

2- Resolução de Sistemas Não-lineares.

2.1- Método de Newton.

2.2- Método da Iteração.

2.3- Método do Gradiente.

2- Sistemas Não Lineares de Equações

Considere um sistema não linear de n equações com n incógnitas.

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad \text{ou } \mathbf{f}(\mathbf{x}) = \mathbf{0} \quad \text{com } \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad \text{e } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Note que como caso particular podemos ter um sistema linear de equações algébricas:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - b_1 = 0 \\ f_2(x_1, x_2, \dots, x_n) = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n - b_2 = 0 \\ \dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n - b_n = 0 \end{cases} \quad \text{ou } \mathbf{Ax} - \mathbf{b} = \mathbf{0}$$

2.2- Método da Iteração

converge rapidamente se a norma de $\varphi'(\mathbf{x})$ é pequena. Logo é conveniente escolher a matriz Λ tal que

$$\varphi'(\mathbf{x}^0) = \mathbf{E} + \Lambda \mathbf{f}'(\mathbf{x}^0) = \mathbf{0}$$

e se a matriz $\mathbf{f}'(\mathbf{x}^0)$ é não singular segue que

$$\Lambda = -[\mathbf{f}'(\mathbf{x}^0)]^{-1} \Rightarrow \varphi(\mathbf{x}) = \mathbf{x} - [\mathbf{f}'(\mathbf{x}^0)]^{-1} \mathbf{f}(\mathbf{x}) \Rightarrow \mathbf{x} = \mathbf{x} - [\mathbf{f}'(\mathbf{x}^0)]^{-1} \mathbf{f}(\mathbf{x})$$

Aplicando o método da iteração a este sistema segue

$$\mathbf{x}^{k+1} = \mathbf{x}^k - [\mathbf{f}'(\mathbf{x}^0)]^{-1} \mathbf{f}(\mathbf{x}^k) \quad (4)$$

Note que (4) coincide com o **Método de Newton Modificado** para o sistema (3), já que $[\mathbf{f}'(\mathbf{x}^0)]^{-1} = \mathbf{W}(\mathbf{x}^0)^{-1}$ (Jacobiana). Ou seja,

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{W}(\mathbf{x}^0)^{-1} \mathbf{f}(\mathbf{x}^k) \quad (4')$$

Se a matriz $\mathbf{f}'(\mathbf{x}^0)$ é singular ($\det(\mathbf{f}'(\mathbf{x}^0)) = 0$), então deve ser escolhida uma aproximação inicial diferente de \mathbf{x}^0 .

2.2.1- Convergência do Método da Iteração

Teorema: Seja um sistema não linear de equações $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{x})$ com funções $\boldsymbol{\varphi}(\mathbf{x})$ e $\boldsymbol{\varphi}'(\mathbf{x})$ contínuas em Ω tais que em Ω se verifica a desigualdade $\|\boldsymbol{\varphi}'(\mathbf{x})\|_I \leq q < 1$ (q é uma constante). Se todas as aproximações sucessivas $\mathbf{x}^{k+1} = \boldsymbol{\varphi}(\mathbf{x}^k)$ pertencem a Ω , então este processo iterativo converge e seu limite é uma solução do sistema $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{x})$ em Ω .

Ou seja, $\lim_{k \rightarrow \infty} \mathbf{x}^k = \mathbf{x}^* = \boldsymbol{\varphi}(\mathbf{x}^*)$.

Aqui as normas são $\|\boldsymbol{\varphi}'(\mathbf{x})\|_I = \max_{\mathbf{x} \in \Omega} \|\boldsymbol{\varphi}'(\mathbf{x})\|_m$ e $\|\boldsymbol{\varphi}'(\mathbf{x})\|_m = \max_i \sum_{j=1}^n \left| \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j} \right|$

Corolário: O **Método da Iteração** converge se

$$\sum_{j=1}^n \left| \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j} \right| \leq q < 1 \quad \forall i = 1, \dots, n \text{ quando } \mathbf{x} \in \Omega.$$

Estimativa de erro na aproximação k para o **Método da Iteração**:

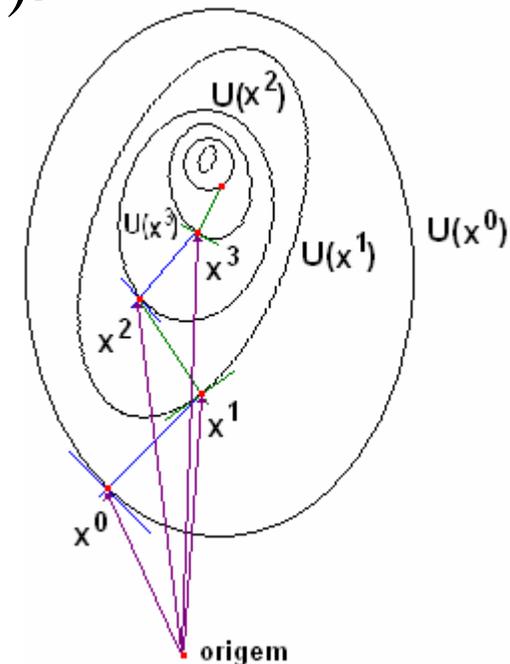
$$\|\mathbf{x}^* - \mathbf{x}^k\|_m \leq \frac{q^k}{1-q} \|\mathbf{x}^1 - \mathbf{x}^0\|_m = \frac{q^k}{1-q} \|\boldsymbol{\varphi}(\mathbf{x}^0) - \mathbf{x}^0\|_m$$

já que $\mathbf{x}^1 = \boldsymbol{\varphi}(\mathbf{x}^0)$

2.3 – Método do Gradiente (Descida mais Íngreme)

Seja \mathbf{x}^* a raiz de (3) e \mathbf{x}^0 a aproximação inicial. Construa uma superfície de nível para $U(\mathbf{x})$ em \mathbf{x}^0 . Se \mathbf{x}^0 está suficientemente perto de \mathbf{x}^* esta superfície de nível $U(\mathbf{x}) = U(\mathbf{x}^0)$ deve se assemelhar a um elipsóide. Parta de \mathbf{x}^0 na direção da normal da superfície $U(\mathbf{x}) = U(\mathbf{x}^0)$ até tocar alguma outras superfície de nível tal que $U(\mathbf{x}) = U(\mathbf{x}^1) < U(\mathbf{x}^0)$. Parta de \mathbf{x}^1 na direção da normal da superfície $U(\mathbf{x}) = U(\mathbf{x}^1)$ até tocar alguma outras superfície de nível tal que $U(\mathbf{x}) = U(\mathbf{x}^2) < U(\mathbf{x}^1)$.

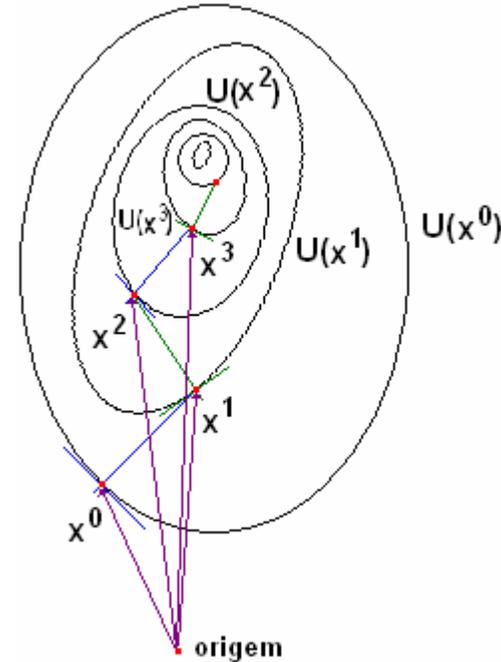
A repetição deste processo nos leva ao ponto de menor valor de $U(\mathbf{x})$, que corresponde à raiz do sistema (3).



2.3 – Método do Gradiente (Descida mais Íngreme)

Lembrando que o gradiente da função $U(\mathbf{x})$ tem a direção da normal \mathbf{n} da superfície de nível em cada ponto \mathbf{x}^k .

$$\nabla U(\mathbf{x}) = \begin{bmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \vdots \\ \frac{\partial U}{\partial x_n} \end{bmatrix}$$



Logo $\mathbf{x}^{k+1} = \mathbf{x}^k - \lambda^k \nabla U(\mathbf{x}^k)$ ($k = 0, 1, \dots$). Resta determinar λ^k para cada passo. Para isto considere a função que descreve a variação contínua da superfície de nível entre o ponto \mathbf{x}^k e \mathbf{x}^{k+1} .

$$\Phi(\lambda) = U(\mathbf{x}^k - \lambda \nabla U(\mathbf{x}^k))$$

2.3 – Método do Gradiente (Descida mais Íngreme)

O parâmetro $\lambda = \lambda^k$ deve ser escolhido para que $\Phi(\lambda)$ tenha um mínimo. Ou seja,

$$\frac{\partial \Phi(\lambda)}{\partial \lambda} = 0 = \frac{\partial U(\mathbf{x}^k - \lambda \nabla U(\mathbf{x}^k))}{\partial \lambda} = \frac{\partial \sum_{i=1}^n [f_i(\mathbf{x}^k - \lambda \nabla U(\mathbf{x}^k))]^2}{\partial \lambda} \quad (5)$$

Em geral, esta equação deve ser resolvida por métodos numéricos. Aqui apresentamos um método para determinar de forma aproximada o parâmetro λ^k . Para isto suponha que λ é pequeno e que potências deste parâmetro podem ser desprezadas se comparada com a parte linear. Logo expandindo as funções f_i em potências de λ e retendo apenas o termo linear segue

$$f_i(\mathbf{x}(\lambda)) = f_i(\mathbf{x}^k - \lambda \nabla U(\mathbf{x}^k)) \approx f_i(\mathbf{x}(0)) + \left. \frac{\partial f_i}{\partial \lambda} \right|_{\lambda=0} \lambda =$$

$$f_i(\mathbf{x}^k) + \lambda \frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \lambda} = f_i(\mathbf{x}^k) - \lambda \frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k)$$

2.3 – Método do Gradiente (Descida mais Íngreme)

Note que na equação anterior deve ser entendido por

$$\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} = \left[\frac{\partial f_i(\mathbf{x}^k)}{\partial x_1} \quad \frac{\partial f_i(\mathbf{x}^k)}{\partial x_2} \quad \dots \quad \frac{\partial f_i(\mathbf{x}^k)}{\partial x_n} \right] \text{ (um vetor)}$$

$$\text{Logo } \Phi(\lambda) = \sum_{i=1}^n \left(f_i(\mathbf{x}^k) - \lambda \frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right)^2 \text{ e}$$

$$\frac{\partial \Phi(\lambda)}{\partial \lambda} = 0 = -2 \sum_{i=1}^n \left(f_i(\mathbf{x}^k) - \lambda \frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right) \left(\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right)$$

ou

$$\lambda^k = \frac{\sum_{i=1}^n f_i(\mathbf{x}^k) \left(\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right)}{\sum_{i=1}^n \left(\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right)^2}$$

2.3 – Método do Gradiente (Descida mais Íngreme)

Como $U(\mathbf{x}) = \sum_{i=1}^n [f_i(\mathbf{x})]^2$ segue que

$$\nabla U(\mathbf{x}) = \begin{bmatrix} \frac{\partial U}{\partial x_1} \\ \frac{\partial U}{\partial x_2} \\ \vdots \\ \frac{\partial U}{\partial x_n} \end{bmatrix} = 2 \begin{bmatrix} \sum_{i=1}^n \frac{\partial f_i(\mathbf{x})}{\partial x_1} f_i(\mathbf{x}) \\ \sum_{i=1}^n \frac{\partial f_i(\mathbf{x})}{\partial x_2} f_i(\mathbf{x}) \\ \vdots \\ \sum_{i=1}^n \frac{\partial f_i(\mathbf{x})}{\partial x_n} f_i(\mathbf{x}) \end{bmatrix} = 2 \underbrace{\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}}_{[\mathbf{W}(\mathbf{x})]^T} \underbrace{\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}}_{\mathbf{f}(\mathbf{x})}$$

onde $\mathbf{W}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$ é a Jacobiana da função vetor $\mathbf{f}(\mathbf{x})$

2.3 – Método do Gradiente (Descida mais Íngreme)

Logo $\nabla U(\mathbf{x}^k) = 2[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)$ e

$$\sum_{i=1}^n f_i(\mathbf{x}^k) \left(\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right) = [\mathbf{f}(\mathbf{x}^k)]^T \mathbf{W}(\mathbf{x}^k) 2[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)$$

$$\sum_{i=1}^n \left(\frac{\partial f_i(\mathbf{x}^k)}{\partial \mathbf{x}} \nabla U(\mathbf{x}^k) \right)^2 = \left[2\mathbf{W}(\mathbf{x}^k)[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]^T \left[2\mathbf{W}(\mathbf{x}^k)[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]$$

Portanto

$$\lambda^k = \frac{[\mathbf{f}(\mathbf{x}^k)]^T \mathbf{W}(\mathbf{x}^k)[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)}{2 \left[\mathbf{W}(\mathbf{x}^k)[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]^T \left[\mathbf{W}(\mathbf{x}^k)[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]}$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mu^k [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \quad \text{onde} \quad \mu^k = 2\lambda^k \quad (6)$$

Note que o parâmetro μ^k deve ser calculado para cada passo.

2.3 – Método do Gradiente (Descida mais Íngreme)

Resumindo o método, já que $[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) = \left[[\mathbf{f}(\mathbf{x}^k)]^T \mathbf{W}(\mathbf{x}^k) \right]^T$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mu^k [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)$$

$$\mu^k = \frac{[\mathbf{f}(\mathbf{x}^k)]^T \mathbf{W}(\mathbf{x}^k) [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)}{\left[\mathbf{W}(\mathbf{x}^k) [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]^T \left[\mathbf{W}(\mathbf{x}^k) [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]}$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \frac{\left[[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]^T \left[[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right] \left[\mathbf{W}(\mathbf{x}^k) \right]^T \mathbf{f}(\mathbf{x}^k)}{\left[\mathbf{W}(\mathbf{x}^k) [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]^T \left[\mathbf{W}(\mathbf{x}^k) [\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k) \right]}$$

Observe que para otimizar o tempo de computo é conveniente para cada passo fazer os cálculos na seguinte ordem:

- 1- Calcule $\mathbf{f}(\mathbf{x}^k)$, $\mathbf{W}(\mathbf{x}^k)$ ou $[\mathbf{W}(\mathbf{x}^k)]^T$
- 2- Calcule $[\mathbf{W}(\mathbf{x}^k)]^T \mathbf{f}(\mathbf{x}^k)$ e termine as outras operações

2.1.3- Exemplo 2 pag. 462 do Demidovich

Use o Método da Iteração e do Gradiente para encontrar a solução aproximada positiva

do sistema de equações:

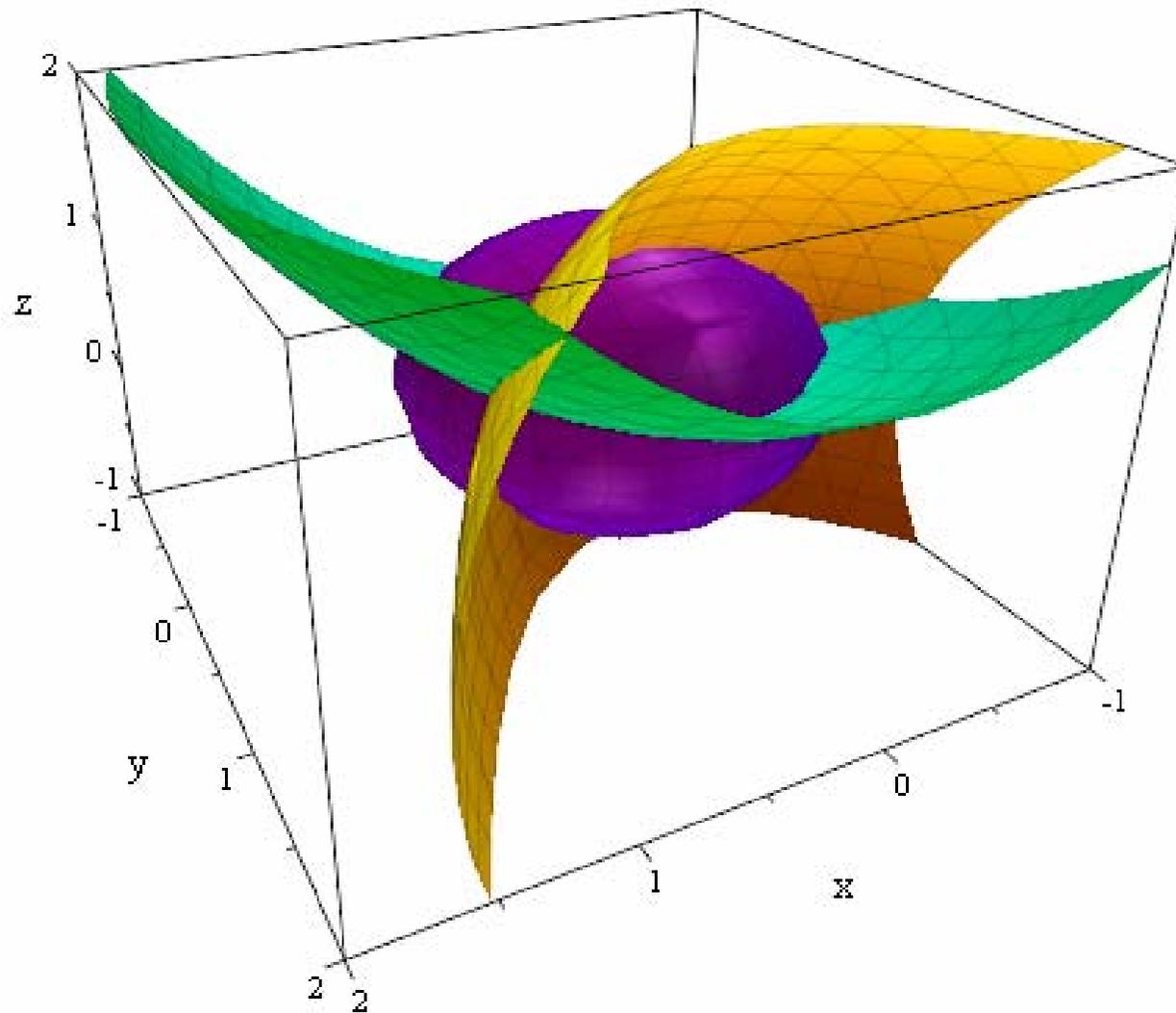
$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ 2x^2 + y^2 - 4z = 0 \\ 3x^2 - 4y + z^2 = 0 \end{cases}$$

começando com a aproximação inicial $x_0 = y_0 = z_0 = 0.5$.

Visualização gráfica do problema na próxima tela.

2.1.3- Exemplo 2 pag. 462 do Demidovich

Visualização gráfica do problema



Frases do Dia

“... the progress of physics will to a large extent depend on the progress of nonlinear mathematics, of method to solve nonlinear equations ... and therefore we can learn by comparing different nonlinear problems.”

Werner Heisenberg