

# Reprezentacja trójwymiarowej, ruchomej sceny

M. K. Karpiński

Uniwersytet Wrocławski  
Instytut Informatyki

Wrocław, dnia 17 października 2012 r.

## 1 Wstęp

Badanie geometrycznych zależności między trójwymiarową sceną a dwuwymiarowymi obrazami generowanymi poprzez ruch kamery po scenie, leży w sercu dwóch fundamentalnych typów przekształceń:

1. ruchu Euklidesowego (czyli ruchu brył sztywnych)
2. rzutu perspektywicznego

Pierwsze z nich zostanie omówione na dzisiejszym seminarium, drugie zaś zostanie omówione za tydzień przez Maćka. Obydwa tematy będą wymagały znajomości podstawowych obiektów i notacji z dziedziny algebry. Tak na prawdę, całe dzisiejsze seminarium, to będzie czysta algebra, ale mam nadzieję, że was tym nie zanudzę. Agenda:

1. Świat, którym się zajmujemy (przestrzeń  $\mathbb{E}^3$ )
2. Obiekty, które żyją w tym świecie (bryły sztywne)
3. Ruch obrotowy i jego reprezentacje
4. Reprezentacje pełnego ruchu bryły sztywnej
5. Przekształcenia współrzędnych i prędkości

## 2 Przestrzeń euklidesowa

Niech  $\mathbb{E}^3$  będzie trójwymiarową przestrzenią euklidesową. Jest to naturalny sposób modelowania świata rzeczywistego. W ogólnym przypadku przestrzeń euklidesowa zdefiniowana jest poprzez 5 aksjomatów Euklidesa:

- dowolne dwa punkty można połączyć odcinkiem
- dowolny odcinek można przedłużyć nieograniczenie, uzyskując prostą
- dla danego odcinka można zaznaczyć okrąg o środku w jednym z jego końców i promieniu równym jego długości
- wszystkie kąty proste są przystające
- dwie proste, które przecinają przecią w taki sposób, że suma kątów wewnętrznych po jednej stronie jest mniejsza od dwóch kątów prostych, przetną się z tej właśnie strony

Dla naszych potrzeb, trójwymiarową przestrzeń euklidesową będziemy reprezentować za pomocą układu współrzędnych kartezjańskich: każdy punkt  $p \in \mathbb{E}^3$  może być zidentyfikowany poprzez punkt w  $\mathbb{R}^3$ :

$$X = [x_1, x_2, x_3]^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \quad (1)$$

Zamiast  $x_1, x_2, x_3$  często będziemy pisać  $x, y, z$ . Dzięki temu mamy zależność 1 – 1 między  $\mathbb{E}^3$  a  $\mathbb{R}^3$ , więc bezpiecznie możemy mówić o punktach i ich współrzędnych, jakby były tym samym bytem. Przeszliśmy pierwszy krok w stronę mierzenia odległości i kątów. Czego nam brakuje, to związać  $\mathbb{E}^3$  z metryką. Precyzyjna definicja metryki opiera się na definicji wektora:

**Definicja 1.** (wektor) W przestrzeni euklidesowej, wektor  $v$  jest reprezentowany przez parę punktów  $p, q \in \mathbb{E}^3$  i zdefiniowany jako strzałka skierowana łącząca  $p$  do  $q$ . Piszemy  $v = \vec{pq}$ .

Wektor  $v$  zapisujemy także jako  $[v_1, v_2, v_3]^T \in \mathbb{R}^3$ , gdzie  $v = Y - X$ .

Wektor jest zaczerpiony, jeśli podamy punkt  $p$  explicite. W przeciwnym wypadku mówimy o wektorach swobodnych.

Zbiór wszystkich wektorów swobodnych tworzy liniową przestrzeń wektorową z liniową kombinacją dwóch wektorów  $v, u \in \mathbb{R}^3$  zdefiniowaną jako:

$$\alpha v + \beta u = [\alpha v_1 + \beta u_1, \alpha v_2 + \beta u_2, \alpha v_3 + \beta u_3]^T, \quad \forall \alpha, \beta \in \mathbb{R} \quad (2)$$

Metryka Euklidesowa dla  $\mathbb{E}^3$  jest wtedy zdefiniowana poprzez produkt skalarny na przestrzeni wektorowej  $\mathbb{R}^3$ .

$$\langle u, v \rangle = u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3, \quad \forall u, v \in \mathbb{R}^3 \quad (3)$$

Niektóre własności:

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2 + v_3^2} \quad (4)$$

**Definicja 2.** Jeśli  $\langle u, v \rangle = 0$ , wtedy  $u$  i  $v$  są ortogonalne.

Podsumowując: przestrzeń euklidesowa  $\mathbb{E}^3$  może być zdefiniowana jako przestrzeń, która identyfikuje się ze zbiorem  $\mathbb{R}^3$  i posiada metrykę w swojej przestrzeni wektorowej daną poprzez powyższy iloczyn skalarny.

Jeszcze o wektorach:

**Definicja 3** (Iloczyn wektorowy). Mając dane dwa wektory  $u, v \in \mathbb{R}^3$  ich iloczyn wektorowy (cross product) jest wektorem ze współzrędnymi:

$$u \times v = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} \in \mathbb{R}^3 \quad (5)$$

Można łatwo sprawdzić, że:

$$\langle u \times v, u \rangle = \langle u \times v, v \rangle = 0, u \times v = -v \times u \quad (6)$$

Z tego wynika, że iloczyn wektorowy jest prostopadły do  $u$  oraz do  $v$  oraz kolejność czynników definiuje zwrot wektora wynikowego.

Jeśli ustalimy z góry  $u$ , to iloczyn wektorowy może być reprezentowany przez funkcję z  $\mathbb{R}^3$  do  $\mathbb{R}^3$ , która bierze  $v$  i zwraca  $u \times v$ . Taka funkcja jest zdefiniowana jako:

$$u \times v = \hat{u}v, \quad \hat{u} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (7)$$

Widać, że  $\hat{u}$  jest macierzą antysymetryczną tzn.  $\hat{u}^T = -\hat{u}$ . Jest oczywiste, że następujący lemat zachodzi:

**Lemat 4.** Macierz  $M \in \mathbb{R}^{3 \times 3}$  jest antysymetryczna wtedy i tylko wtedy, gdy  $M = \hat{u}$  dla pewnego  $u \in \mathbb{R}^3$

Dzięki czemu przestrzeń wektorowa  $\mathbb{R}^3$  i przestrzeń wszystkich macierzy antysymetrycznych  $3 \times 3$  (zwana  $so(3)$ ) są izomorficzne.

### 3 Ruch bryły sztywnej

Wyobraźmy sobie obiekt ruszający się przed kamerą. Aby precyzyjnie zdefiniować jego ruch powinniśmy ustalić trajektorię każdego punktu leżącego na obiekcie. Np. specyfikując współrzędne punktu poprzez funkcję od czasu  $X(t)$ . Na nasze szczęście nie musimy specyfikować ruchu każdego punktu. Jak za chwilę się przekonamy, do zdefiniowania ruchu obiektu wystarczy ustalić ruch jednego punktu oraz ruch trzech osi współrzędnych zaczepionych w tym punkcie. Powodem tego jest to, że dla każdej bryły sztywnej, odległość między dwoma dowolnymi punktami leżącymi na tej bryle nie zmienia się w czasie podczas ruchu tej bryły. Więc jeśli  $X(t)$  i  $Y(t)$  są współrzędnymi dowolnych dwóch punktów  $p$  i  $q$ , odległość między tymi punktami jest stała:

$$\|X(t) - Y(t)\| = \text{constant}, \quad \forall t \in R \quad (8)$$

Ruch bryły sztywnej (lub transformacja bryły sztywnej) jest rodziną funkcji, które określają w jaki sposób współrzędne każdego punktu na bryle sztywnej zmieniają się w czasie (respektując powyższą zasadę). Zdefiniujmy taką funkcję jako:

$$g(t) : \mathbb{R}^3 \rightarrow \mathbb{R}^3; X \rightarrow g(t)(X). \quad (9)$$

Zamiast patrzeć na całą ciągłą ścieżkę ruszającego się obiektu, skoncentrujemy się na funkcji między jego początkową a końcową konfiguracją (rigid-body displacement) zdefiniowaną jako:

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}^3; X \rightarrow g(X). \quad (10)$$

Oprócz transformacji na współrzędnych punktu,  $g$  można również wykorzystać do transformacji na wektorach. Niech  $v$  będzie wektorem zdefiniowanym przez punkty  $p$  i  $q$  o współrzędnych  $v = Y - X$ . Wtedy, po transformacji  $g$ , otrzymujemy nowy wektor:

$$u = g_*(v) = g(Y) - g(X) \quad (11)$$

Jako że  $g$  zachowuje odległości pomiędzy punktami, w oczywisty sposób otrzymujemy  $\|g_*(v)\| = \|v\|$  dla każdego swobodnego wektora  $v \in \mathbb{R}^3$ .

Zauważmy, że własność zachowania odległości między punktami nie jest wystarczająca aby poprawnie zdefiniować ruch bryły sztywnej. Np. taka transformacja:

$$f : [x, y, z]^T \rightarrow [x, y, -z]^T \quad (12)$$

Zachowuje odległość, ale nie zwrot! Aby w naszej rodzinie wykluczyć tego typu funkcje, wymagamy, aby zachowywać nie tylko odległość, ale też zwrot. Formalnie, oprócz zachowania modułu wektora, musimy także zachować produkt wektorowy.

**Definicja 5** (Ruch bryły sztywnej (rigid-body motion)). *Funkcja  $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  jest ruchem bryły sztywnej wtedy, gdy zachodzą:*

1.  $\|g_*(v)\| = \|v\|, \forall v \in \mathbb{R}^3$
2.  $g_*(u) \times g_*(v) = g_*(u \times v), \forall u, v \in \mathbb{R}^3$

Zbiór wszystkich takich ruchów nazywamy  $SE(3)$ .

Nie jest do końca oczywiste, że powyższa definicja zachowuje kąty między wektorami. Ale można to udowodnić w taki sposób: iloczyn skalarny  $\langle \cdot, \cdot \rangle$  definiujemy poprzez moduł  $\|\cdot\|$  (polarization identity):

$$\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2) \quad (13)$$

Więc skoro  $\|u + v\| = \|g_*(u) + g_*(v)\|$ , to otrzymujemy, że:

$$\langle u, v \rangle = \langle g_*(u), g_*(v) \rangle, \quad \forall u, v \in \mathbb{R}^3 \quad (14)$$

Innymi słowy, do definicji ruchu bryły sztywnej można dodać zachowanie iloczynu skalarnego.

W jaki sposób te własności pomagają nam precyzyjnie opisać ruch bryły sztywnej? Fakt, że odległości i zwroty są zachowane przez ruch bryły sztywnej oznacza, że punkty nie mogą ruszać się relatywnie do siebie. Konsekwencją tego faktu jest to, że ruch bryły sztywnej może być opisany przez ruch jednego wybranego punktu na bryle i rotacji układu współrzędnych związanego z tym punktem. Będziemy reprezentowali konfigurację bryły sztywnej związując koratezjański układ współrzędnych (frame) do jakiegoś punktu na obiekcie i śledząc ruch tego układu relatywnie do ustalonego układu współrzędnych świata (uład globalny). World (reference) frame.

Weźmy układ współrzędnych wyznaczony przez trzy ortonormalne wektory  $e_1, e_2, e_3 \in \mathbb{R}^3$ , które spełniają:

$$e_i^T e_j = \delta_{ij} = \begin{cases} 1 & \text{jeśli } i = j \\ 0 & \text{wpw} \end{cases} \quad (15)$$

Wektory te są ułożone tak, aby tworzyły prawoskrętny układ:  $e_1 \times e_2 = e_3$ . Ruch bryły sztywnej  $g$  daje nam:

$$g_*(e_i)^T g_*(e_j) = \delta_{ij}, \quad g_*(e_1) \times g_*(e_2) = g_*(e_3) \quad (16)$$

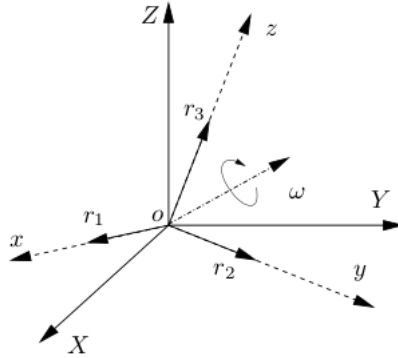
To znacza, że wektory wynikowe  $g_*(e_1), g_*(e_2), g_*(e_3)$  wciąż tworzą prawoskrętny ortonormalny układ. Więc bryła sztywna zawsze może być rozpoznawana przez prawoskrętny, ortonormalny układ, który będziemy nazywali 'object coordinate frame' lub 'body coordinate frame' oraz jej ruch może być w całości opisany przez ruch takiego układu.

W szczególności powinniśmy ustalić relację między kamerą a resztą świata. Możemy np. ustalić, że 'world frame' jest związany z obiektem kamery i reszta obiektów porusza się względem niej.

## 4 Ruch obrotowy i jego reprezentacje

### 4.1 Reprezentacja obrotu poprzez macierze ortogonalne

Założmy, że mamy bryłę sztywną obracającą się wokół ustalonego punktu  $o \in \mathbb{E}^3$ . Założmy, że w tym punkcie znajduje się też początek układu globalnego  $W$ . Zaczepimy do  $o$  jeszcze jeden układ  $C$ , będący obracającą się kamerą. Relację między tymi dwoma układami współrzędnych ilustruje poniższy rysunek:



Konfiguracja układu  $C$  względem układu  $W$  jest wyznaczona przez współrzędne trzech ortonormalnych wektorów

$$r_1 = g_*(e_1), r_2 = g_*(e_2), r_3 = g_*(e_3) \in \mathbb{R}^3 \quad (17)$$

$r_1, r_2, r_3$  są po prostu wektorami jednostkowymi położonymi na współrzędnych  $x, y, z$  układu  $C$ . Konfiguracja obracającego się układu jest wtedy w całości wyznaczona przez macierz  $3 \times 3$ :

$$R_{wc} = [r_1, r_2, r_3] \in \mathbb{R}^{3 \times 3} \quad (18)$$

gdzie  $r_1, r_2, r_3$  są ułożone kolumnami. Z tego że  $r_1, r_2, r_3$  tworzą układ ortogonalny, mamy:

$$r_i^T r_j = \delta_{ij} = \begin{cases} 1 & \text{jeśli } i = j \\ 0 & \text{wpp} \end{cases} \quad \forall i, j \in 1, 2, 3 \quad (19)$$

co może być zapisane w postaci macierzy:

$$R_{wc}^T R_{wc} = R_{wc} R_{wc}^T = I \quad (20)$$

Każda macierz spełniająca powyższą zależność nazywa się macierzą ortogonalną. Z definicji wynika również, że:

$$R_{wc}^{-1} = R_{wc}^T. \quad (21)$$

Z tego że  $r_1, r_2, r_3$  tworzą układ prawoskrętny, wymagamy aby wyznacznik macierzy  $R_{wc}$  był równy  $+1$ . Przestrzeń wszystkich takich ortogonalnych macierzy należących do  $\mathbb{R}^{3 \times 3}$  definiujemy jako:

$$SO(3) = \{R \in \mathbb{R}^{3 \times 3} | R^T R = I, \det(R) = +1\} \quad (22)$$

Takie macierze ortogonalne tradycyjnie nazywają się macierzami rotacji (z wiadomych powodów). (rotation matrices). Side note:  $SO(3)$  z operacją mnożenia macierzy tworzą grupę. Dlatego  $SO(3)$  nazywa się czasem grupą obrotową.

Przykład:

$$R_Z(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \quad (23)$$

Wracając do rysunku, każda macierz obrotu  $R_{wc} \in SO(3)$  reprezentuje możliwą konfigurację obiektu obruczonego wokół punktu  $o$ . Ponadto taka macierz reprezentuje transformację z układu współrzędnych  $C$  do  $W$ . Konfigurację ciągle obracającego się obiektu można opisać jako funkcję trajektorii  $R(t) : t \rightarrow SO(3)$  w przestrzeni  $SO(3)$ . Jeśli  $t \neq 0$ , obrót pomiędzy czasem  $t_2$  i  $t_1$  opiszemy jako  $R(t_2, t_1)$ . W prosty sposób można wywnioskować iż:

$$R(t_2, t_0) = R(t_2, t_1)R(t_1, t_0) \quad \forall t_0 < t_1 < t_2 \in \mathbb{R} \quad (24)$$

## 4.2 Reprezentacja obrotu przez współrzędne eksponentne

Wg autora jest bardziej intuicyjny. Obrót jest parametryzowany poprzez dwie rzeczy:

- wektor jednostkowy wychodzący z układu
- kąt obrotu wektora (prawoskrętnie)

Wygląda obiecująco, ale czy na pewno można w ten sposób reprezentować każdy obrót bryły sztywnej? Trzeba to udowodnić. Zaczynamy od początku.

Mamy trajektorię  $R(t) : \mathbb{R} \rightarrow SO(3)$ , która wyznacza ciągły ruch obrotowy. Obrót musi spełniać warunek:

$$R(t)R^T(t) = I \quad (25)$$

Obliczając pochodną w powyższym równaniu po zmiennej  $t$  i zauważysz, że prawa strona jest macierzą stałą, otrzymujemy:

$$\dot{R}(t)R^T(t) + R(t)\dot{R}^T(t) = 0 \Rightarrow \dot{R}(t)R^T(t) = -(\dot{R}(t)R^T(t))^T \quad (26)$$

Oznacza to, że macierz  $\dot{R}(t)R^T(t) \in \mathbb{R}^{3 \times 3}$  jest antysymetryczna. Czyli musi istnieć wektor  $\hat{\omega}(t) \in \mathbb{R}^3$  taki, że:

$$\dot{R}(t)R^T(t) = \hat{\omega}(t) \quad (27)$$

a po pomnożeniu obustronnie przez  $R(t)$ :

$$\dot{R}(t) = \hat{\omega}(t)R(t) \quad (28)$$

Teraz wykorzystamy tę wiedzę do stworzenia użytecznej reprezentacji macierzy rotacyjnej. Zaczniemy od założenia, że  $\hat{\omega}$  jest ustalone. Oczywiście jest, że rozwiązanie tego równania jest podane przez:

$$x(t) = e^{\hat{\omega}t}x(0) \quad (29)$$

gdzie  $e^{\hat{\omega}t}$  jest eksponentą macierzy:

$$e^{\hat{\omega}t} = I + \hat{\omega}t + \frac{(\hat{\omega}t)^2}{2!} + \dots + \frac{(\hat{\omega}t)^n}{n!} + \dots \quad (30)$$

Zakładając że  $R(0) = I$  jest stanem początkowym rotacji otrzymujemy:

$$R(t) = e^{\hat{\omega}t} \quad (31)$$

Aby sprawdzić, czy rzeczywiście  $e^{\hat{\omega}t}$  jest macierzą rotacyjną, można wprost z definicji eksponenty pokazać, że:



$$(e^{\hat{\omega}t})^{-1} = e^{-\hat{\omega}t} = e^{\hat{\omega}^T t} = (e^{\hat{\omega}t})^T \quad (32)$$

Czyli  $(e^{\hat{\omega}t})^T e^{\hat{\omega}t} = I$ . Fizyczna interpretacja  $R(t) = e^{\hat{\omega}t}$  jest taka, o której mówiliśmy na początku tego działu, czyli jeśli założymy, że  $\|\omega\| = 1$ , wtedy  $R(t) = e^{\hat{\omega}t}$  jest obrotem dookoła osi  $\omega \in \mathbb{R}^3$  o kąt  $t$  radianów. Ogólnie możemy  $t$  absorbować do  $\omega$ , więc mamy  $R = e^{\hat{\omega}}$  dla  $\omega$  o dowolnej długości.

Eksponenta definiuje funkcję z przestrzeni  $so(3)$  do przestrzeni  $SO(3)$ , zwie się to “exponential map”.

Zauważmy że otrzymaliśmy wzór  $R(t) = e^{\hat{\omega}t}$  zakładając, że  $\omega(t)$  jest stały. Ale oczywiście nie zawsze tak jest. Wtedy naturalnie powstaje pytanie: czy każda macierz rotacyjna  $R \in SO(3)$  może być wyrażona w formie eksponenty? Odpowiedź brzmi: tak, i fakt ten pokazuje poniższe twierdzenie:

**Twierdzenie 6** (logarytm  $SO(3)$ ). *Dla każdego  $R \in SO(3)$ , istnieje  $\omega \in \mathbb{R}^3$  takie, że  $R = e^{\hat{\omega}}$ .*

Jest to tzw. funkcja z przestrzeni  $SO(3)$  do  $so(3)$ . Zapisujemy  $\hat{\omega} = \log(R)$

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (33)$$

$$\theta = \|\omega\| = \arccos\left(\frac{\text{trace}(R) - 1}{2}\right) \quad (34)$$

$$\omega = \frac{1}{2 \sin(\theta)} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix} \quad (35)$$

Dzięki temu twierdzeniu możemy w efektywny sposób obliczać wektor obrotu mając daną macierz obrotu, ale w jaki sposób efektywnie obliczyć macierz obrotu  $R = e^{\hat{\omega}}$  mając dany wektor  $\omega$ ? Z definicji eksponenty to tak nie bardzo. Możemy użyć następującego twierdzenia:

**Twierdzenie 7** (wzór Rodrigues’a dla macierzy rotacyjnych). *Mając dany  $\omega \in \mathbb{R}^3$ , eksponenta macierzy  $R = e^{\hat{\omega}}$  jest równa:*

$$e^{\hat{\omega}} = I + \frac{\hat{\omega}}{\|\omega\|} \sin(\|\omega\|) + \frac{\hat{\omega}^2}{\|\omega\|^2} (1 - \cos(\|\omega\|)) \quad (36)$$

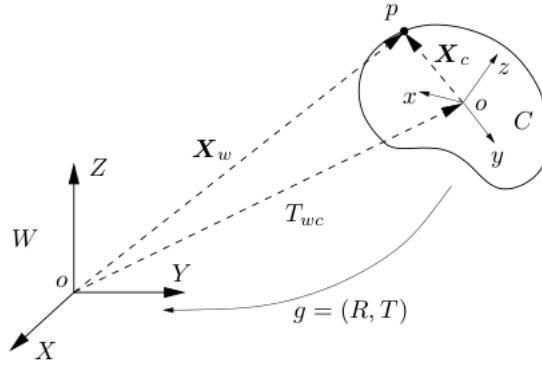
Wprost z definicji eksponenty możemy zapisać:

$$e^{\hat{\omega}t} = I + \hat{\omega}\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots\right) + \hat{\omega}^2\left(\frac{t^2}{2!} - \frac{t^4}{4!} + \frac{t^6}{6!} - \dots\right) \quad (37)$$

Co w zasadzie kończy dowód.

Oczywiste jest, że jeśli  $\|w\| = 1$  i  $t = 2k\pi$ , to dostaniemy identyczność  $(\forall k)$ . Więc transformacja z  $so(3)$  do  $SO(3)$  nie jest bijekcją, bo ma nieskończenie wiele elementów w  $SO(3)$ .

## 5 Reprezentacje pełnego ruchu bryły sztywnej



Czyli ruch z rotacją oraz translacją. Rysunek 3 przedstawia ruszający się obiekt z układem  $C$  do niego przypisanym. Aby opisać współrzędne punktu  $p$  na obiekcie odnośnie do układu globalnego  $W$ , jest oczywiste że należy dodać do siebie wektor translacji  $T_{wc}$  i wektor  $X_c$ , ale opisany relatywnie do  $W$ . Z tego że  $X_c$  są współrzędnymi punktu  $p$  relatywnymi do układu  $C$ , mamy  $R_{wc}X_c$ , gdzie  $R_{wc} \in SO(3)$  jest rotacją między układami  $W$  i  $C$ . Wtedy współrzędne  $X_w$  można otrzymać tak:

$$X_w = R_{wc}X_c + T_{wc} \quad (38)$$

Zwykle będziemy oznaczali pełny ruch bryły sztywnej przez  $g_{wc} = (R_{wc}, T_{wc})$  lub po prostu  $g = (R, T)$  jeśli układy współrzędnych będą w oczywisty sposób wynikały z kontekstu. Wtedy możemy reprezentować na raz rotację i translację pomiędzy dwoma układami współrzędnych. W zwięzłej formie piszemy:

$$X_w = g_{wc}(X_c) \quad (39)$$

Zbiór wszystkich możliwych konfiguracji bryły sztywnej może być opisany przez przestrzeń ruchów brył sztywnych:

$$SE(3) = \{g = (R, T) | R \in SO(3), T \in \mathbb{R}^3\} \quad (40)$$

Zauważmy, że  $g = (R, T)$  nie jest jeszcze reprezentacją macierzową dla  $SE(3)$ . Żeby tak się stało złączenie dwóch ruchów bryły sztywnej musiałyby być iloczynem dwóch macierzy. Aby uzyskać taką reprezentację musimy wprowadzić pojęcie tzw. współrzędnych jednorodnych.

## 5.1 Reprezentacja jednorodna

Jak niektórzy mogli już zauważyć, w odróżnieniu od samych rotacji obiektów, transformacja pełnego ruchu bryły sztywnej nie jest liniowa, lecz afiniczna. W ramach przypomnienia: powiemy że dwa wektory -  $u$  i  $v$  - są związane liniową transformacją jeśli  $u = Av$  dla pewnej macierzy  $A$ , i są związane transformacją afiniczną jeśli  $u = Av + b$  dla pewnej macierzy  $A$  i wektora  $b$ .

Nie boimy się mimo wszystko, bo możemy dokonać konwersji z transformacji afinicznej do liniowej wykorzystując tzw. współrzędne jednorodne. Zamieniamy punkt  $X = [x_1, x_2, x_3]^T \in \mathbb{R}^3$  na  $\bar{X} = [x_1, x_2, x_3, 1]^T \in \mathbb{R}^4$

W efekcie, takie rozszerzenie współrzędnych osadza przestrzeń euklidesową  $\mathbb{E}^3$  w hiperpłaszczyźnie  $\mathbb{R}^4$  zamiast  $\mathbb{R}^3$ . Współrzędne jednorodne wektora  $v = X(q) - X(p)$  są zdefiniowane jako różnica między jednorodnymi współrzędnymi dwóch punktów  $q$  i  $p$ , stąd otrzymujemy:

$$\bar{v} = \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} X(q) \\ 1 \end{bmatrix} - \begin{bmatrix} X(p) \\ 1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \end{bmatrix} \in \mathbb{R}^4 \quad (41)$$

Zauważmy, że wektory w powyższej formie tworzą podprzestrzeń, w której wszystkie liniowe struktury oryginalnych wektorów  $v \in \mathbb{R}^3$  są zachowane w niezminionej formie przez nową reprezentację. Teraz, używając nowej notacji możemy transformację afiniczną  $X_w = R_{wc}X_c + T_{wc}$  przepisać w postaci liniowej:

$$\bar{X}_w = \begin{bmatrix} X_w \\ 1 \end{bmatrix} = \begin{bmatrix} R_{wc} & T_{wc} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ 1 \end{bmatrix} = \bar{g}_{wc} \bar{X}_c \quad (42)$$

gdzie macierz  $\bar{g}_{wc} \in \mathbb{R}^{4 \times 4}$  zwie się jednorodną reprezentacją ruchu brył sztywnych  $g_{wc} = (R_{wc}, T_{wc}) \in SE(3)$ . Generalnie, jeśli  $g = (R, T)$  wtedy jego reprezentacją jednorodną jest macierz

$$\bar{g} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (43)$$

Zauważmy, że dzięki temu możemy reprezentować transformację współrzędnych bryły sztywnej przez liniowe mnożenie macierzy. Reprezentacja jednorodna  $g$  daje nam naturalną macierzową reprezentację  $SE(3)$ :

$$SE(3) = \{\bar{g} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \mid R \in SO(3), T \in \mathbb{R}^3\} \subset \mathbb{R}^{4 \times 4} \quad (44)$$

Używając takiej reprezentacji w łatwy sposób można sprawdzić, że  $SE(3)$  spełnia wszystkie wymagania do bycia grupą. W szczególności  $\forall g_1, g_2, g \in SE(3)$  mamy:

$$\bar{g}_1 \bar{g}_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1 R_2 & R_1 T_2 + T_1 \\ 0 & 1 \end{bmatrix} \in SE(3) \quad (45)$$

$$g^{-1} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^T & -R^T T \\ 0 & 1 \end{bmatrix} \in SE(3) \quad (46)$$

W reprezentacji jednorodnej transformacja o wektor  $v = X(q) - X(p) \in \mathbb{R}^3$  wygląda tak:

$$\bar{g}_*(\bar{v}) = \bar{g}\bar{X}(q) - \bar{g}\bar{X}(p) = \bar{g}\bar{v} \quad (47)$$

Czyli akcja jest po prostu opisana przez iloczyn macierzy. Słuchacz może sobie sam zweryfikować, że  $\bar{g}$  faktycznie zachowuje iloczyn skalarny i iloczyn wektorowy, co dowodzi, że  $\bar{g}$  jest ruchem bryły sztywnej zgodnym z podaną wcześniej definicją. Jak widać, ruch brył sztywnej zachowuje się inaczej na punktach (które można obracać i przemieszczać) niż na wektorach (które można jedynie obracać)

## 5.2 Reprezentacja przez współrzędne eksponentne

W punkcie 4.2 niniejszej prezentacji studiowaliśmy współrzędne eksponentne dla rotacji macierzy  $R \in SO(3)$ . Podobna reprezentacja współrzędnych również istnieje dla reprezentacji jednorodnej ruchu bryły sztywnej  $g \in SE(3)$ . Cieszymy się, oto teraz dojdziemy do formy ostatecznej, która będzie wykorzystywana przez wszystkie następne prezentacje. Pokażemy jak rozszerzyć to, co uzyskaliśmy w części 4.2 na pełny ruch bryły sztywnej. Postępujemy analogicznie:

Założmy, że mamy ciągle poruszającą się bryłę sztywną opisaną przez trajektorię na  $SE(3)$ :  $g(t) = (R(t), T(t))$ . W postaci jednorodnej mamy:

$$g(t) = \begin{bmatrix} R(t) & T(t) \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (48)$$

Analogicznie jak w przypadku rotacji eksponentnej, zbadajmy taką macierz:

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} \dot{R}(t)R^T(t) & \dot{T}(t) - \dot{R}(t)R^T(t)T(t) \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (49)$$

Wiemy, że  $\dot{R}R^T(t)$  jest macierzą antysymetryczną, czyli istnieje  $\hat{w}(t) \in so(3)$  taki że  $\hat{w}(t) = \dot{R}R^T(t)$ . Zdefiniujmy także wektor  $v(t) \in \mathbb{R}^3$  taki że  $v(t) = \dot{T}(t) - \hat{w}(t)T(t)$ . Wtedy powyższe równanie zapiszemy tak:

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} \hat{w}(t) & v(t) \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (50)$$

Jeśli następnie zdefiniujemy  $\hat{\xi} \in \mathbb{R}^{4 \times 4}$ :

$$\hat{\xi}(t) = \begin{bmatrix} \hat{w}(t) & v(t) \\ 0 & 0 \end{bmatrix} \quad (51)$$

To otrzymamy:

$$\dot{g}(t) = (\dot{g}(t)g^{-1}(t))g(t) = \hat{\xi}(t)g(t) \quad (52)$$

Na  $\hat{\xi}$  można wtedy patrzeć jako wektor stycznej do krzywej w  $g(t)$  i może być wykorzystany do aproksymowania  $g(t)$  lokalnie:

$$g(t + dt) \approx g(t) + \hat{\xi}(t)g(t)dt = (I + \hat{\xi}(t)dt)g(t) \quad (53)$$

macierz  $4 \times 4$  w formie  $\hat{\xi}$  nazywamy twistem (skrętem). Zbiór wszystkich skrętów opisujemy:

$$se(3) = \left\{ \hat{\xi} = \begin{bmatrix} \hat{\omega} & v \\ 0 & 0 \end{bmatrix} \mid \hat{\omega} \in so(3), v \in \mathbb{R}^3 \right\} \subset \mathbb{R}^{4 \times 4} \quad (54)$$

Zbiór  $se(3)$  nazywamy przestrzenią styczną (lub algebrą Lie) grupy  $SE(3)$ . We współrzędnych skrętu  $\xi$  nazwiemy  $v$  jako prędkość liniową,  $\omega$  jako prędkość kątową (co będzie wskazywało na powiązanie z częścią translacyjną lub rotacyjną pełnego ruchu bryły sztywnej). Załóżmy teraz, że  $\hat{\xi}$  jest stały, w równaniu:

$$\dot{g}(t) = \hat{\xi}g(t) \quad (55)$$

znów otrzymujemy liniowe równanie różniczkowe zwyczajne, które po scałkowaniu daje nam:

$$g(t) = e^{\hat{\xi}t}g(0) \quad (56)$$

Zakładając, że początkowy stan  $g(0) = I$ , dochodzimy do wniosku iż:

$$g(t) = e^{\hat{\xi}t} \quad (57)$$

gdzie eksponenta skrętu jest dana jako:

$$e^{\hat{\xi}t} = I + \hat{\xi}t + \frac{(\hat{\xi}t)^2}{2!} + \dots + \frac{(\hat{\xi}t)^n}{n!} + \dots \quad (58)$$

Wykorzystując wzór Rodrigues'a i inne ciekawe własności o których autor nie wspomniał, można wyprowadzić taki oto wzór na  $e^{\hat{\xi}}$ :

$$e^{\hat{\xi}} = \begin{bmatrix} e^{\hat{\omega}} & \frac{(I - e^{\hat{\omega}}\hat{\omega}v + \omega\omega^T v)}{\|\omega\|} \\ 0 & 1 \end{bmatrix}, \quad \text{jeśli } \omega \neq 0 \quad (59)$$

Jeśli  $\omega = 0$ , eksponenta jest po prostu równa  $e^{\hat{\xi}} = \begin{bmatrix} I & v \\ 0 & 1 \end{bmatrix}$ . Patrząc na ten wspaniały wzór od razu jesteśmy w stanie stwierdzić, że faktycznie  $e^{\hat{\xi}}$  jest macierzą transformacji bryły sztywnej w  $SE(3)$ . Analogicznie jak w sekcji 4.2 możemy zdefiniować funkcję z  $se(3)$  do  $SE(3)$  jako:

$$\exp : se(3) \rightarrow SE(3); \quad \hat{\xi} \rightarrow e^{\hat{\xi}} \quad (60)$$

Znów trzeba sobie zadać pytanie, z uwagi na założenie że  $\hat{\xi}$  jest stałą, czy każdy ruch bryły sztywnej  $g \in SE(3)$  może być reprezentowany w formie eksponenty? Innymi słowy czy istnieje funkcja odwrotna/logarytmiczna. Odpowiedź oczywiście brzmi tak.

**Twierdzenie 8** (Logarytm  $SE(3)$ ). *Dla każdego  $g \in SE(3)$ , istnieją współrzędne skrętu  $\xi = (v, \omega)$  takie, że  $g = \exp(\hat{\xi})$ .*

Dowód przez konstrukcję. Niech  $g = (R, T)$ . Z tw. o logarytmie  $SO(3)$  wiemy, że dla macierzy rotacyjnej  $R \in SO(3)$  możemy znaleźć  $\omega$ , że  $e^{\hat{\omega}} = R$ . Jeśli  $R \neq I$ , czyli  $\|\omega\| \neq 0$ , wektor  $v \in \mathbb{R}^3$  można wyliczyć rozwiązując następujące równanie liniowe:

$$\frac{(I - e^{\hat{\omega}}\hat{\omega}v + \omega\omega^T v)}{\|\omega\|} = T \quad (61)$$

Jeśli  $R = I$ , wtedy  $\|\omega\| = 0$ . W takim wypadku, mamy po prostu  $\omega = 0$  i  $v = T$ .

## 6 Przekształcenia współrzędnych i prędkości

W następnych rozdziałach często będziemy chcieli wiedzieć jak współrzędne punktu i jego prędkość zmieniają się gdy kamera się porusza. Taka wiedza jest potrzebna, gdyż zwykle wygodniej jest wybrać układ kamery jako 'reference frame' i opisywać ruch kamery i punktu relatywnie do tego układu. Jako że kamera może się poruszać, musimy wiedzieć jak zamienić wartości, takie jak współrzędne czy prędkość, z jednego układu kamery do drugiego. W szczególności chcemy wiedzieć jak poprawnie wyrażać położenie i prędkość punktu względem kamery. Teraz zaprezentuję konwencję, która będzie stosowana do końca seminarium.

### 6.1 Zasady transformacji współrzędnych

Czas  $t \in R$  będzie zwykle wykorzystany jako index ruchu kamery. Nawet w przypadku dyskretnym, gdy będzie podanych tylko klika 'snapshotów', wciąż będziemy wykorzystywać  $t$  jako index pozycji kamery i obrazka z jakim jest związany (czyli jaki generuje w danym momencie). Więc będziemy używali  $g(t) = (R(t), T(t)) \in SE(3)$  lub

$$g(t) = \begin{bmatrix} R(t) & T(t) \\ 0 & 1 \end{bmatrix} \in SE(3) \quad (62)$$

aby oznaczyć przemieszczenie między jakimś ustalonym układem  $W$  a kamerą  $C$  w czasie  $t \in R$ . Zauważcie, że pominąłem notację  $wc$  w subindexach w notacji  $g_{cw}(t)$ , bo to oczywiście wynika z kontekstu. Ponadto zakładamy, że  $g(0) = I$ , czyli w czasie  $t = 0$  kamera jest przystająca do układu świata  $W$ . Więc jeśli współrzędne punktu  $p \in \mathbb{E}^3$  ustawione względem układu  $W$  są podane jako  $X_0 = X(0)$ , to współrzędne ustawione względem kamery w czasie  $t$  są podane jako

$$X(t) = R(t)X_0 + T(t) \quad (63)$$

lub w postaci jednorodnej:

$$X(t) = g(t)X_0 \quad (64)$$

Jeśli kamera znajduje się w miejscach  $g(t_1), g(t_2) \dots g(t_m)$  w czasach  $t_1, t_2 \dots t_m$ , to współrzędne punktu  $p$  są podane jako  $X(t_i) = g(t_i)X_0, i = 1, 2 \dots m$ .

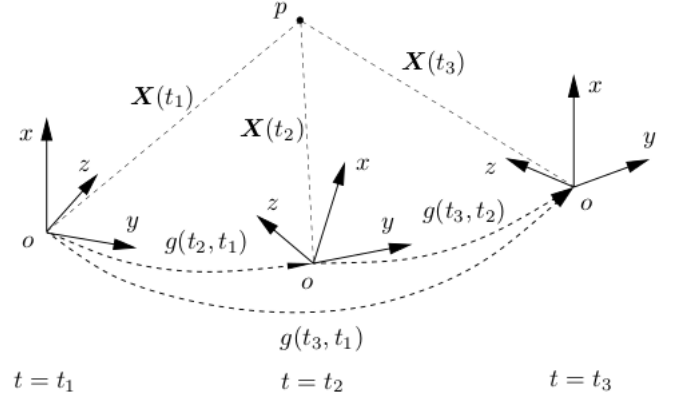
Jeśli tylko pozycja, a nie czas będzie w danej chwili nas interesować, będziemy pisali  $g_i$  zamiast  $g(t_i)$  i podobnie  $R_i$  zamiast  $R(t_i)$ ,  $T_i$  zamiast  $T(t_i)$  i  $X_i$  zamiast  $X(t_i)$ . Więc mamy:

$$X_i = R_i X_0 + T_i \quad (65)$$

Ponadto jeśli kamera nie startuje z  $t = 0$ , ruch między kamerą w czasie  $t_2$  a  $t_1$  będziemy oznaczać  $g(t_2, t_1) \in SE(3)$ . Wtedy otrzymujemy następującą zależność między współrzędnymi punktu  $p$  w różnych czasach:

$$X(t_2) = g(t_2, t_1)X(t_1), \quad \forall t_2, t_1 \in \mathbb{R} \quad (66)$$

Weźmy teraz trzecią pozycję kamery  $t = t_3 \in \mathbb{R}$ , tak jak na poniższym rysunku:



Ruch między kamerą w  $t_3$  a kamerą w  $t_2$  oznaczamy  $g(t_3, t_2)$  a między kamerą w  $t_3$  a  $t_1$  oznaczamy  $g(t_3, t_1)$ . Mamy wtedy następującą zależność między współrzędnymi:

$$X(t_3) = g(t_3, t_2)X(t_2) = g(t_3, t_2)g(t_2, t_1)X(t_1) \quad (67)$$

przyrównując to do bezpośredniej relacji między współrzędnymi w czasie  $t_3$  a  $t_1$ :

$$X(t_3) = g(t_3, t_1)X(t_1) \quad (68)$$

widzimy że zachowana jest następująca zasada:

$$g(t_3, t_1) = g(t_3, t_2)g(t_2, t_1) \quad (69)$$

Zasada złożenia opisuje współrzędne  $X$  punktu  $p$  relatywnie do każdej pozycji kamery. Implikacją tej zasady jest również zasada inwersji:

$$g^{-1}(t_2, t_1) = g(t_1, t_2) \quad (70)$$



gdyż  $g(t_2, t_1)g(t_1, t_2) = g(t_2, t_2) = I$ .

W przypadku gdy czas nie ma dla nas dużego znaczenia, możemy zamienić oznaczenie  $g(t_i, t_j)$  na  $g_{ij}$ . Powyższe zasady zapisujemy wtedy jako:

$$X_i = g_{ij}X_j, \quad g_{ik} = g_{ij}g_{jk}, \quad g_{ij}^{-1} = g_{ji} \quad (71)$$

## 6.2 Zasady transformacji prędkości

Rozumiejąc już jak działa transformacja współrzędnych, należy teraz prze-studiować jaki ma ona wpływ na prędkość. Wiemy, że współrzędne  $X(t)$  punktu  $p \in \mathbb{E}^3$  względem do ruszającej się kamery są funkcją od czasu  $t$ :

$$X(t) = g_{wc}(t)X_0 \quad (72)$$

oraz prędkość w punkcie  $p$  chwilowego układu kamery jest:

$$\dot{X}(t) = \dot{g}_{cw}(t)X_0 \quad (73)$$

Aby wyrazić  $\dot{X}(t)$  w formie wartości ruszającego się układu zamienimy  $X_0$  na  $g_{cw}^{-1}X(t)$  i używając notacji skrętu zdefiniujemy:

$$\hat{V}_{cw}^c = \dot{g}_{cw}(t)g_{cw}^{-1}(t) \in se(3) \quad (74)$$

Co można jeszcze inaczej zapisać jako

$$\dot{X}(t) = \hat{V}_{cw}^c(t)X(t) \quad (75)$$

z tego że  $\hat{V}_{cw}^c(t)$  jest postaci

$$\hat{V}_{cw}^c(t) = \begin{bmatrix} \hat{\omega}(t) & v(t) \\ 0 & 0 \end{bmatrix} \quad (76)$$

Możemy przepisać prędkość punktu w przestrzeni 3d (zamiast współrzędnych jednorodnych) jako:

$$\dot{X}(t) = \hat{\omega}(t)X(t) + v(t) \quad (77)$$

Fizyczną interpretacją sybolu  $\hat{V}_{cw}^c$  jest prędkość układu globalnego ruszającego się względnie do układu kamery, potrząc z widoku kamery. Zwykle, by jasno ustalić fizyczne znaczenie prędkości, musimy podać prędkość którego układu rusza się względem którego, i z którego układu patrzymy. Jeśli zmienimy pozycję, z której obserwujemy prędkość, wyrażenie zmienia się wraz ze zmianą widoku. Np. założmy że obserwator jest w innym układzie współrzędnych przemieszczonych względem układu kamery poprzez ruch bryły sztywnej  $g \in SE(3)$ . Wtedy współrzędne tego samego punktu  $p$  względem tego

układu są  $Y(t) = gX(t)$ . Wyznaczamy prędkość w nowym układzie i otrzymujemy:

$$\dot{Y}(t) = g\dot{g}_{cw}(t)g_{cw}^{-1}(t)g^{-1}Y(t) = g\hat{V}_{cw}^c g^{-1}Y(t) \quad (78)$$

Więc nowa prędkość (skręt) jest

$$\hat{V} = g\hat{V}_{cw}^c g^{-1} \quad (79)$$

To jest ta sama wartość fizyczna ale obserwowana z innego punktu widokowego. Widzimy, że prędkości są połączone funkcją zdefiniowaną przez ruch  $g$ :

$$ad_g : se(3) \rightarrow se(3); \quad \hat{\xi} \rightarrow g\hat{\xi}g^{-1} \quad (80)$$

To jest tzw. funkcja sprzężenia w przestrzeni  $se(3)$ . Używając tej notacji w poprzednim przykładzie mamy  $\hat{V} = ad_g(\hat{V}_{cw}^c)$ . Zauważmy, że sprzężenie przekształca prędkość z jednego układu do drugiego. Wykorzystując fakt że  $g_{cw}(t)g_{wc}(t) = I$  oczywistym jest, iż:

$$\hat{V}_{cw}^c = \dot{g}_{cw}g_{cw}^{-1} = -g_{wc}^{-1}\dot{g}_{wc} = -g_{cw}(\dot{g}_{wc}g_{wc}^{-1})g_{cw}^{-1} = ad_{g_{cw}}(-\hat{V}_{wc}^w) \quad (81)$$

więc  $\hat{V}_{cw}^c$  można interpretować jako negację prędkości kamery ruszającej się względem układu globalnego, patrząc z układu kamery.