Optimisation

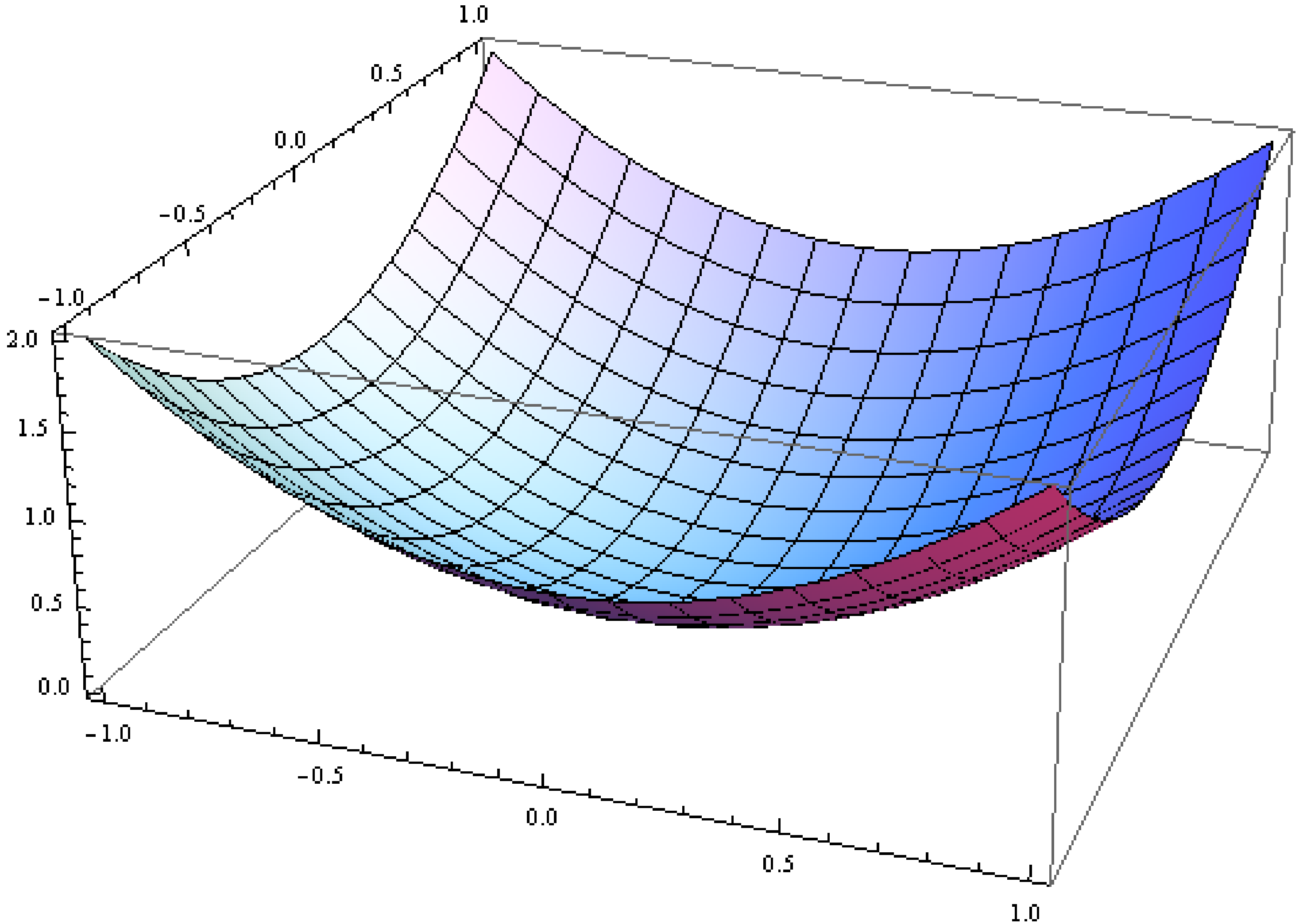
Définition

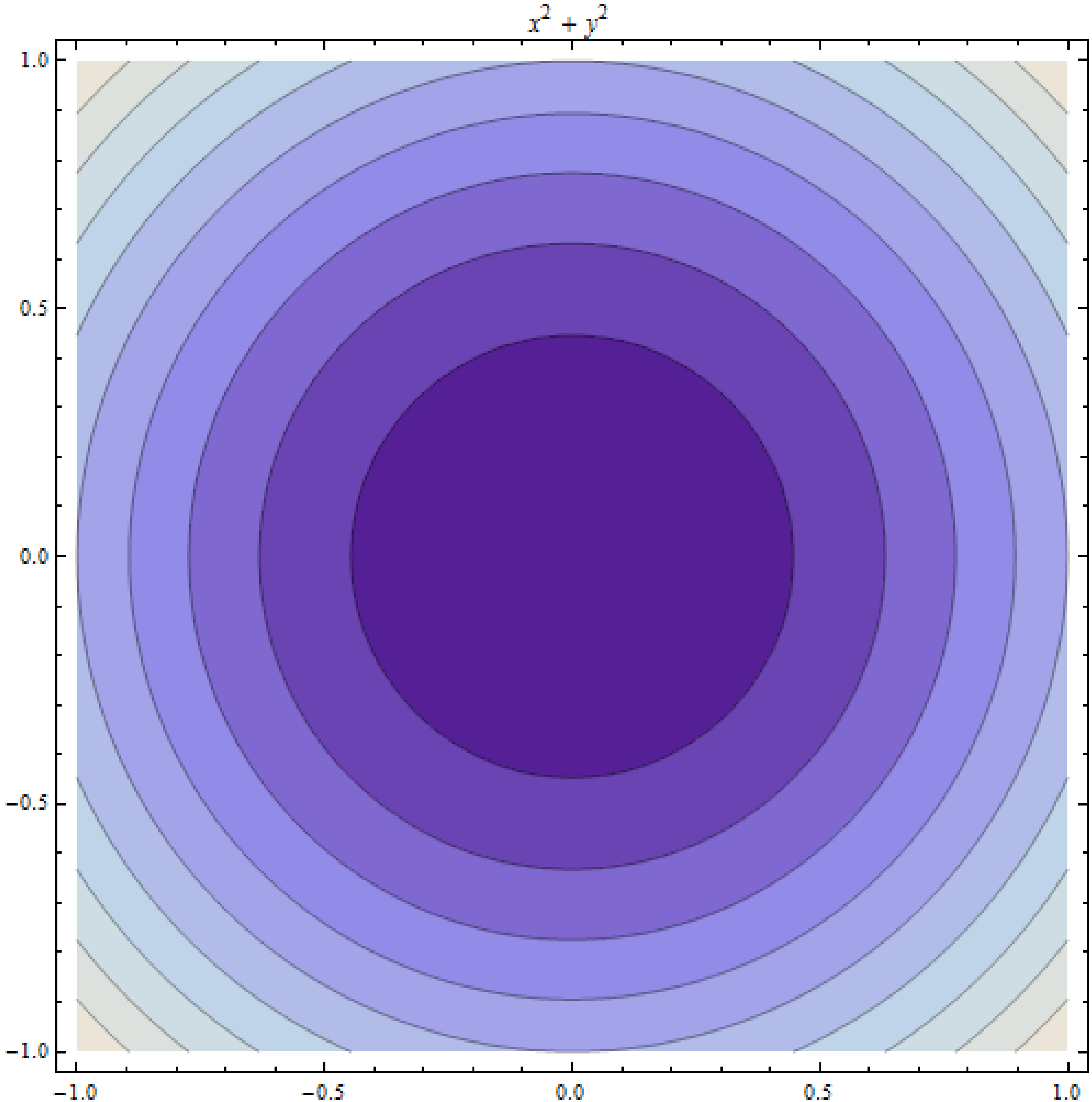
- L’optimisation d’un problème ou d’une fonction consiste à chercher la valeur des variables de la fonction qui minimise (ou qui maximise) sa valeur.
- L’optimisation est généralement un problème compliqué, et il existe de multiples techniques, dont l’utilisation doit être adapté au type du problème.

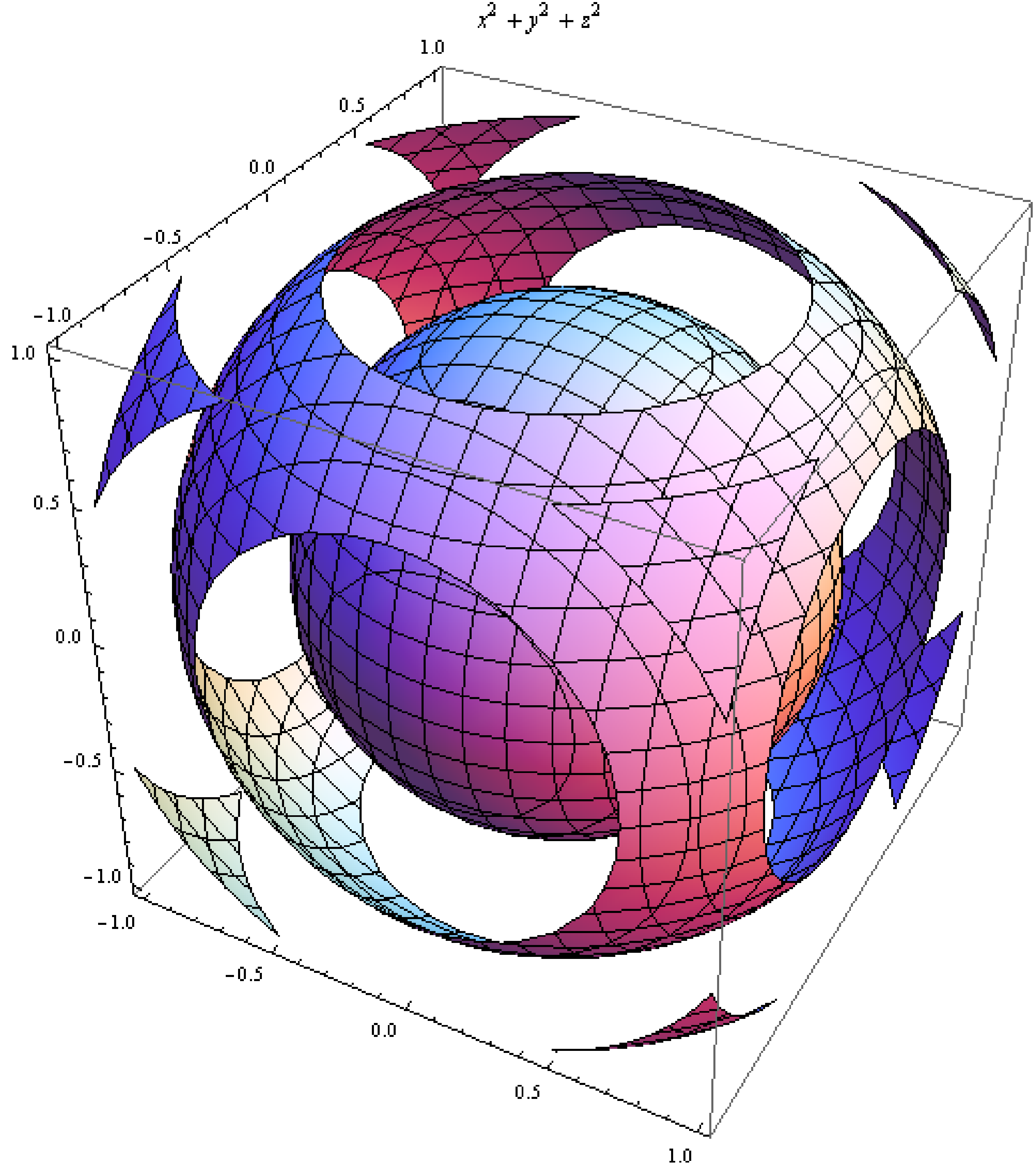
Caractérisation des techniques d'optimisation

- Optimisation globale/locale
  - L'optimisation globale consiste à chercher le maximum de la fonction sur l'ensemble de définition
  - L'optimisation locale consiste à chercher le maximum de la fonction au voisinage d'un point.
- Méthode déterministe/stochastique
  - Une méthode stochastique parcourt l'espace de recherche de façon « aléatoire ». Deux exécutions successives peuvent donner deux résultats différents.
  - Une méthode déterministe parcourt toujours l'espace de recherche de la même façon.

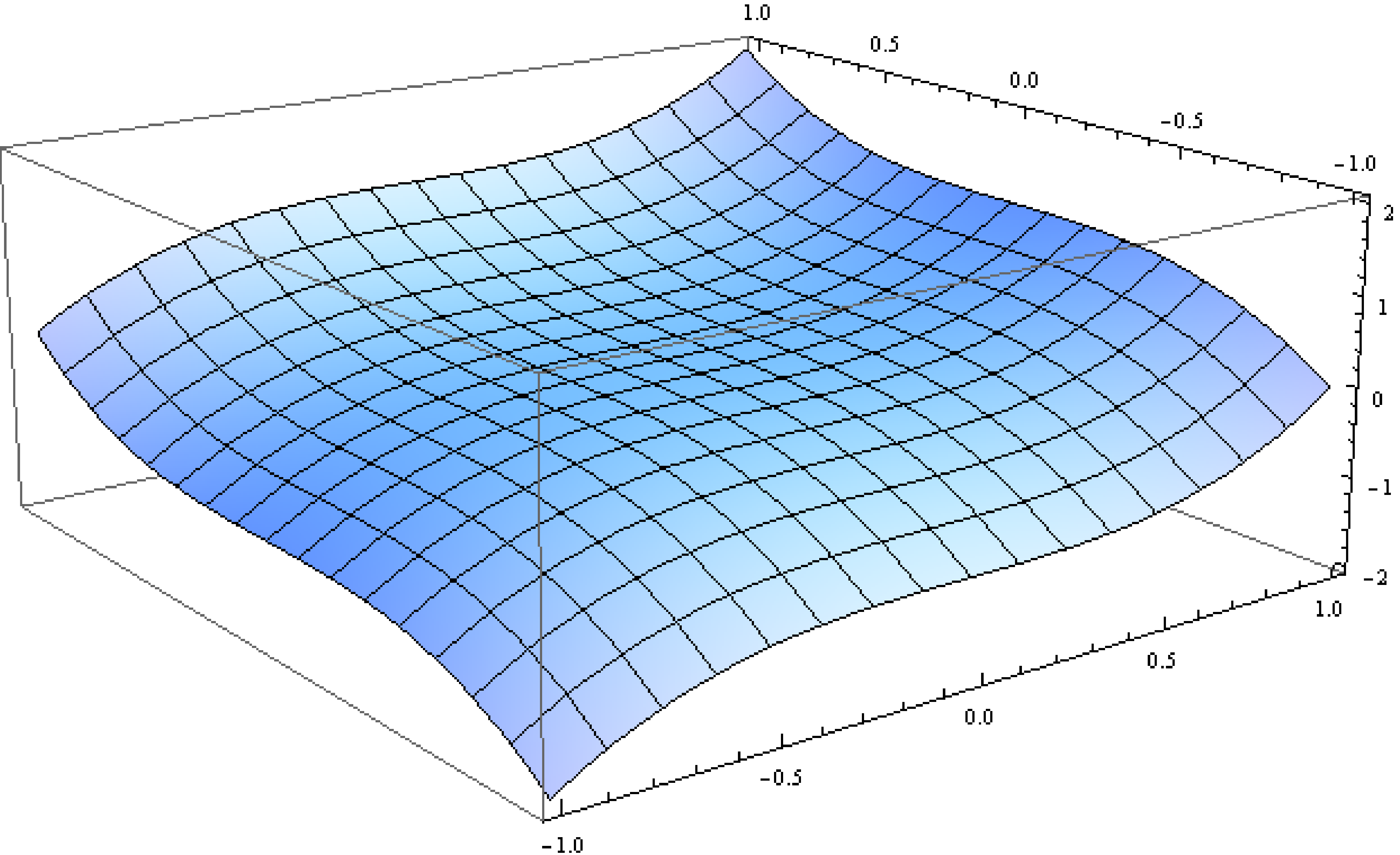
x^2+y^2

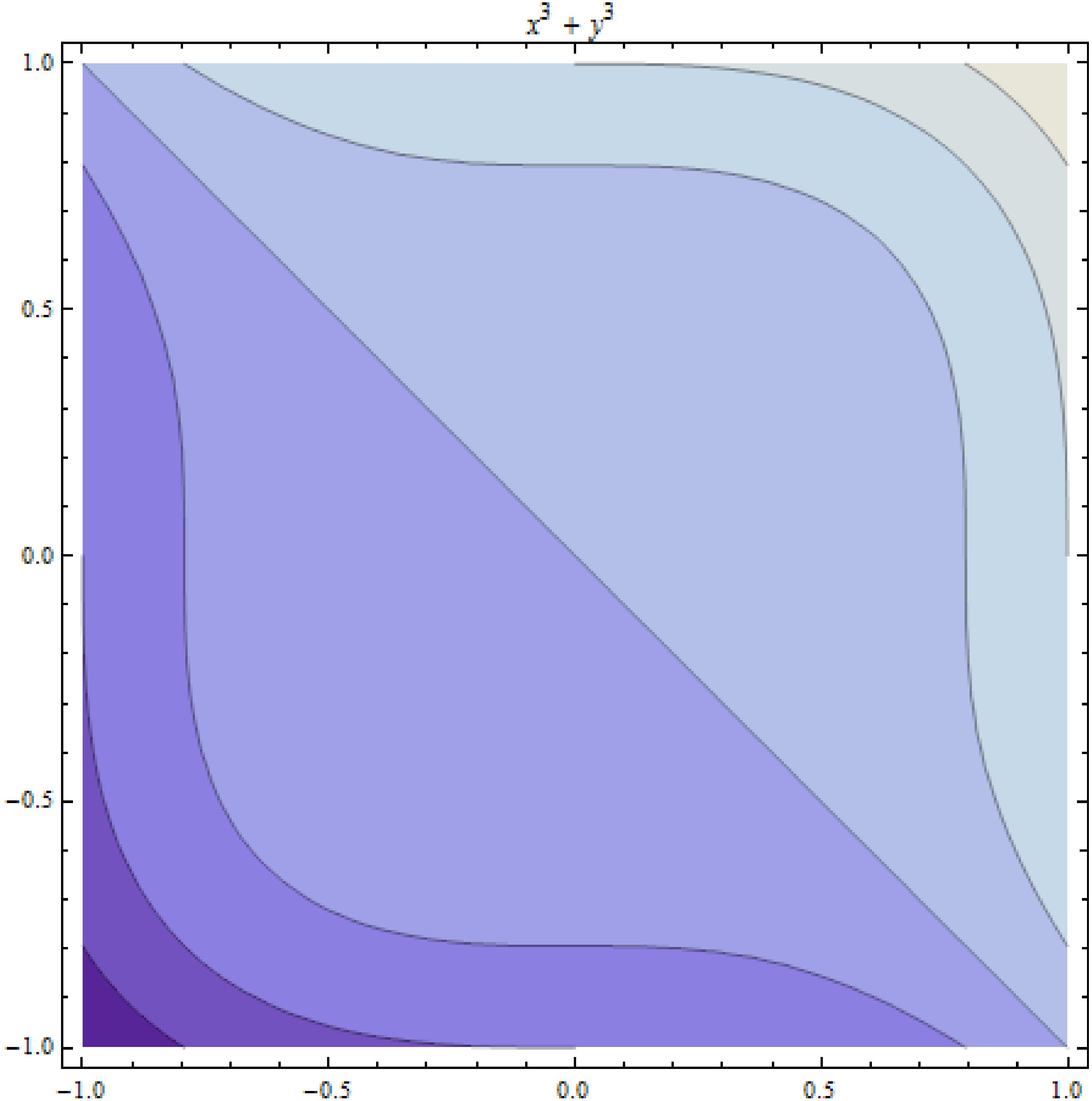



$x^2 + y^2$
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0



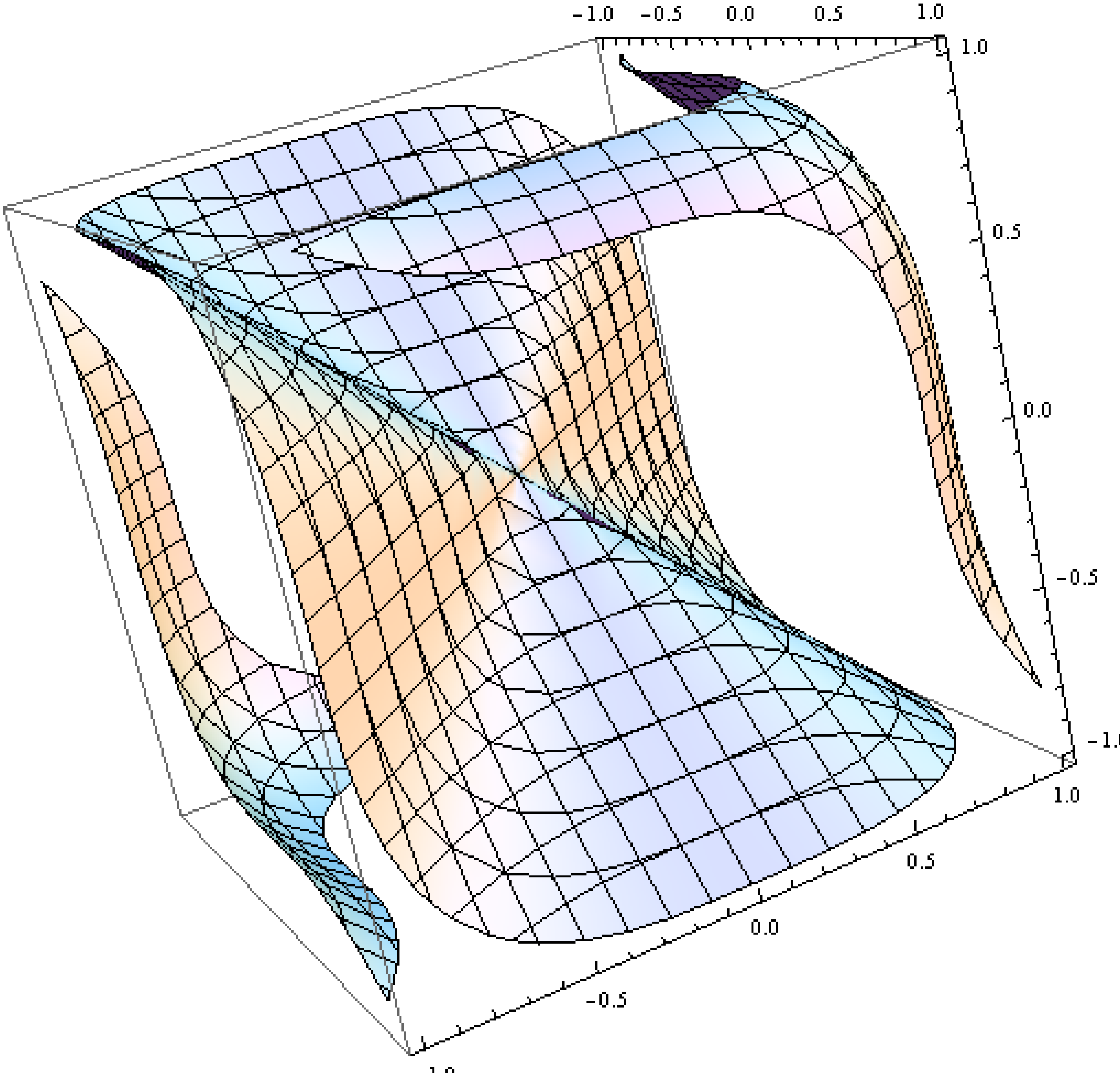
$x^2 + y^2 + z^2$
1.0
0.5
0.0
-0.5
-1.0
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0

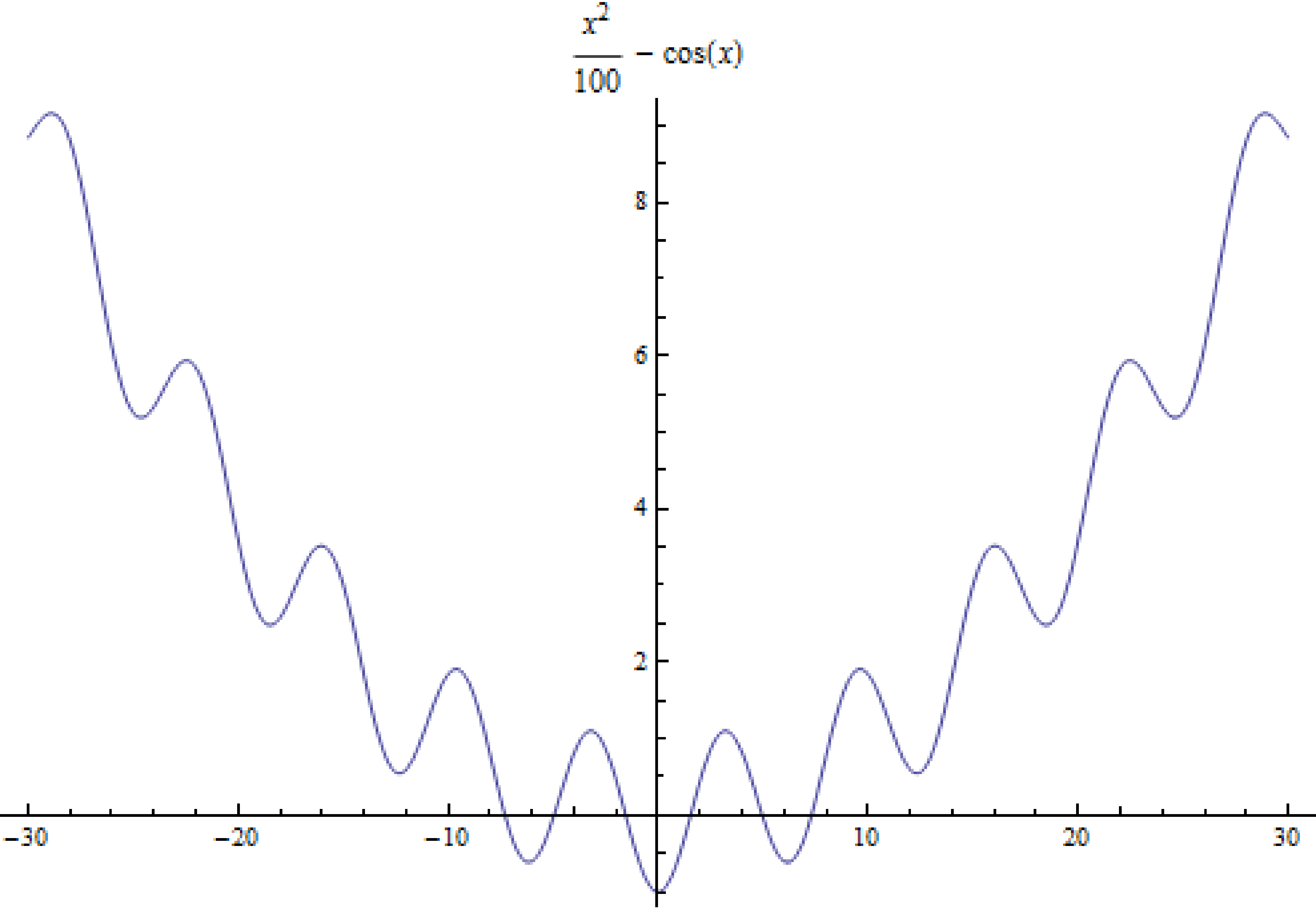

1.0
0.5
0.0
-0.5
-1.0
2
1
0
-1
-2
1.0
0.5
0.0
-0.5
-1.0



$x^3 + y^3$
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

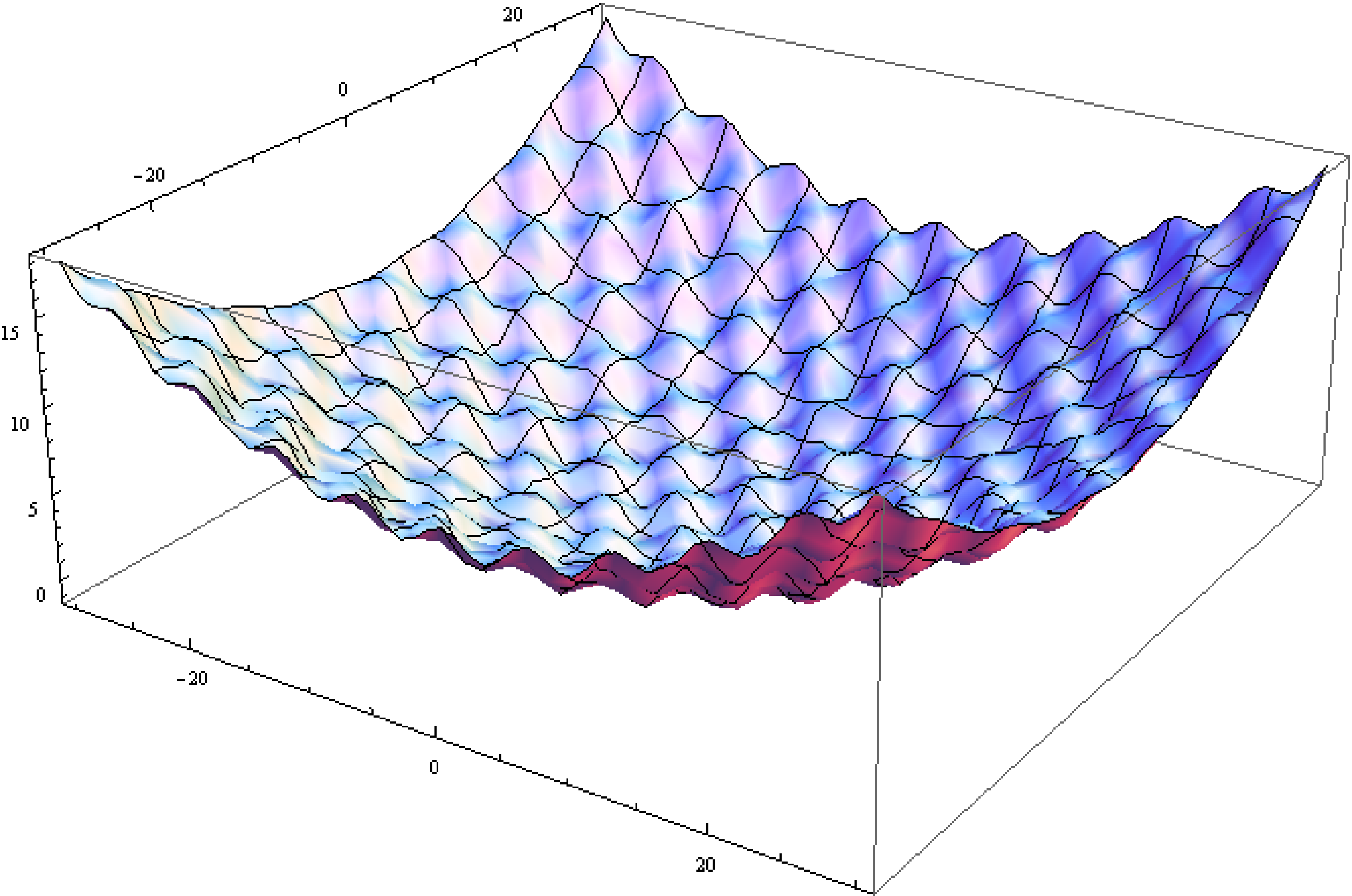
$x^3 + y^3 + z^3$

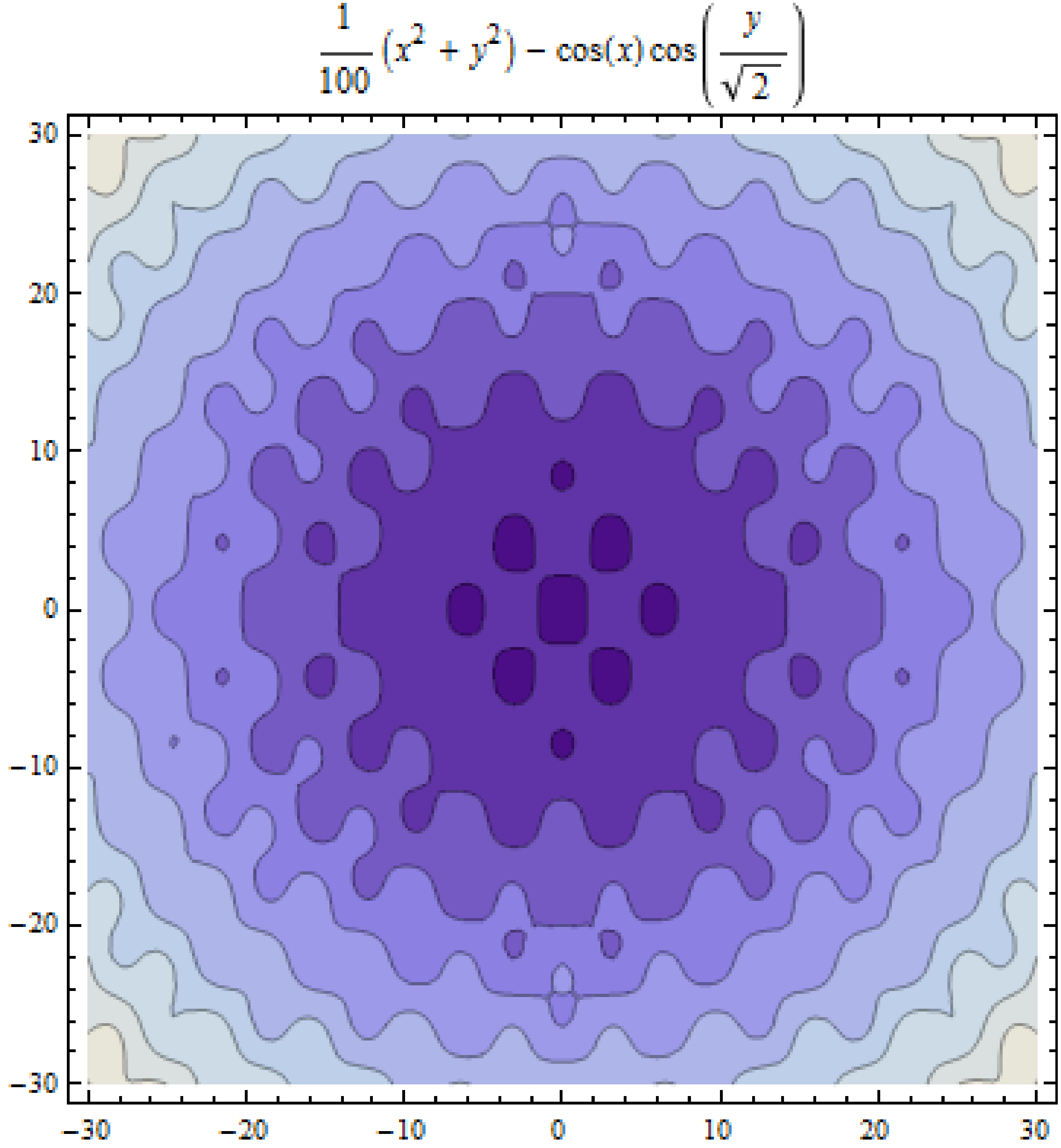




$\frac{x^2}{100} - \cos(x)$
8
6
4
2
−30
−20
−10
10
20
30

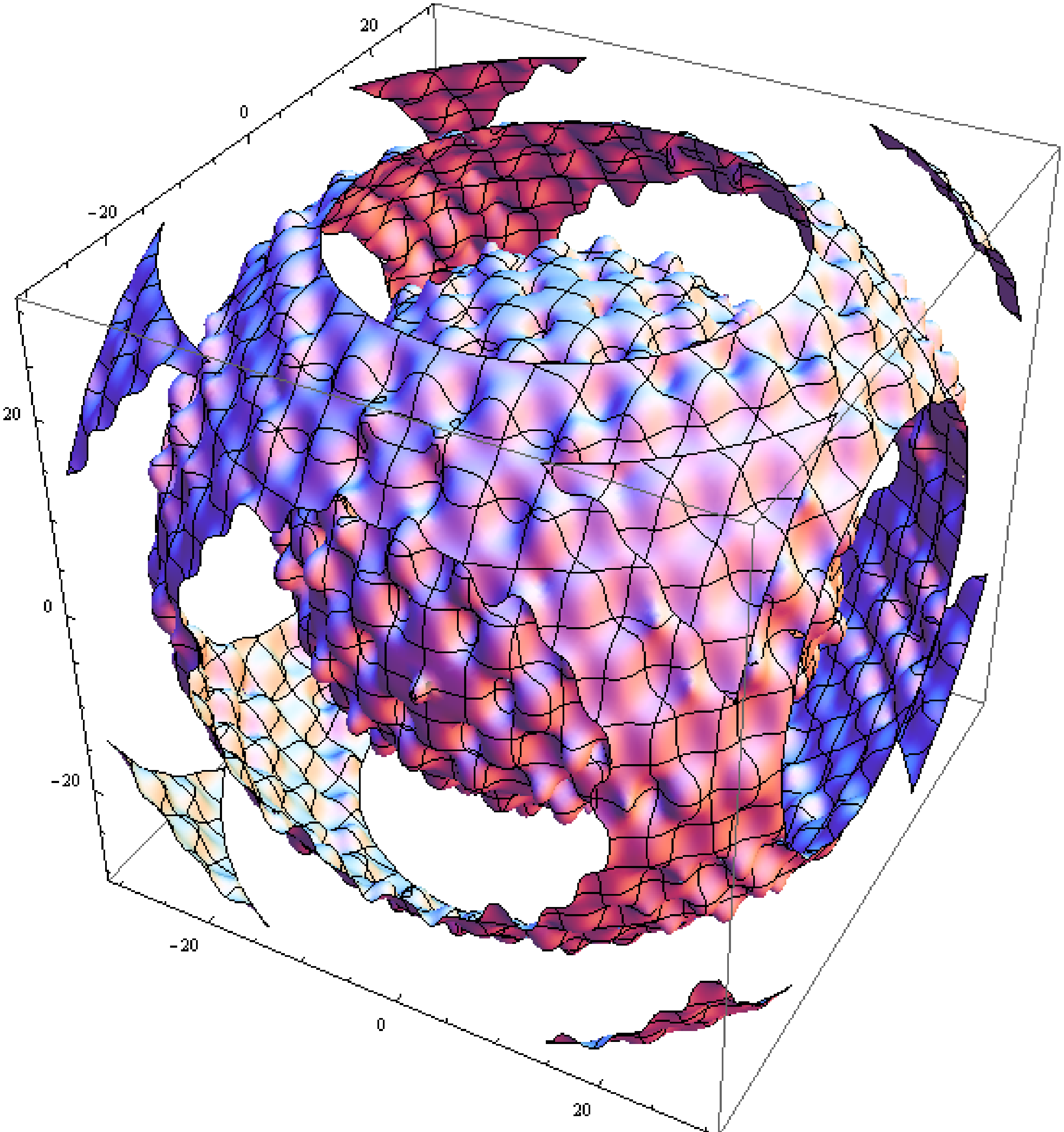
$$\frac{1}{100}\left(x^2+y^2\right)-\cos(x)\cos\left(\frac{y}{\sqrt{2}}\right)$$





$\frac{1}{100}(x^2+y^2)-\cos(x)\cos\left(\frac{y}{\sqrt{2}}\right)$
30
20
10
0
−10
−20
−30
−30
−20
−10
0
10
20
30

$$\frac{1}{100}\left(x^2+y^2+z^2\right)-\cos(x)\cos\left(\frac{y}{\sqrt{2}}\right)\cos\left(\frac{z}{\sqrt{3}}\right)$$



Méthode de la dérivée

- Lorsque l'on sait calculer explicitement les zéros de la dérivée (ou du gradient), on sait que $f'(X)=0$ est une condition nécessaire pour que X soit un point extrémal hors des frontières de l'espace de définition.
- Cette méthode ne fonctionne que pour des fonctions analytiques simples.

Méthodes déterministes locales: le gradient

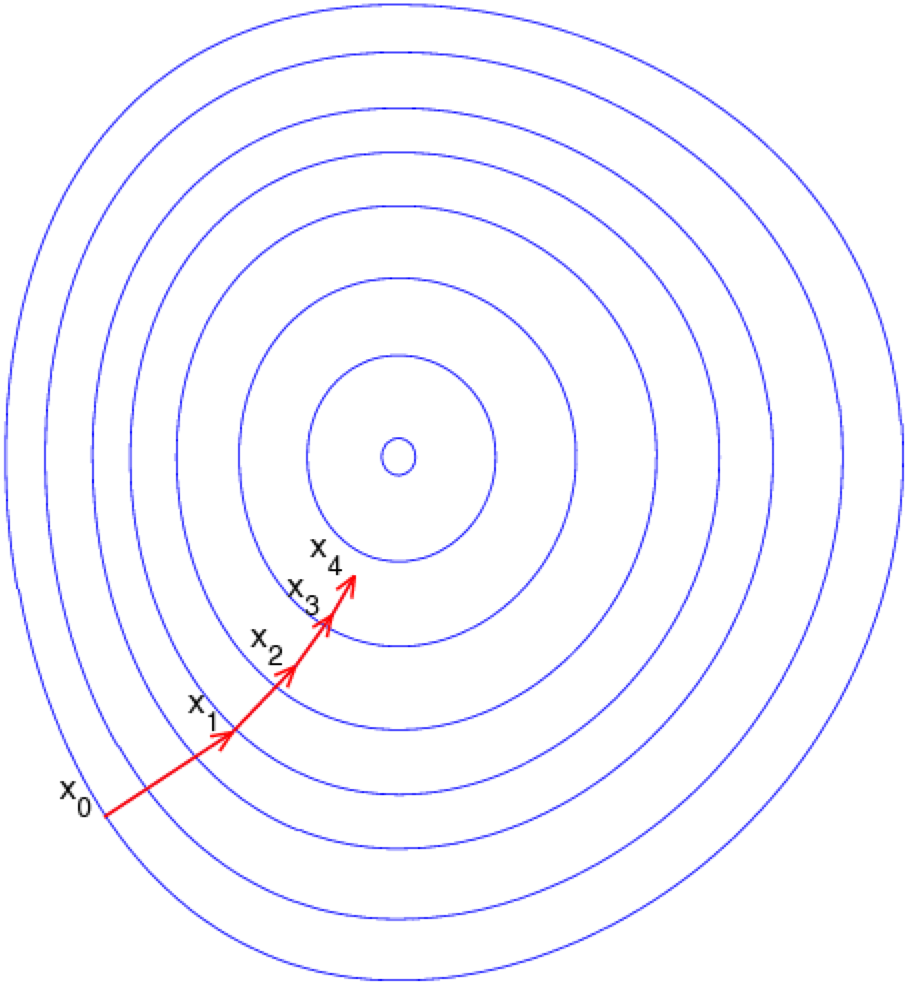
Soit f(X) fonction d'un vecteur X, à valeurs réelles, et son gradient f'(X). On calcule la suite:

$X_{n+1} = X_n - a\, f'(X_n)$, $a>0$

Le choix de a est idéalement fait en minimisant pour a>0:

$G(a)=f(X_n - a\, f'(X_n))$

Ceci n'est généralement pas possible, et on utilise pour trouver a des méthodes approchées.



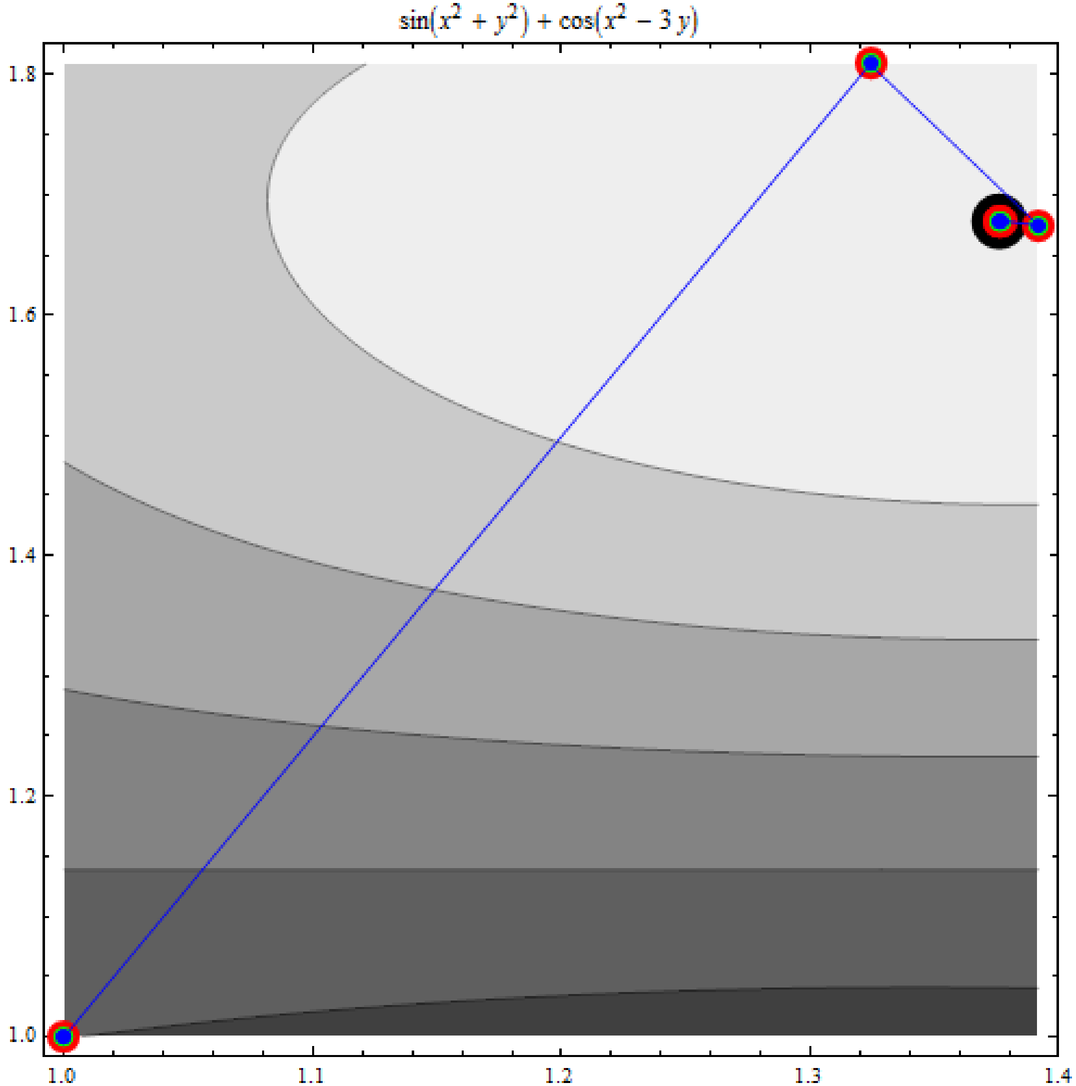
Méthodes déterministes locales d'ordre 2

- Pour accélérer la descente, on utilise les informations apportées par la dérivée seconde de la fonction (le Hessien pour les fonctions à plusieurs variables)
- Ces méthodes nécessitent de calculer la dérivée et le Hessien simultanément.

Méthodes déterministes locales d'ordre 2

- $f(y) = f(x) + f'(x)\,(y-x) + \frac{1}{2} f''(x)\,(y-x)^2 + d$
- En minimisant la forme quadratique en y:
  - $f'(x)+f''(x)(y-x)=0$ soit $y = x - f'(x)/f''(x)$
- Algorithme:
  - $x_{n+1} = x_n - f'(x_n) \,/\, f''(x_n)$
- Méthode connue en général sous le nom de méthode de Newton.
- Sa convergence est plus rapide que la méthode de gradient.

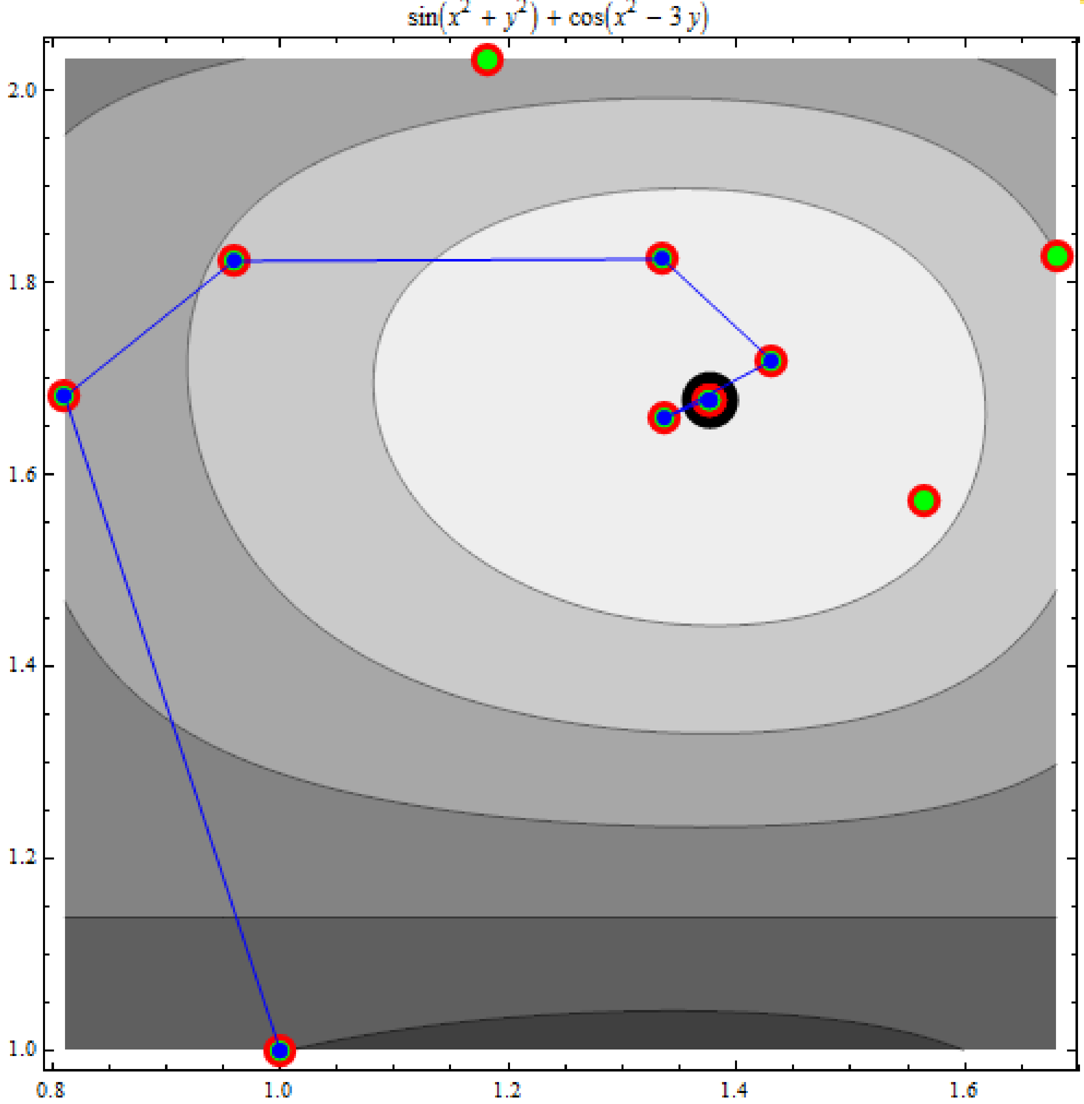
Méthodes déterministes locales d'ordre 2 (Newton)



Méthodes déterministes locales: BFGS

- BFGS est une méthode qui approxime la matrice du hessien par analyse des gradients successifs
- Elle ne nécessite que la connaissance du gradient.
- Elle est plus rapide que le gradient, et moins rapide que la méthode de Newton
- C'est une méthode très utilisée en pratique.

BFGS



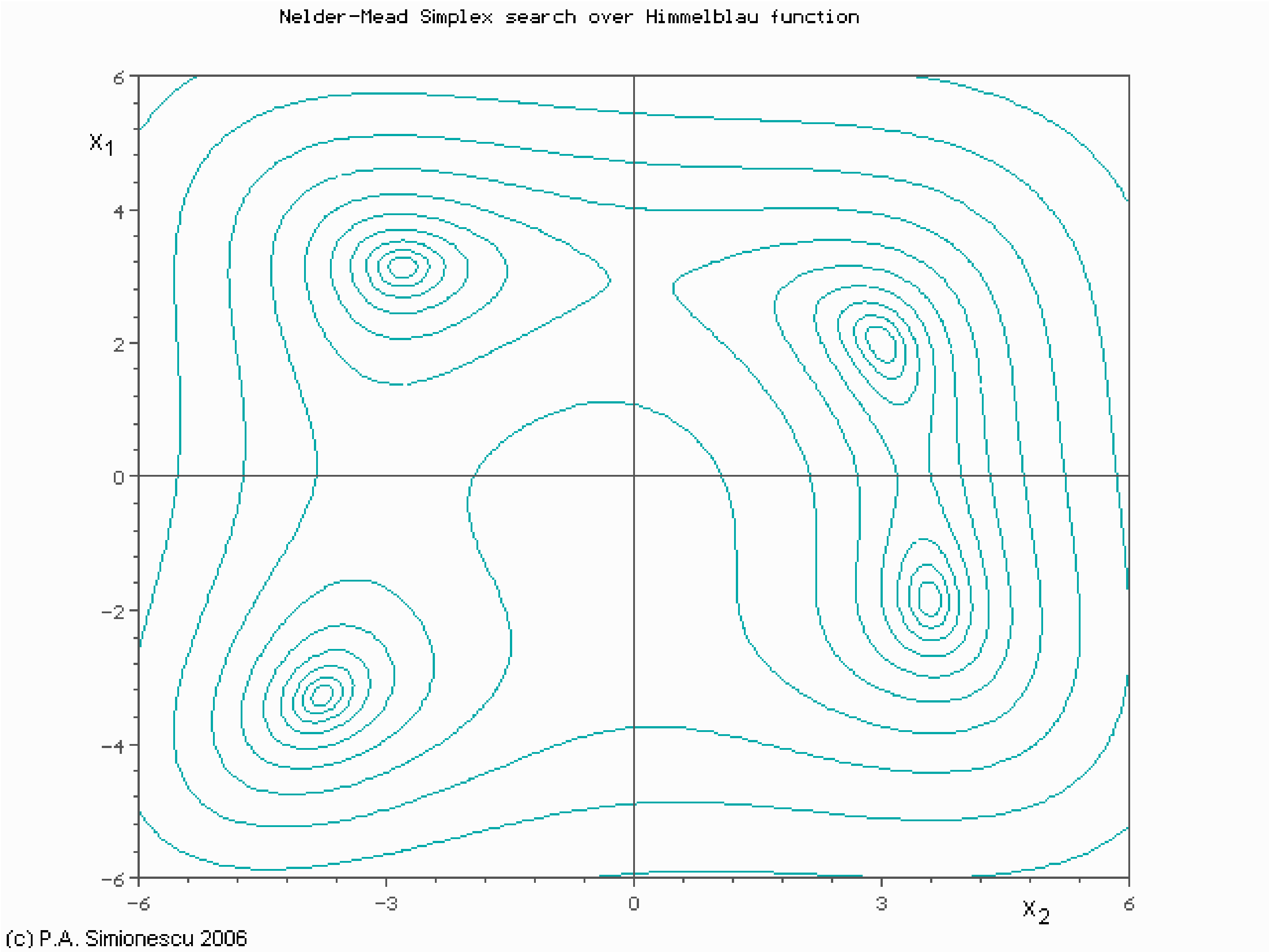
Méthodes locales: le simplexe de Nelder-Mead

- Méthode travaillant sur un polytope: pour une fonction de n variables, on choisit n+1 points et on les fait évoluer par extension et contraction.
- Méthode qui ne nécessite pas le calcul du gradient ou du hessien, et donc très pratique sur les fonctions complexes.
- Les preuves de convergence sont peu claires
- L'algorithme est simple et pratique.

Le simplexe de Nelder-Mead

- Choix de n+1 points $(x_1,..x_{n+1})$
- Tri du simplexe: $f(x_1)<f(x_2)...<f(x_{n+1})$
- Barycentre de n points: $x_0 = (x_1+...+x_n)/n$
- Réfléchi de x_{n+1}/x_0: $x_r=x_0+(x_0-x_{n+1})$
- Si $f(x_r)<f(x_1)$, $x_e=x_0+2(x_0-x_{n+1})$. Si $f(x_e)<f(x_r)$, x_{n+1}<-x_e, sinon x_{n+1}<-x_r. Retour au tri.
- Si $f(x_n)<f(x_r)$, $x_c=x_{n+1}+(x_0-x_{n+1})/2$.Si $f(x_c)<f(x_r)$ x_{n+1}<-x_c, retour au tri
- Sinon: x_i <- $x_0+(x_i-x_1)/2$. Retour au tri.

Nelder Mead



Nelder-Mead Simplex search over Himmelblau function
6
4
2
0
-2
-4
-6
x_1
-6
-3
0
3
x_2
6
(c) P.A. Simionescu 2006

Méthodes *locales* (BFGS sur Griewank)