^2 + (y - b_j)^2 + c_j]^{-1}$

2.19 Shekel's foxholes

This is a multimodal test function. It has the following definition

$$f(x) = -\sum_{i=1}^{m}(\sum_{j=1}^{n}[(x_j - a_{ij})^2 + c_j])^{-1}, \qquad (20)$$

where (c_i, $i = 1, \ldots, m$), (a_{ij}, $j = 1, \ldots, n$, $i = 1, \ldots, m$) are constant numbers fixed in advance. It is recommended to set $m = 30$.



Rysunek 36: Zoom on Shekel's foxholes in 2D. $f(x,y) = -\sum_{i=1}^{m}[(x-a_j)^2 + (y-b_j)^2 + c_j]^{-1}$

2.20 Deceptive functions

A deceptive problem is a class of problems in which the total size of the basins for local optimum solutions is much larger than the basin size of the global optimum solution. Clearly, this is a multimodal function. The general form of deceptive function is given by the following formulae

$$f(x) = -\left[\frac{1}{n}\sum_{i=1}^{n} g_i(x_i)\right]^{\beta}, \tag{21}$$

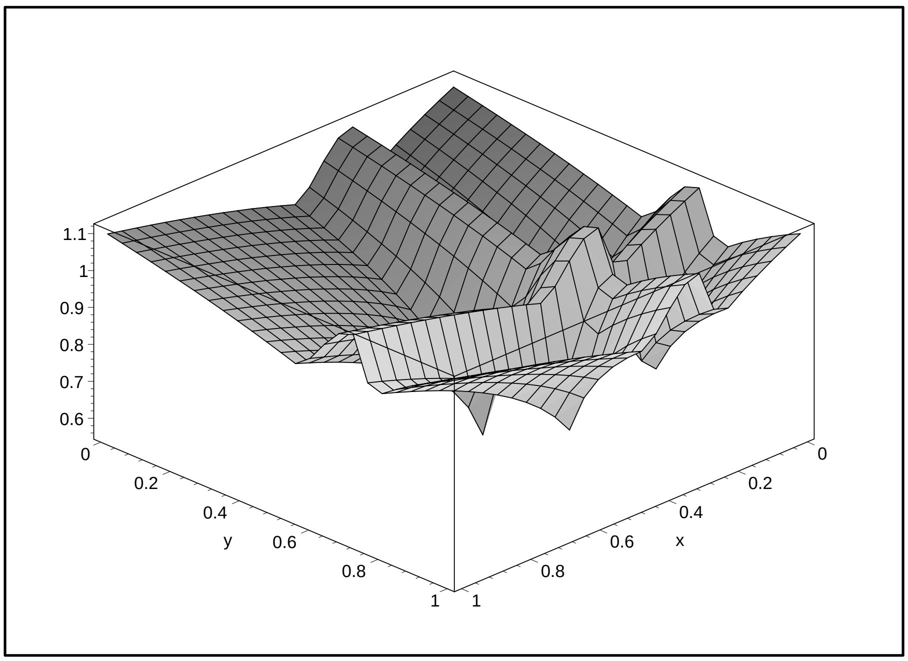
where β is an fixed non-linearity factor.

It has been defined in the literature at least three types of deceptive problems, depending the form of $g_i(x_i)$. A complex deceptive problem (Type III), in which the global optimum is located at $x_i = \alpha_i$, where α_i is a unique random number between 0 and 1 depending on the dimension i. To this aim the following form of auxiliary functions has been proposed

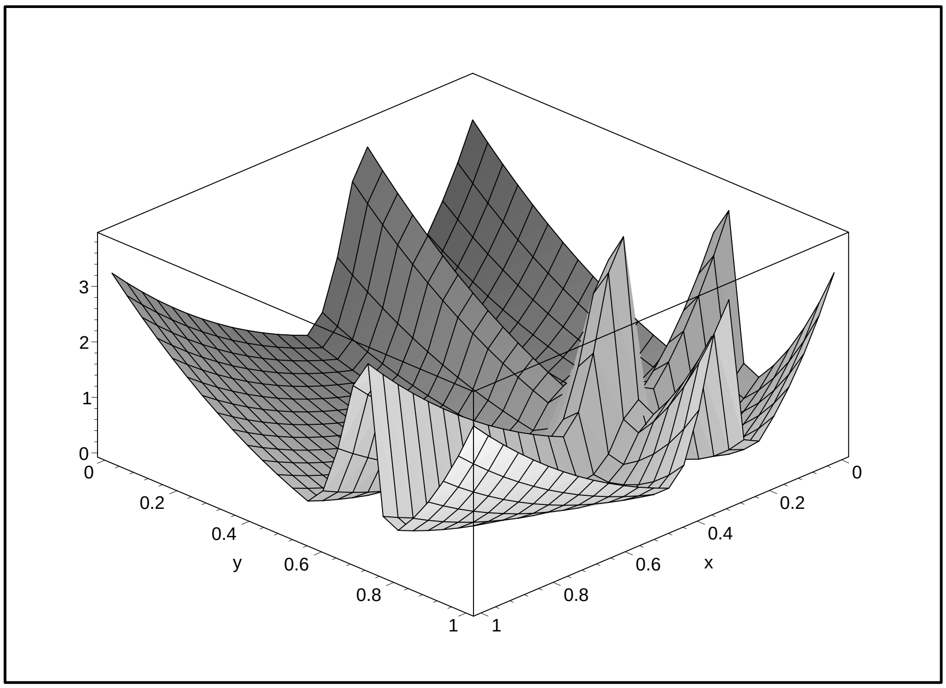
$$g_i(x_i) = \begin{cases} -\frac{x}{\alpha_i} + \frac{4}{5} & \text{if} \quad 0 \leq x_i \leq \frac{4}{5}\alpha_i \\ \frac{5x}{\alpha_i} - 4 & \text{if} \quad \frac{4}{5}\alpha_i < x_i \leq \alpha_i \\ \frac{5(x-\alpha_i)}{\alpha_i - 1} + 1 & \text{if} \quad \alpha_i < x_i \leq \frac{1+4\alpha_i}{5} \\ \frac{x-1}{1-\alpha_i} + \frac{4}{5} & \text{if} \quad \frac{1+4\alpha_i}{5} < x_i \leq 1 \end{cases} \tag{22}$$

The two other types of deceptive problems (Types I and II) are special cases of the complex deceptive problem, with $\alpha_i = 1$ (Type I), or $\alpha_i = 0$ or 1 at random (Type II) for each dimension i, $i = 0, \ldots, n$. Clearly formulae (22) should be suitable adjusted for type I and II.

For all three types of $g_i(x_i)$, the region with local optima is $5^n - 1$ times larger than the region with a global optimum in the n-dimensional space. The number of local optima is $2^n - 1$ for Type I and Type II deceptive problems and $3^n - 1$ for Type III.



Rysunek 37: Deceptive function of Type III in 2D. $\alpha_1 = 0.3$, $\alpha_2 = 0.7$, $\beta = 0.2$



Rysunek 38: Deceptive function of Type III in 2D. $\alpha_1 = 0.3$, $\alpha_2 = 0.7$, $\beta = 2.5$