

## 5. Wprowadzenie do problemu odkształceń na przykładzie samorosnących kropek kwantowych

### *Samorosnące kropki kwantowe*

Jedną z najważniejszych metod wytwarzania nanostruktur jest tzw. epitaksja z wiązek molekularnych (MBE, ang. molecular beam epitaxy). W metodzie tej, w warunkach ultra-wysokiej próżni, na odpowiednio przygotowane podłoże krystaliczne, nanoszone są kolejne warstwy atomów lub molekuł (Rys.5.1). Źródłem strumienia atomów różnych pierwiastków lub bądź molekuł (stąd nazwa MBE) są specjalne piecyki o stabilizowanej temperaturze, tzw. komórki efuzyjne (zwane też komórkami Knudsen). Proces MBE w połączeniu z, na przykład, spektroskopią RHEED (dyfrakcja wysokoenergetycznych elektronów, ang. Reflection High-Energy Electron Diffraction) umożliwia kontrolę wzrostu kolejnych warstw z dokładności do pojedynczych odległości międzyatomowych.

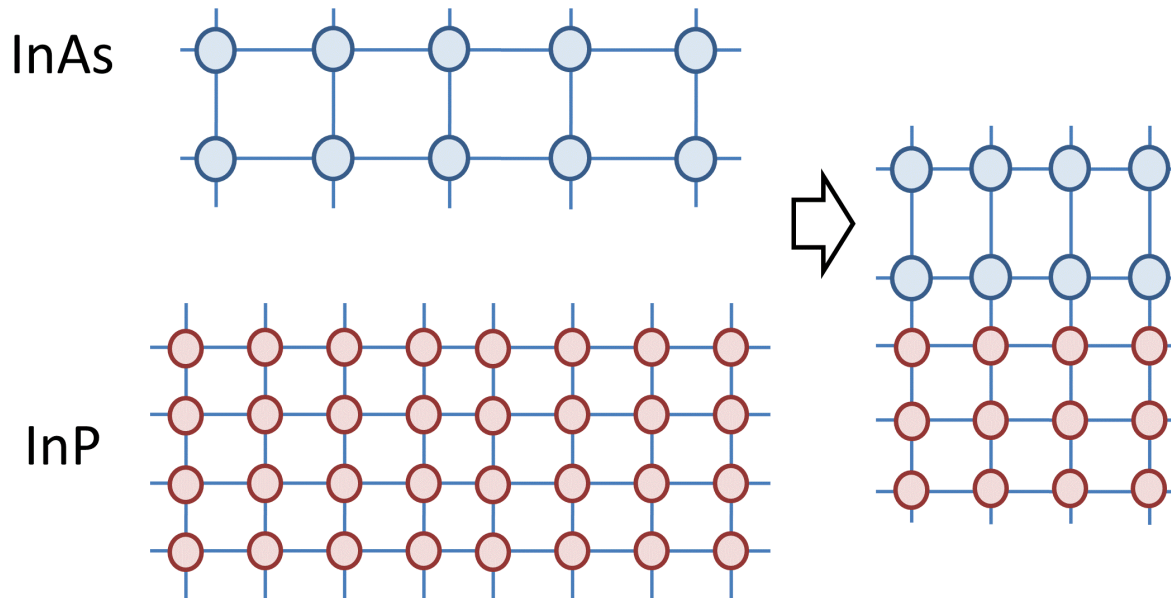


*Rys.5.1 Wygląd zewnętrzny komory MBE (źródło: Internet).*

W procesie wzrostu można zmienić skład chemiczny reagentów, np. poprzez zamknięcie i otwarcie odpowiednich komórek efuzyjnych. I tak, przykładowo, po osadzeniu kilkunastu monowarstw (ang. monolayers) arsenku galu (GaAs) można nanieść kolejne warstwy innego materiału, np. arsenku aluminium (AlAs). Jako, że materiały osadzane oraz podłoża różnią się własnościami elektronowymi, w szczególności szerokością przerwy energii wzbronionej, to opisany proces zmiany składników umożliwia wytwarzanie studni kwantowych oraz tzw. supersieci.

Nie wszystkie materiały półprzewodnikowe można ze sobą łączyć, tak aby obszar złącza wolny był od dyslokacji oraz innych defektów struktury krystalicznej. W szczególności stałe sieci krystalicznych (związane z długością wiązań międzyatomowych) materiałów podłoża oraz osadzanego nie mogą zbyt różnić się zbyt różnić. Dobrym przykładem są tutaj GaAs oraz AlAs, których stałe sieci różnią się o mniej niż 0.2%.

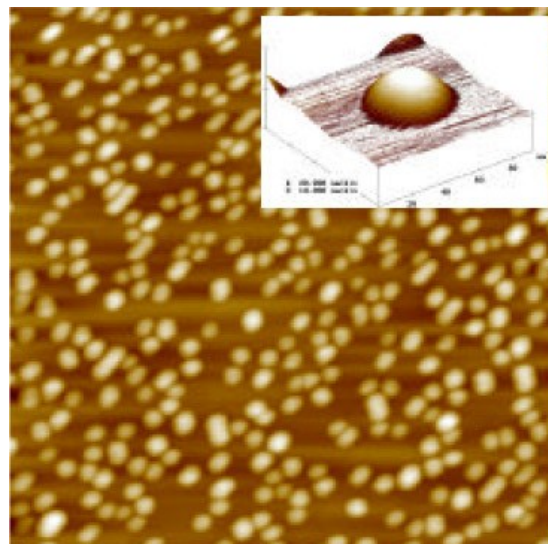
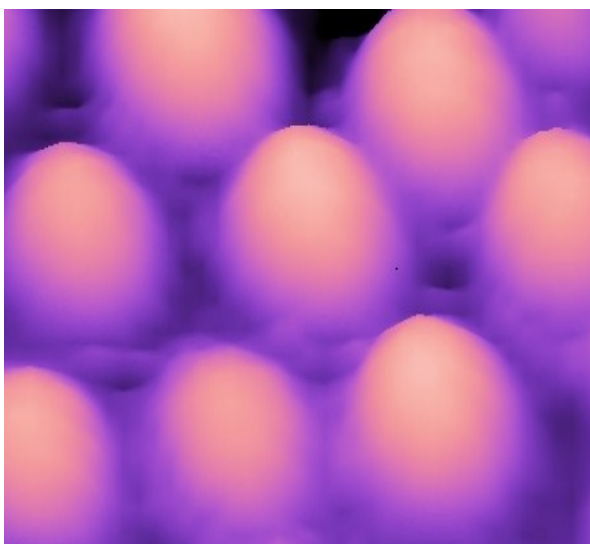
Bardzo ciekawa sytuacja ma miejsce w przypadku często stosowanych kombinacji materiałów półprzewodnikowych np. dla InAs/InP różnica stałych sieci wynosi około 3%, natomiast dla InAs/GaAs jest to aż 7%. Ze względu na wspomnianą różnicę stałych sieci, nanoszona warstwa oraz podłoże muszą się w jakiś sposób „dopasować” się (rys. 5.2).



*Rys.5.2 Schemat ilustrujący dopasowanie stałej sieci epitaksjalnie nanoszonej warstwy do stałej sieci podłoża.*

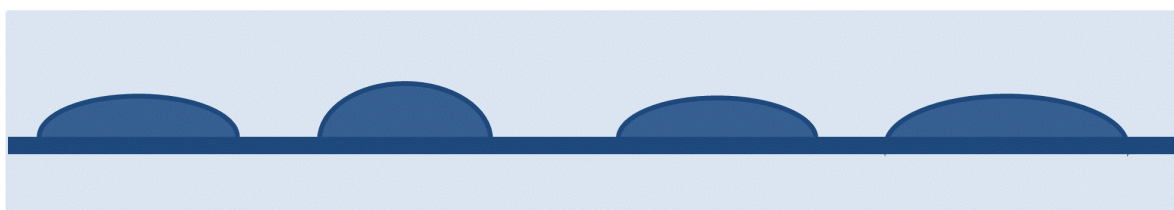
Dla materiałów różniących się istotnie stałą sieci, w procesie tzw. wzrostu Stranskiego-Krastanowa, pierwsza z osadzanych warstw (tzw. warstwa zwilżająca) dopasowuje swoją stałą sieci w płaszczyźnie złącza do stałej materiału podłoża (jest to tzw. pseudomorfizm). Dla materiałów widocznych na rysunku schematycznym oznacza to, że materiał warstwy (InAs) o większej stałej sieci krystalicznej, jest ściśnięty w płaszczyźnie złącza, a towarzyszy temu jednocześnie zwiększenie grubości warstwy (rozciągnięcie) w kierunku wzrostu, podczas gdy podłoże pozostaje praktycznie nie odkształcone. W takiej sytuacji mówimy, że warstwa ulega silnej deformacji lub odkształceniu, gdyż zmianie ulegają odległości międzyatomowe w stosunku do odległości charakterystycznych dla nieodkształconego kryształu objętościowego. Bardziej ilościowy opis tego zjawiska znajdziemy w dalszej części wykładu.

Kontynuacja procesu jednorodnego wzrostu materiałów niedopasowanych sieciowo doprowadziłaby do silnego odkształcenia nowo rosnącej warstwy. W rzeczywistości bardziej energetycznie korzystna jest sytuacja w której energia odkształceń minimalizowana jest w procesie wzrostu poprzez powstanie niejednorodnych zgrubień „wysp” nanoszonego materiału (Rys 5.3). Powstałe „wyspy” mają rozmiary rzędu nanometrów (średnica 10 -30 nanometrów; wysokość 1-5 nanometrów) oraz rozmieszczone są w nieregularny, przypadkowy sposób na powierzchni podłoża. Charakteryzują się jednak dość podobnym kształtem przypominającym soczewkę, a ich parametry takie jak skład chemiczny, rozmiar, czy gęstość rozmieszczenia na podłożu mogą być w dużym stopniu kontrolowalne poprzez dobór warunków wzrostu.



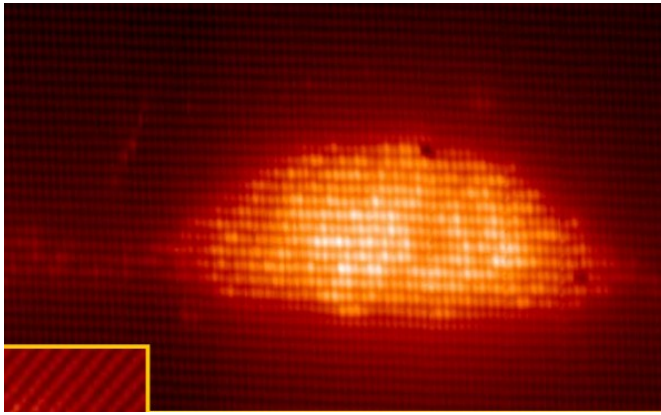
*Rys.5.3 Obrazy z mikroskopu sił atomowych ilustrujące powstawanie „wysp” (kropek kwantowych) na podłożu o innej stałej sieci (źródło: Internet).*

Zwykle proces wzrostu „wysp” jest przerywany, natomiast ponownie nanoszony jest materiał podłoża. W rezultacie otrzymuje się strukturę zbudowaną z „wtrąceń” materiału półprzewodnikowego otoczonego przez materiał podłoża (Rys. 5.4 oraz Rys. 5.5).



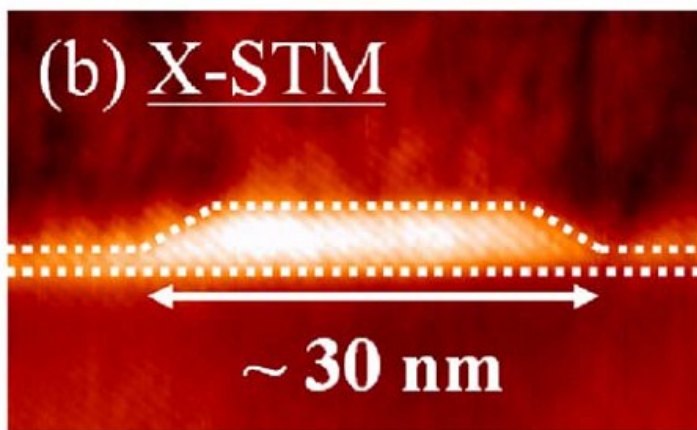
*Rys.5.4 Schematyczny przekrój przez warstwę samorosnących kropek kwantowych.*

Odpowiedni dobór składu chemicznego, np. otaczającej matrycy o szerokiej przerwie energetycznej (GaAs lub InP) oraz wąsko-przerwowego materiału „wysp” (InAs) prowadzi do przestrzennego ograniczenia ruchu nośników ładunku (elektronów i dziur) do obszaru danej „wyspy”. Tego rodzaju nanostruktury nazywamy samorosnącymi kropkami kwantowymi (ang. self-assembled quantum dot), gdzie „samorosnący” lub „samoorganizujący” ilustruje charakter procesu wzrostu, a słowo „kropka kwantowa” opisuje niewielki obszar przestrzeni ograniczony w trzech wymiarach barierami potencjału.



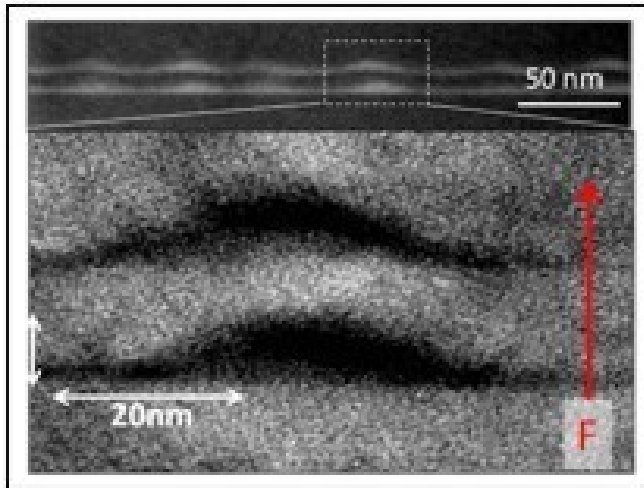
Rys.5.5 Obraz ze skaningowego mikroskopu tunelowego, przekrój (X-STM, ang. cross-section scanning tunneling microscope) przez pojedynczą samorosnącą kropkę kwantową InAs/GaAs o kształcie soczewki (źródło: prace P.M. Koenraad et al).

Kontrola parametrów wzrostu umożliwia także uzyskiwanie nanostruktur o innych niż soczewka kształtach. W szczególności technika „indium-flush” pozwala na tworzenie samorosnących kropek kwantowych o kształcie dysku (Rys. 5.6).

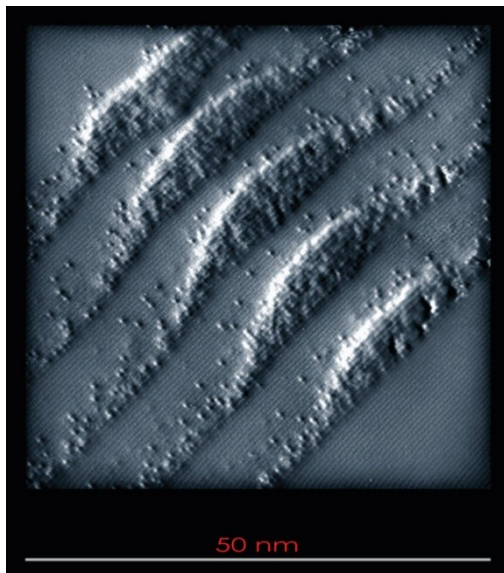


Rys.5.6 Obraz ze skaningowego mikroskopu tunelowego, przekrój (X-STM) przez pojedynczą samorosnącą kropkę kwantową InAs/GaAs o kształcie dysku (źródło: prace P.M. Koenraad et al).

Obok pojedynczych kropek kwantowych możliwe jest również stworzenie układów podwójnych kropek kwantowych (Rys. 5.7) lub nawet stosu kropek kwantowych (Rys. 5.8).



Rys.5.7 Obraz ze transmisyjnego mikroskopu elektronowego (TEM, ang. transmission electron microscopy), przekrój przez układ dwóch samorosnących kropek kwantowych InAs/GaAs (źródło: Internet).



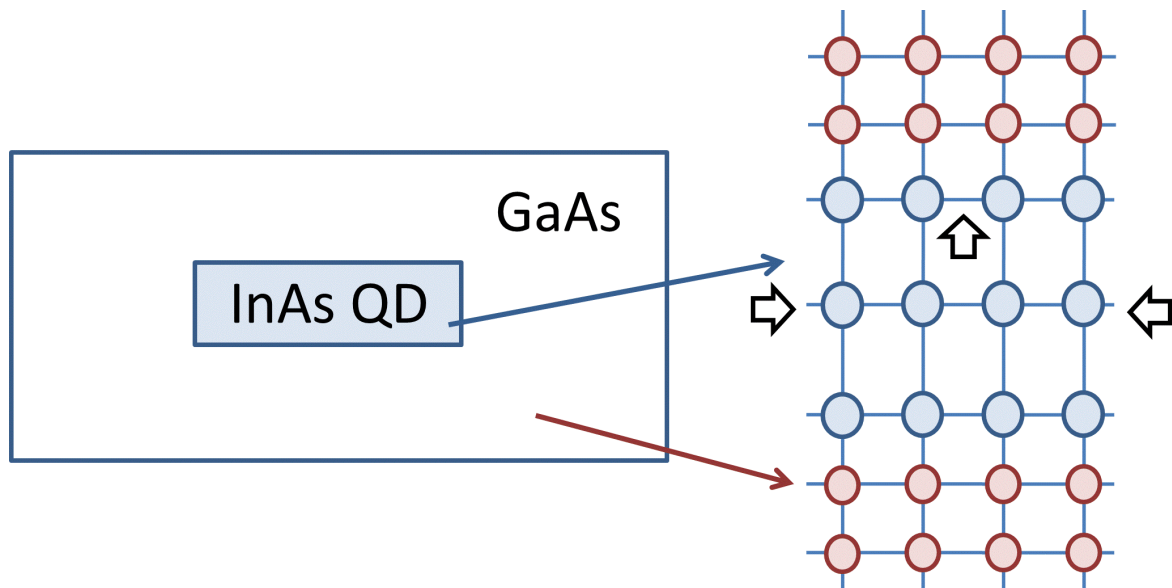
Rys.5.8 Obraz ze skaningowego mikroskopu tunelowego, przekrój (X-STM) przez pionowy stos samorosnących kropek kwantowych InAs/GaAs (źródło: Internet).

Własności widmowe pojedynczych kropek kwantowych przypominają własności atomów (dyskretne widmo energii), stąd czasem stosowana jest nazwa „sztuczne atomy”. Analogicznie, układy dwóch i więcej kropek kwantowych nazywa się często mianem „sztucznych molekuł”.

## 6. Rodzaje i źródła odkształceń

### Tensorowy opis odkształceń

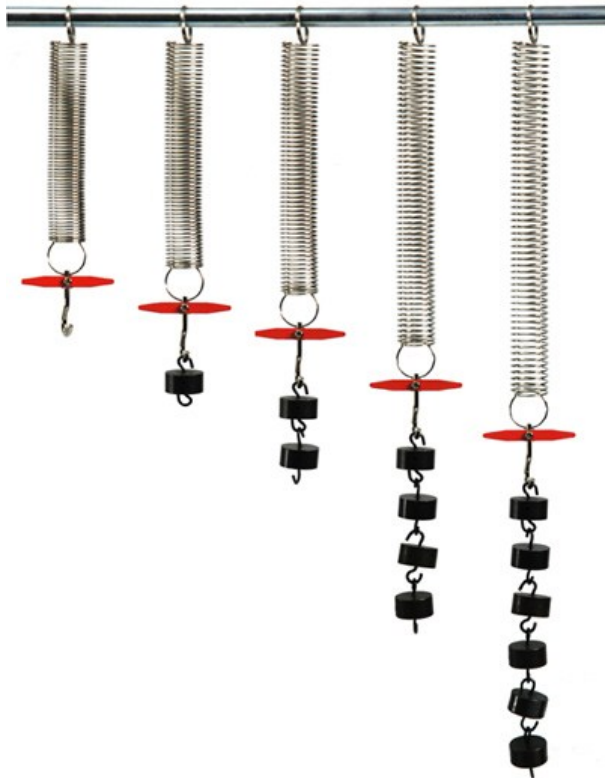
Omówione do tej pory samorosnące kropki kwantowe zawdzięczają swoje istnienie niedopasowaniu sieci krystalicznych materiału kropki i podłoża i związanym z tym efektom odkształceń. Z drugiej strony materiał kropki jest silnie odkształcony (Rys. 5.9), i jak należy przypuszczać efekty odkształceń mogą mieć istotny wpływ na własności widmowe danej kropki kwantowej.



Rys.5.9 Schemat deformacji odległości międzyatomowych dla kropki kwantowej InAs w matrycy GaAs.

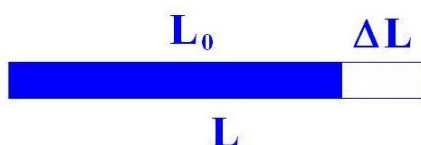
Samorosnąca kropka kwantowy stanowi dość skomplikowany, trójwymiarowy układ, w pierwszej kolejności zastanówmy się więc jak ilościowo opisać deformacje (odkształcenia) dla zdecydowanie prostszego zagadnienia jednowymiarowego.

Zacznijmy od przykładu metalowej sprężyny na której zawieszamy coraz większe obciążenie (Rys. 5.10). Zgodnie z prawem Hooke'a oraz naszą intuicją stopień rozciągnięcia sprężyny zależy od jej rodzaju (np. grubości oraz materiału drutu z którego jej wykonana) oraz zawieszanej masy.



Rys.5.10 Deformacja (rozciąganie) sprężyny w funkcji rosnącego obciążenia.

Na użytek aktualnych rozważań nie interesują nas jednak ani siły związane z tym procesem, ani naprężenia występujące w drucie sprężyny, a jedynie jakaś miara jej deformacji. Wprowadzimy więc definicję odkształcenia sprężyny poprzez porównanie długości sprężyny zdeformowanej i nieodkształconej (Rys. 5.11) oraz podzielenie tej różnicy przez długość nieodkształconą.

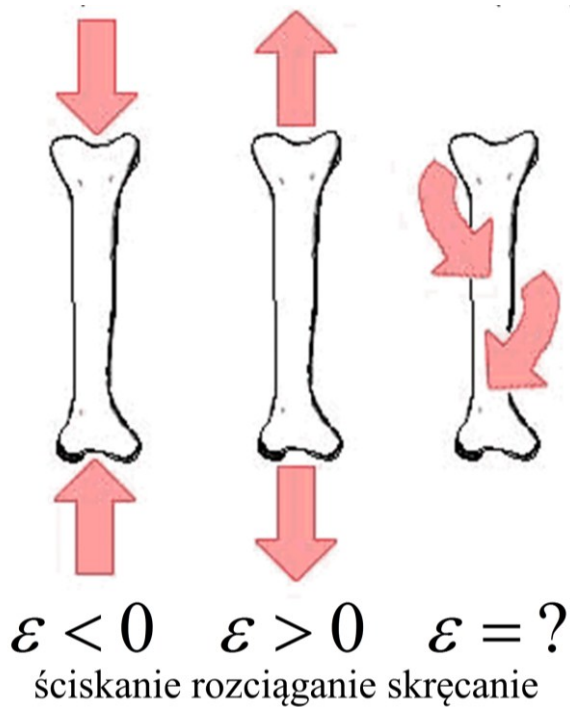


$$\epsilon = \frac{L - L_0}{L_0} = \frac{\Delta L}{L_0}$$

Rys.5.11 Ilustracja i definicja odkształcenia dla przypadku jednowymiarowego (źródło: Internet).

Definicja ta ma taką zaletę, że dla struktury nieodkształconej  $\epsilon$  przyjmuje wartość równą zero. Znak  $\epsilon$  opisuje charakter odkształceń:  $\epsilon > 0$  odpowiada odkształceniom o charakterze rozciągającym, natomiast  $\epsilon < 0$  odpowiada odkształceniom o charakterze ściskającym. Dodatkowo dzielenie przez długość sprężyny nieodkształconej pozwala wprowadzić jednakową miarę odkształceń dla obiektów różnej długości. Przykładowo metrowy drut rozciągnięty o 1 centymetr jest tak samo silnie odkształcony (1%), jak kilometrowy drut rozciągnięty o dziesięć metrów.

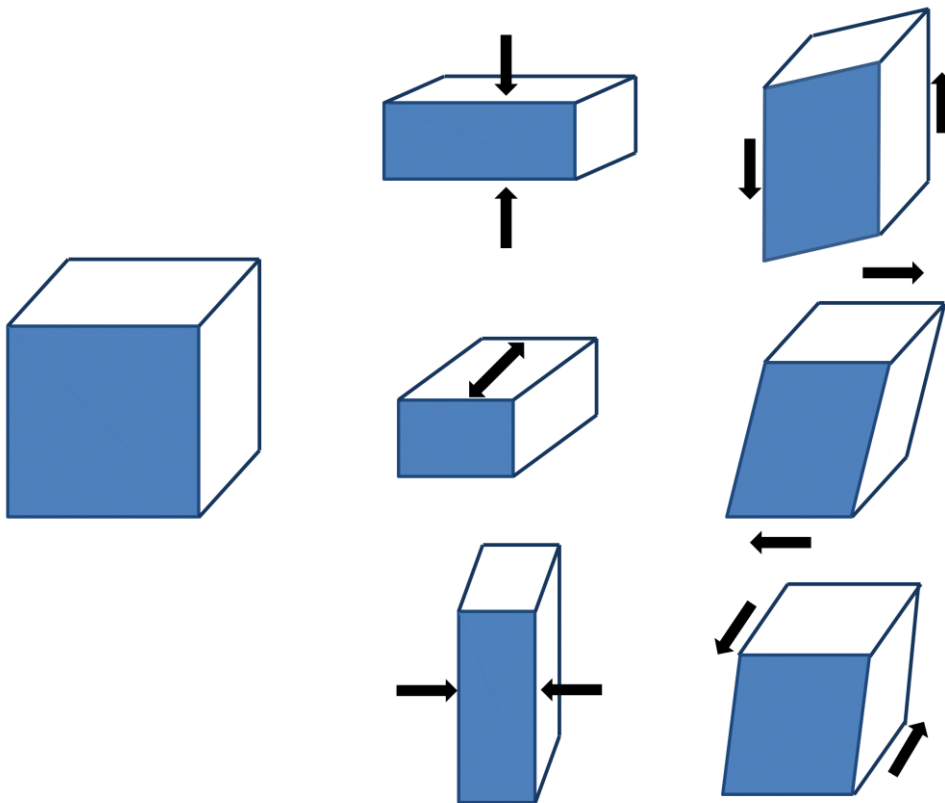
Definicja  $\epsilon$  jest w pełni wystarczająca dla deformacji o charakterze jednowymiarowym. Problem pojawia się jednak dla bardziej skomplikowanych deformacji przestrzennych (Rys. 5.11).



Rys.5.11 Ilustracja znaku odkształcenia dla przypadku jednowymiarowego oraz „kłopotu” z prostym opisem jednowymiarowym dla bardziej skomplikowanych deformacji o charakterze skręcającym.

W ogólnym przypadku w opisie należy uwzględnić istnienie pozostałych dwóch wymiarów. Ilustruje to rysunek Rys. 5.12, który pokazuje wiele (często skomplikowanych) możliwości w jakie można zdeformować jednostkowy sześcian. Możliwe jest więc rozciąganie i ściskanie wzdłuż określonych krawędzi sześcianu, a co więcej deformacja o charakterze skręcającym, bądź ścinającym.





Rys.5.12 Schemat różnych odkształceń dla trójwymiarowego sześcianu.

Wszystkie możliwe deformacje sześcianu można opisać za pomocą jednej wielkości macierzowej, tzw. tensora odkształceń:

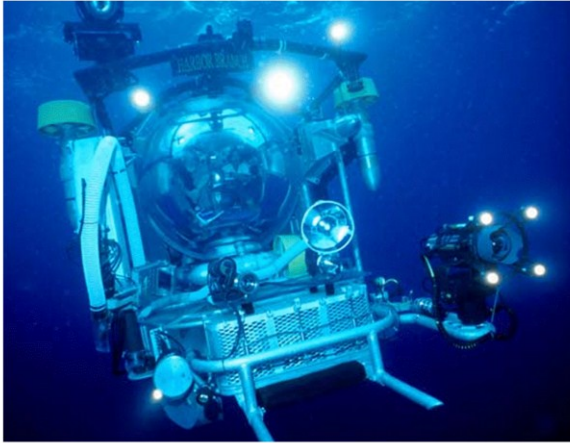
$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix}$$

Tensor ten jest symetryczny ( $\varepsilon_{ij}=\varepsilon_{ji}$ ) oraz, co ważne, opisuje deformacje kształtu jednostkowego sześcianu, a nie jego obrót (wyrażony za pomocą macierzy obrotów) czy translacje (przesunięcia) w przestrzeni.

#### Rodzaje odkształceń

Bez wchodzenia w szczegóły matematyczne zilustrujemy (Rys. 5.13) główne rodzaje odkształceń oraz odpowiadające im postaci tensora odkształceń. W szczególności deformacja o charakterze hydrostatycznym (której doświadcza np. kabina łodzi podwodnej) oznacza jednorodne ściskanie ( $\varepsilon < 0$ ), zatem żaden z kierunków przestrzennych nie jest wyróżniony ( $\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon_{zz}$ ). Ponieważ kształt deformowanej struktury jest zachowany to elementy niediagonalne tensora są dokładnie równe zero.

Z kolei dla odkształceń o charakterze osiowym, ściskaniu (rozciąganiu) w jednym z kierunków (np.  $\epsilon_{zz} < 0$ ) towarzyszy rozciąganie (ściskanie) w dwóch pozostałych kierunkach ( $\epsilon_{xx} = \epsilon_{yy} > 0$ ). Wybierając osie układu współrzędnych wzdłuż kierunku deformacji ponownie otrzymujemy diagonalną postać tensora odkształceń, przy czym elementy na diagonalu są różne (np.  $\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}$ ).



### Hydrostatyczne

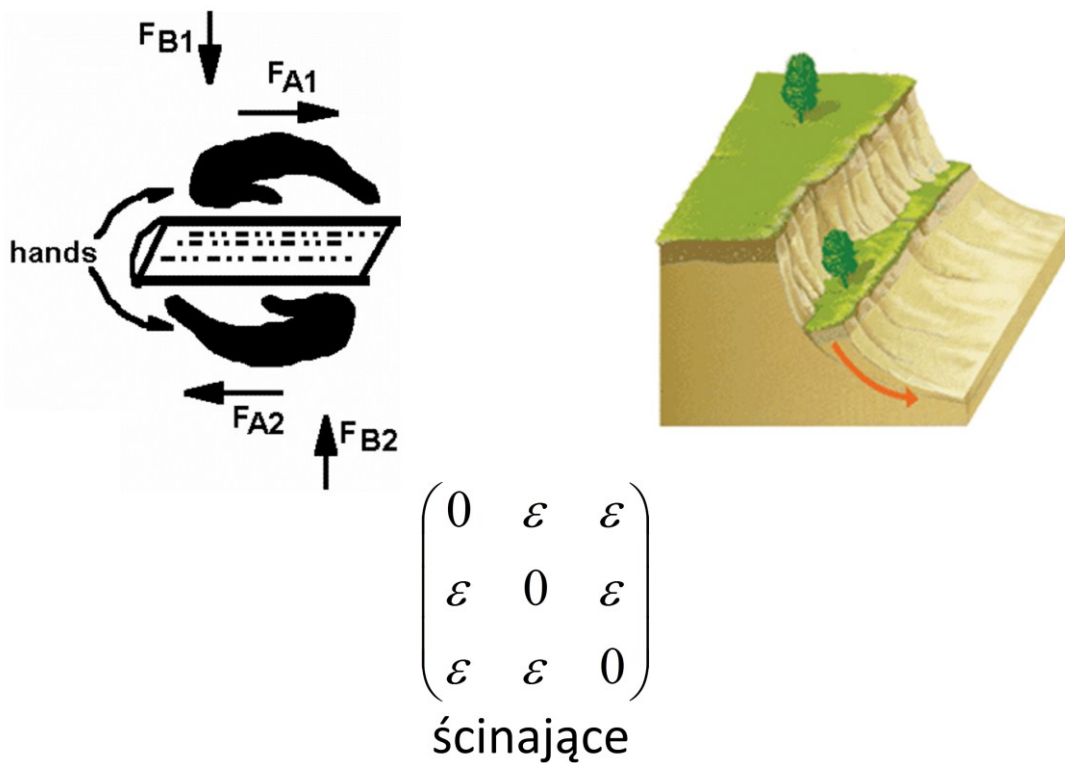
$$\begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{pmatrix}$$

### Osiowe

$$\begin{pmatrix} \epsilon_{\leftrightarrow} & 0 & 0 \\ 0 & \epsilon_{\leftrightarrow} & 0 \\ 0 & 0 & \epsilon_{\downarrow} \end{pmatrix}$$

Rys.5.13 Przykład odkształceń o charakterze hydrostatycznym i osiowym oraz odpowiadających im postaci tensor odkształceń (Źródło: Internet).

Jeszcze innym rodzajem odkształceń są te o charakterze ścinającym (ang. shear strain), zilustrowane na Rys. 5.14, w takim przypadku zmianie nie ulegają liniowe wymiary struktury, a odkształcenie ciała spowodowane przyłożeniem siły stycznej do jej powierzchni. Należy w tym miejscu dodać, że trzy wspomniane wyżej rodzaje deformacji mogą występować jednocześnie. Na przykład w przypadku „typowych” samorosnących kropek kwantowych, dominująca rolę odgrywają odkształcenia o zarówno hydrostatycznym, jak i osiowym, natomiast te o charakterze ścinającym są zdecydowanie mniej istotne.



Rys.5.14 Przykład odkształceń o ścinającym (Źródło: Internet).

Z elementów macierzy tensora odkształceń można skonstruować użyteczne wielkości pomocnicze będące miarą określonych typów deformacji. W szczególności ślad tensora odkształceń (suma elementów diagonalnych macierzy) jest ilościową miarą odkształceń hydrostatycznych:

$$Tr(\varepsilon) = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$$

Dla niewielkich deformacji  $Tr(\varepsilon)$  odpisuje zmianę objętości jednostkowego sześcianu:

$$V \approx V_0 [1 + Tr(\varepsilon)]$$

Ujemne wartości  $Tr(\varepsilon)$  opisują zatem odkształcenia o charakterze ściskającym, a dodatnie o charakterze rozciągającym. Użyteczności  $Tr(\varepsilon)$  jest tym większa, że jest on niezmiennikiem, a więc wielkością niezależną od wyboru układu współrzędnych.

Miarą odkształceń osiowych może być np. wielkość zdefiniowana następująco:

$$B(\varepsilon) = \sqrt{(\varepsilon_{xx} - \varepsilon_{yy})^2 + (\varepsilon_{yy} - \varepsilon_{zz})^2 + (\varepsilon_{zz} - \varepsilon_{xx})^2}$$

Dla oddziaływań o charakterze czysto hydrostatycznym ( $\epsilon_{xx}=\epsilon_{yy}=\epsilon_{zz}$ ) mamy zatem  $B(\epsilon)=0$ , gdyż wszystkie wymiary liniowe odkształcone się w ten sam sposób. Z kolei dla odkształceń o charakterze czysto osiowym jest  $\text{Tr}(\epsilon)=0$ , przy czym  $B(\epsilon)\neq 0$ , co odpowiada deformacji kształtu zachowującej objętość. Przypomnijmy raz jeszcze, że w ogólności możliwe są odkształcenia będą kombinacją wszystkich wspomnianych podstawowych typów odkształceń.

### Odształcenia a naprężenia.

W naszych rozważaniach dotyczących kropek kwantowych opisaliśmy już jeden przykład: odkształcenia pochodzące od niedopasowania stałych sieci krystalicznych dwóch materiałów półprzewodnikowych. Odształcenia występują jednak w wielu innych układach, w szczególności makroskopowych m.in. konstrukcjach budowlanych, urządzeniach technicznych itp. Należy też doprecyzować, że mówiąc od odkształceniach odnosimy się do odkształceń sprężystych, a więc takich, które ustępuje po usunięciu siły, która je spowodowała. W tym wykładzie nie interesują nas odkształcenia plastyczne charakterystyczne dla takich materiałów jak np. plastelina czy glina.

Należy w tym miejscu rozróżnić także „odkształcenia” od „naprężeń”. Te pierwsze opisują deformację danego ciała w stosunku do przypadku niedokształconego, przy czym  $\epsilon$  jest wielkością bezwymiarową. Naprężenia (oznaczenie  $\sigma$ ) są z kolei pewną miarą sił związanych z tą deformacją, a ich jednostką jest paskal ( $\text{N/m}^2$ , jednostka siły na jednostkę powierzchni). Przykładowo intuicja podpowiada nam, że 1% odkształcenie stalowego drutu o przekroju  $1\text{mm}^2$  wymaga przyłożenia zdecydowanie większej siły niż gumowego paska o identycznym przekroju.

Relacja pomiędzy dwoma wspomnianym wyżej wielkościami jest niczym innym jak prawem Hooke’a:

$$\sigma = E \epsilon$$

gdzie współczynnik proporcjonalności  $E$  jest parametrem materiałowym zwanym modułem sprężystości (sztywności) podłużnej  $E$  lub krócej modułem Younga. Wielkość ta przyjmuje wartości np. 210 GPa dla stali i 0.1 GPa dla gumy, co dobrze ilustruje fundamentalnie różną sprężystość obu materiałów.

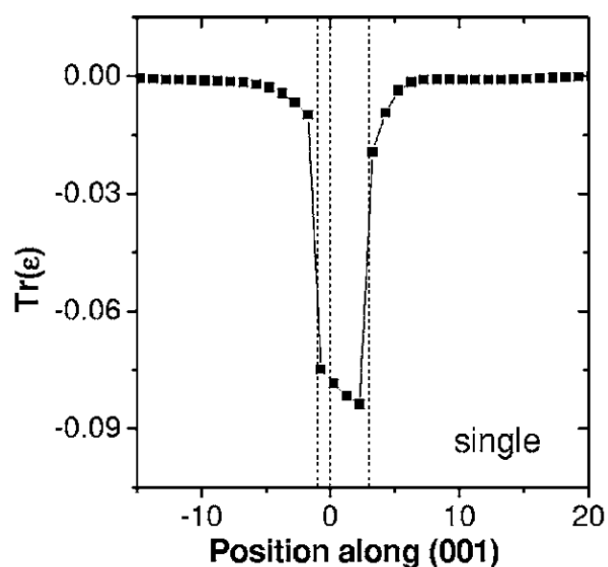
Podobnie jak odkształcenia, naprężenia są w ogólnym przypadku (dla materiałów anizotropowych) opisane wielkością tensorową ( $\sigma_{ij}$ ), a relacja między obiema wielkościami zwana jest uogólnionym prawem Hooke’a:

$$\sigma_{kl} = \sum_{ij} c_{klij} \epsilon_{ij}$$

Ponieważ naprężenia oraz odkształcenia są w tensorami drugiego rzędu, związek pomiędzy nimi jest zadany tensorem czwartego rzędu,  $c_{ijkl}$ , zwanym tensorem sztywności, którego składowe zwane są modułami sprężystości lub współczynnikami sztywności sprężyste. Tensor sztywności ma w ogólności  $3^4=81$  składowych, jednak symetria sieci krystalicznej istotnie redukuje liczbę niezależnych składowych i tak dla struktury o symetrii regularnej (kubicznej) mamy jedynie 3 niezerowe składowe  $c_{ijkl}$ , zwyczajowo oznaczane jako  $C_{11}$ ,  $C_{12}$  oraz  $C_{44}$ , przy czym zastosowano tutaj notację w której indeksy 1, 2 oraz 4 odpowiadają parom wskaźników: 11, 22, 23.

### Przybliżenie ośrodka ciągłego oraz podejście atomistyczne.

Do tej pory omawialiśmy deformacje obiektu (sieci krystalicznej) jako całości. Na wstępie opisywaliśmy jednak przykład samorosnącej kropki kwantowej oraz zasugerowaliśmy, że kropka kwantowa jest silnie odkształcona, podczas, gdy podłoże krystaliczne jest zdeformowane w mniejszym stopniu. Oznacza to, że dla w ogólnym przypadku deformacja, a więc i składowe tensora odkształceń, jest funkcją położenia. Jest to widoczne na Rys. 5.15, gdzie wewnątrz kropki kwantowej doznaje silnych odkształceń o charakterze ściskającym ( $\text{Tr}(\epsilon) \approx -0.08$ ), podczas, gdy obszar w odległości 10 stałych sieci od kropki kwantowej pozostaje praktycznie niezdeformowany.



Rys.5.15 Ślad tensora odkształceń  $\text{Tr}(\epsilon)$  w funkcji położenia atomów; przekrój wzdłuż kierunku wzrostu przez geometryczny środek kropki kwantowej InAs/GaAs o kształcie soczewki (Źródło: *Phys. Rev. B* **74**, 195339 (2006)).

Jednym ze sposobów wyznaczenia przestrzennego rozkładu odkształceń jest wykonanie obliczeń w tzw. przybliżeniu ośrodka ciągłego (ang. continuum elasticity), w którym zaniedbujemy atomową budowę materii, to znaczy ośrodek dzielimy na bardzo małe części, ale o takich samych własnościach jak cała badana objętość.

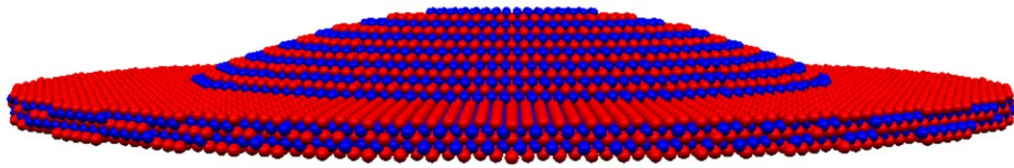
W interesującym nas przypadku definiujemy wartości energii sprężystej (odkształceń) w każdym punkcie dyskretnej, przestrzennej siatki. Przykładowo dla struktury o symetrii regularnej:

$$E_{CE}(x, y, z) = \frac{V}{2} \left[ C_{11}(x, y, z)(\epsilon_{xx}^2 + \epsilon_{yy}^2 + \epsilon_{zz}^2) + C_{44}(x, y, z)(\epsilon_{yz}^2 + \epsilon_{zx}^2 + \epsilon_{xy}^2) + 2C_{12}(x, y, z)(\epsilon_{xx}\epsilon_{yy} + \epsilon_{zz}\epsilon_{xx} + \epsilon_{xx}\epsilon_{yy}) \right],$$

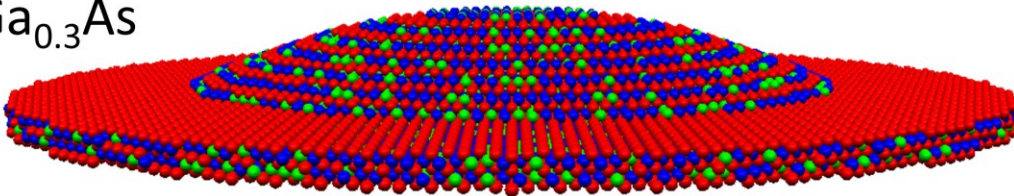
gdzie  $C_{11}$ ,  $C_{12}$  oraz  $C_{44}$  to zdefiniowane wcześniej zależne od położenia (zatem inne dla kropki kwantowej oraz inne dla otaczającej matrycy) współczynniki sztywności sprężystej, a  $V$  to objętość związana z dyskretyzacją siatki (objętość „oczka” siatki). Wartości  $\epsilon$  w każdym węzle siatki są znajdowane w ten sposób, aby energia całkowita układu, tj. suma  $E_{CE}$  po wszystkich punktach siatki, przyjmowała minimum.

W podejściu tym widoczny jest jednak pewien fundamentalny problem. Po pierwsze krok siatki (a więc i objętość  $V$ ) musi być tak dobrany, aby zminimalizować błędy dyskretyzacji. Z drugiej strony, nawet jeśli krok siatki byłby bardzo mały, np. porównywalny ze stałą sieci krystalicznej, to i tak symetria dyskretnej siatki obliczeniowej jest w ogólności inna od symetrii sieci krystalicznej. W szczególności, dla materiałów takich jak InAs czy InP, może to być całkiem skomplikowana symetria typu blendy cynkowej lub wurcytu. Co więcej, jak widzieliśmy wcześniej na obrazach X-STM kropek kwantowych, atomy w kropce kwantowej i na jej brzegu mogą być umieszczone w sposób losowy. Fluktuacje składu chemicznego np. dla kropek kwantowych o mieszanym składzie chemicznym ( $\text{In}_x\text{Ga}_{1-x}\text{As}$ , Rys. 5.16) odgrywają istotną rolę i nie mogą być w sposób prosty modelowane w przybliżeniu ośrodka ciągłego.

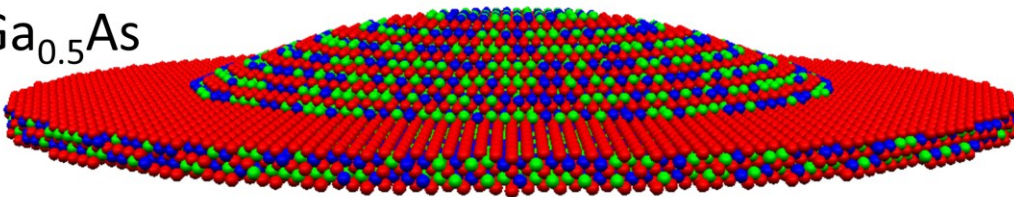
InAs



$\text{In}_{0.7}\text{Ga}_{0.3}\text{As}$



$\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$



Rys.5.16 Położenia atomów arsenu (kolor czerwony), indu (niebieski) oraz galu (zielony) dla warstwy zwilżającej oraz kropki kwantowej  $\text{In}_x\text{Ga}_{1-x}\text{As}$  o kształcie soczewki oraz różnych wartości składu chemicznego. Otaczająca kropkę kwantową matryca GaAs została na rysunku pominięta.

Sposobem na opisane wyżej problemy jest uwzględnienie w obliczeniach atomowego, „ziarnistego” charakteru materii. W odróżnieniu od metod wykorzystujących przybliżenie ośrodka ciągłego tego rodzaju metody noszą ogólne miano „atomistycznych”.

Przykładem takiej właśnie metody jest VFF (ang. Valence Force Field), w której energia sprężystości zdefiniowana jest jako funkcja położenia atomów, a w szczególności długości wiązań międzyatomowych oraz kątów między wiązaniami. W interesującym nas przypadku kryształów o mikroskopowej symetrii tetraedralnej mamy:

$$E_{VFF} = \frac{1}{2} \sum_i^N \sum_{j=1}^4 A_{ij} \left( (\mathbf{R}_i - \mathbf{R}_j)^2 - (d_{ij}^0)^2 \right)^2 + \sum_i^N \sum_{j=1}^3 \sum_{k=j+1}^4 B_{ijk} \left( (\mathbf{R}_j - \mathbf{R}_i) \cdot (\mathbf{R}_k - \mathbf{R}_i) - \cos \theta_{ij}^0 d_{ik}^0 \right)^2$$

W wyrażeniu tym znajdują się dwie sumy. W przypadku pierwszej sumy indeks  $i$  przebiega po wszystkich ( $N$ ) atomach w układzie, a indeks  $j$  po (4) najbliższych sąsiadach danego atom  $i$ ,  $A_{ij}$  oznacza z kolei stałą siłową związaną ze zmianą długości wiązania (ściskanie/rozciąganie) pomiędzy atomami  $i$  oraz  $j$ . Druga suma przebiega po wszystkich atomach oraz parach sąsiadów ( $ij$ ), a stała siłowa  $B_{ijk}$  związana jest ze zmianą kąta między wiązaniami  $ij$ - $ki$  w stosunku do idealnej wartości dla kryształu tetraedrycznego ( $\cos(\theta)=1/3$ ). W obu sumach  $\mathbf{R}_i$  oznacza położenia danego atomu.

Rozwiązanie problemu odkształceń w metodzie VFF polega na znalezieniu takich pozycji atomów ( $\mathbf{R}_i$ ) dla których energia  $E_{\text{vff}}$  osiąga minimum. Dla samorosnących kropek kwantowych liczba atomów w domenie obliczeniowej może sięgać nawet 100 milionów atomów, co prowadzi do złożonego problemu obliczeniowego.

### Znajdowanie minimum funkcji – metoda gradientów prostych

Istnieje wiele metod znajdowania ekstremum funkcji, jedną z nich jest metoda gradientu prostego (ang. *gradient descent*). Jest to metoda iteracyjna w której ekstremum osiągnęte jest w serii kolejnych kroków wykonywanych w kierunku największego wzrostu lub spadku, w zależności od tego czy celem optymalizacji jest minimum funkcji czy jej maksimum (Rys. 5.17).

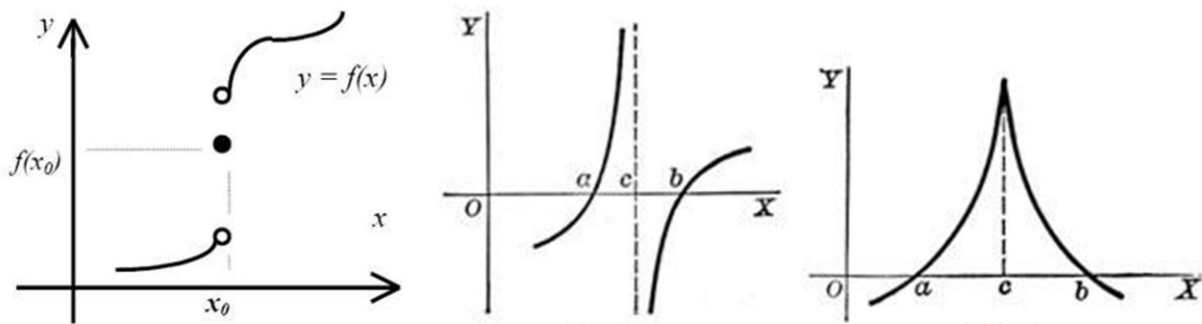


Rys.5.17 Najkrótsza droga na szczyt to zwykle ta najbardziej stroma.

Utożsamiając kierunek największego spadku z (minus) gradientem funkcji, bardziej sformalizowany algorytm wygląda następująco:

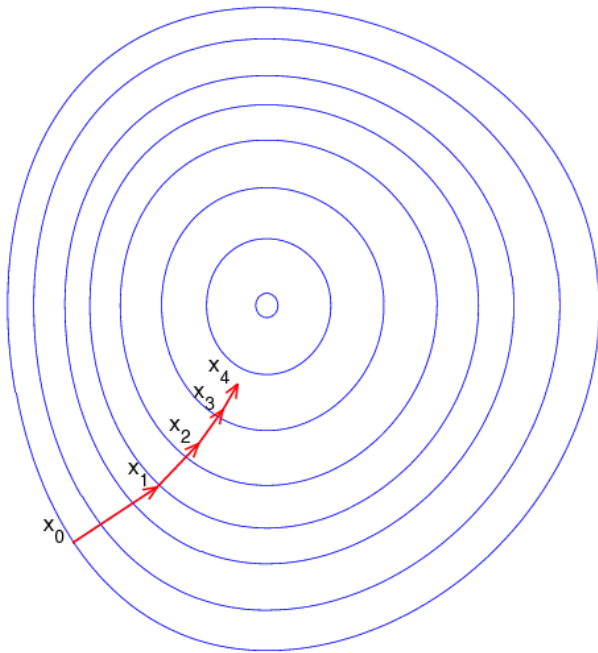
$$x_{k+1} = x_k - \alpha_k \nabla f_k$$

Przy czym  $f : R^N \mapsto R$  jest różniczkowalną i ciągłą (Rys. 5.18) funkcją  $N$  zmiennych.



Rys.5.18 Przykłady funkcji nieciągłych i nieróżniczkowalnych.

Proces iteracyjny (Rys. 5.19) rozpoczyna się od punktu startowego  $x_0$ , którego wybór jest arbitralny i zależy od konkretnego problemu. Długość kroku ( $\alpha_k$ ) powinna być z kolei tak dobrana, aby wartość funkcji w kolejnych punktach nie rosla ( $f(x_0) > \dots > f(x_k) > f(x_{k+1})$ ). Zbyt duży krok sprawi, że metoda nie będzie numerycznie stabilna, z kolei zbyt mały krok spowolni proces znajdowania minimum. Podobnie jak  $x_0$ , wybór  $\alpha_k$  jest w dużym stopniu arbitralny i zależy od konkretnego zagadnienia.



Rys.5.19 Kilka kolejnych kroków w metodzie gradientu prostego.

Jednym ze sposobów na uniknięcie powyższego problemu jest wariant metody gradientu prostego, tzw. metoda najszybszego spadku (ang. steepest descent), w której długość kroku nie jest parametrem, ale jest wyznaczana w każdym kroku iteracji, tak by funkcja osiągnęła minimum w na kierunku gradientu:

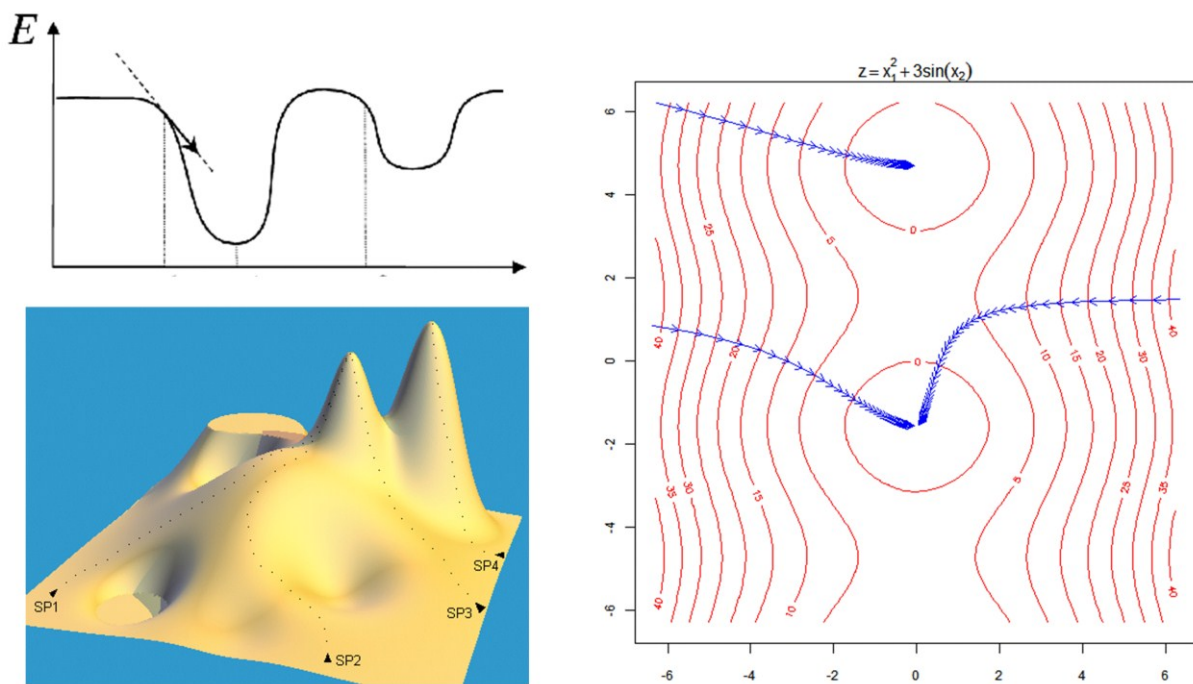
$$f(x_k - \alpha_k \nabla f(x_k)) = \min_{\alpha > 0} f(x_k - \alpha \nabla f(x_k))$$

Zaletą tego podejścia jest uniknięcie związanych kłopotów z wyborem  $\alpha_k$ , z drugiej strony jednak strony metoda ta jest bardziej skomplikowana od metody gradientu prostego oraz przede wszystkim zwiększona jest ilość (czas) wykonywanych obliczeń numerycznych.



Niezależnie od sposobu wyboru  $\alpha_k$  proces iteracyjny kontynuowany jest do momentu osiągnięcia minimum. W celu określenia, czy dany punkt dostatecznie dobrze przybliży minimum funkcji można użyć kilku kryteriów, np. sprawdzić czy długość wektora gradientu jest mniejsza od pewnej zadanej precyzji ( $\|\nabla f(x_k)\| < \varepsilon$ ) lub odległość między kolejnymi punktami jest mniejsza od pewnej odległości minimalnej ( $\|x_{k+1} - x_k\| < \delta$ ). Oba kryteria mogą być też stosowane jednocześnie.

Na zakończenie tej części wykładu należy zwrócić uwagę, że w metodzie gradientu prostego, jak i metodzie największego spadku, znajdowane jest lokalne, a nie globalne minimum funkcji (Rys.5. 20). Dla funkcji posiadających wiele minimów lokalnych, to które z nich zostanie znalezione zależy m.in. od wyboru punktu startowego.



Rys.5.20 Jedno- i dwuwymiarowe funkcje posiadające lokalne minima.