

prof. dr hab. inż. Ignacy Duleba
Katedra Cybernetyki i Robotyki
Wydział Elektroniki, Fotoniki i Mikrosystemów
Politechnika Wrocławska
ignacy.duleba@pwr.edu.pl

Wrocław 2023-07-27

**Recenzja rozprawy doktorskiej mgra inż. Pawła Maciąga
"Optimal Control of Multibody Systems Using an Efficient, Parallelizable
Adjoint Method".**

Podstawa prawna.

recenzję wykonano na zlecenie Przewodniczącego Rady Naukowej Dyscypliny Inżynieria Mechaniczna Politechniki Warszawskiej, prof. dra hab. inż. Roberta Sitnika, zgodnie z uchwałą Rady z dnia 07.06.2023, powołującej mnie na recenzenta w przewodzie doktorskim mgra inż. Pawła Maciąga.

1. Charakterystyka ogólna.

Sterowanie optymalne ma swą długoletnią historię, bowiem zawsze istniała i istnieć będzie chęć nie tylko jakiegokolwiek rozwiązania zadania, ale rozwiązania (sub-) (lokalnie-globalnie) optymalnego. I choć metody i sposoby opisu postępowania są znane dla różnorodnych klas systemów, to praktyczne (a nie akademickie) zadania przez nie generowane są zwykle bardzo trudne: wielowymiarowe i niemożliwe do rozwiązania analitycznego. Postęp technologiczny oferuje coraz tańszą moc obliczeniową, czy to w postaci wieloprocesorowych (wielordzeniowych) komputerów, czy wspomaganie kartami graficznymi (CUDA), czy wreszcie specjalizowanymi układami (RISC). Od wielu lat panuje paradygmat komputerów o architekturze von neumannowskiej z binarnym sposobem zapisu danych i obliczeń. Wprowadzenie systemów obliczeń opartych na niebinarnej logice technologicznie jest ciągle niedostępne, podobnie jak i praktyczne wielkoskalowe obliczenia kwantowe, które są nutą (raczej dalszej niż najbliższej) przyszłości. Zatem należy wykorzystywać obecne możliwości technologiczne oraz ich programistyczne wsparcie w celu analizy czy projektowania coraz bardziej złożonych układów praktycznych. W tym nurcie leży recenzowana praca, która dla klasycznego zadania sterowania optymalnego dla układu wielowymiarowego (układu powiązanych ciał o strukturze drzewiastej) i dynamice opisanej formalizmem hamiltonowskim oferuje przyspieszenie obliczeń przez ich zrównoleglenie. Propozycja dotyczy zarówno przyspieszenia klasycznego zadania dynamiki prostej jak i szybkiego liczenia gradientu funkcji jakości, służącego bądź do optymalizacji parametrów projektowych, bądź do wyznaczania optymalnych sterowań, gdy parametry te są wartościami wektorów sterowań w dyskretnych chwilach na horyzoncie sterowania.

Dysertacja leży na granicy dyscyplin wśród których Inżynieria Mechaniczna jest pierwszoplanowa.

2. Forma i kompozycja rozprawy.

Rozprawa doktorska jest napisana po angielsku i obejmuje 145 stron tekstu podzielonego na osiem rozdziałów, poprzedzonych spisem akronimów i oznaczeń oraz uzupełnionych trzema dodatkami zawierającymi materiał pomocniczy. Całości dopełnia bibliografia licząca 127

klasycznych bądź współczesnych pozycji z zakresu przedmiotu dysertacji (modelowanie układów wielocłonowych, optymalizacja, w tym badanie wrażliwości na parametry, równoleglenie obliczeń). Autor poprzedził bibliografię własnymi pracami współautorskimi, z których dwie są artykułami w czasopiśmie, a dziewięć pozostałych bądź jako rozdziały w materiałach pokonferencyjnych, czy też recenzowane abstrakty konferencyjne lub warsztatowe. Prace są datowane na ostatnie pięć lat.

W krótkim rozdziale 1 opisano motywację pracy, jej cel oraz rozkład materiału w rozdziałach. Rozdział 2. zawiera usytuowanie rozważanego problemu wśród innych zadań sterowania optymalnego, przedstawionym na tle literatury przedmiotu. Sformułowano tezę główną stanowiącą, że połączenie podejścia hamiltonowskiego opisu dynamiki układu wraz z metodą sprzężoną, *adjoint method*, w implementacji równoległej jest efektywnym i niezawodnym sposobem rozwiązania zadań dynamiki prostej i wyliczenia gradientu funkcji jakości. Pokazano jak uzasadnienie tezy będzie realizowane przez szereg wymienionych podcelów szczegółowych.

W rozdziale 3. pokazano jak w ujęciu hamiltonowskim wyrazić wzory na charakterystyczne dane konieczne do wyliczenia dynamiki prostej układu wielocłonowego o strukturze drzewiastej wyrażone w zmiennych globalnych i złączowych, gdy na układ nałożone są dodatkowo ograniczenia równościowe. Szczególną uwagę poświęcono dekompozycji obiektu rzeczywistego na coraz mniejsze obiekty wirtualne oraz transformację w odwrotnym kierunku stanowiącą esencję metody *dziel i zdobywaj*, która ma potencjał do równoleglenia obliczeń. Dzięki wykorzystaniu stałej struktury równań na etapie kompozycji/dekompozycji możliwe staje się łatwe przechodzenie między etapami układów wirtualnych przez modyfikację – zgodnie z odpowiednimi równaniami – jedynie współczynników równań, co zdecydowanie ułatwia algorytmizację i zrównoleglenie obliczeń.

Rozdział 4 poszerza rozważania o wprowadzenie funkcji jakości i przedstawienie jakie równania (różniczkowo-algebraiczne) generuje metoda sprzężona dla rozważanego układu wielocłonowego. Zapis równań podkreśla jak należy je implementować w strategii *dziel i zdobywaj*.

Celem rozdziału 5. było przetestowanie zaproponowanego paradygmatu na kilku wybranych układach mechanicznych dla zadań optymalnego doboru parametrów, optymalnego sterowania czy stabilizacji. Pokazano, że wzory działają poprawnie poprzez porównanie z metodami klasycznymi (metoda różnic skończonych).

Rozdział 6. analizuje różnorodne aspekty zrównoleglania obliczeń. Przedstawiono proponowane algorytmy w czytelnym pseudokodzie (symulującym język C) z jasno wydzieloną częścią zrównoleglaną. W kodzie wskazano także jakie procedury będą oceniane pod względem złożoności czasowej (start i stop timera). Dla łatwiejszej generacji układów wielocłonowych, jako obiekt badania zaproponowano model planarnego n -wahadła umieszczonego na tłumionej platformie jezdnej. Przeprowadzono testy uwzględniające zmienną liczbę członów układu (skomplikowanie układu) oraz liczbę wątków (możliwość sprzętowe zrównoleglenie obliczeń) przedstawiając wyniki szczegółowe oraz ich dyskusję.

W rozdziale 7. paradygmat proponowany w dysertacji poddano weryfikacji praktycznej dla generacji optymalnych trajektorii, które zostały następnie podane do układu sterowania jako etap sprzężenia wyprzedzającego, *feedforward*. Domknięcie pętli sprzężenia zwrotnego zrealizowano klasycznym regulatorem PD.

Ostatni, rozdział 8. zawiera podsumowanie głównych osiągnięć dysertacji ze specyfikacją jak poszczególne podzadania zostały zrealizowane. Po nim następują trzy dodatki stanowiące ilustrację lub uzupełnienie materiału przedstawionego w poprzednich rozdziałach.

Dysertacja jest świetnie napisana po angielsku (choć tu nie jestem ekspertem) z dbałością o formę i treść. Kolorowe ilustracje w sposób graficzny ułatwiają zrozumienie idei transfor-

macji i ich wyników. Autor dba, by najpierw pojawiały się zapowiedzi treści i ich rozłożenie w tekście (początki rozdziałów), a następnie ich właściwe rozwinięcie. W ten sposób czytelnik nie jest zaskakiwany i wie, gdzie znaleźć interesujący go fragment wywodu.

3. Wyniki uzyskane w rozprawie.

Za oryginalne osiągnięcia Autora uważam:

- Zręczne połączenie znanych technik: modelowania w paradygmacie hamiltonowskim (stabilniejsza numerycznie od lagranżowskiego podejścia), metody sprzężonej (*adjoint method*) (wyprowadzenie i uzasadnienie stosownych wzorów umożliwiających stosowanie metody) oraz techniki dziel i zdobywaj (skutecznej w polityce od czasów rzymskich – *divide et impera*) dla rozwiązania trudnego zadania modelowania i wyliczania gradientu funkcji wrażliwości układu wielu ciał połączonych w strukturę drzewiastą z nałożonymi ograniczeniami typu równościowego.
- Dla tej ostatniej metody pokazanie jak należy komponować z ciał fizycznych obiekty wirtualne, a potem jak z warunków brzegowych uzyskiwać w procesie odwrotnym (dekompozycja) z obiektów wirtualnych dane fizyczne (prędkości, momenty) w zmiennych absolutnych i mieszanych przegubowo-absolutnych. Wykorzystano ciekawe podejście umożliwiające, przez wprowadzenie portów wejściowych/wyjściowych ciał (fizycznych i wirtualnych), zachowanie stałej struktury równań opisujących transformacje charakterystyk dynamicznych. Zatem etapy kompozycji/dekompozycji bazują jedynie na odpowiednim redefiniowaniu współczynników równań bez zmiany ich struktury. Podejście takie umożliwiło zrównoleglenie obliczeń, a w konsekwencji znaczne ich przyspieszenie oraz uzyskanie dobrej skalowalności problemu.
- Algorytmizacja i implementacja zaproponowanych metod.
- Przeprowadzenie testów zarówno symulacyjnych (na kilku obiektach) jak i na obiekcie rzeczywistym, właściwie wykorzystując narzędzia, oraz stosując odpowiednie kryteria oceny uzyskanych wyników. Testy dobrano celowo, by uwypuklić poszczególne aspekty proponowanego podejścia w możliwie odseparowany od innych wpływów sposób. Szczególnie cenne są analizy efektywności wynikające ze zrównoleglenia obliczeń i objaśnienie spodziewanych i niespodziewanych efektów.

4. Kwestie do dyskusji, komentarze, uwagi polemiczne.

1. Użycie nawiasów $()$, $[]$ nie zawsze jest konsekwentne. Zwykle $[]$ grupują złożone wyrażenia, ale w (4.1) tak nie jest. W sekcjach dotyczących testów dane wektorowe przedstawiane są konsekwentnie w nawiasach $[]$.
2. Bardzo nieliczne literówki (np. 40^9), czy pominięcia (dt we wzorze (5.9)).
3. Stosowanie transpozycji we wzorach (4.15), (4.16) i dalej przy $\nabla_b J$ jest nieuzasadnione, zob. przyjęta konwencja str. 11 po wzorze (2.4).
4. 50^{15} Wydaje się, że wzór $\nabla_b J = (h_b - \xi^T f_b) \Delta t$ jest przedstawiony zbyt ogólnie i należy się domyślać co oznacza. Pewnie lepiej byłoby zapisać ten wzór jako:

$$\nabla_b J = \sum_{i=0}^{\lfloor t_f/\Delta t \rfloor} (h_b(i\Delta t) - \xi^T(i\Delta t) f_b(i\Delta t)) \Delta t.$$

5. Dla jakiej topologii układu ciał algorytm dziel i zdobywaj jest najbardziej/najmniej efek-

tywny (dla nieograniczonej, skończonej liczby wątków)?

6. Autor ze względu na obszerność (liczbę składowych i liczbę argumentów) wielu wzorów nie pisze argumentów stosowanych funkcji. Jednak w części wstępnej, gdy wzory są krótsze argumenty wypisuje. O ile nie mam wątpliwości np. w zapisie $h(t, y, \dot{y}, b)$ wskazującym, że część zmiennych zależy od czasu, np. $y(t), \dot{y}(t)$ o tyle we wzorze (2.2) explicite występuje czas i $y(t), \dot{y}(t)$. Czy jest to efekt zamierzony? Proszę ewentualnie podać przykład takich ograniczeń jawnie zależnych od czasu.

W bardzo wielu pracach z oznaczeniami zawsze występują tego typu problemy: czy zmienna, czy funkcja, czy wreszcie wartość funkcji w punkcie?

7. Stosowanie metody pierwszej (liniowej) wariacji daje warunki konieczne gradientu funkcji jakości (niezbędna przy optymalizacji parametrów lub wyznaczaniu sterowań optymalnych, gdy są wyrażone jako wartości w punktach kontrolnych), ale bazuje na założeniu, że wpływ wariacji wyższych rzędów jest pomijalnie mały. Jak dobierano dyskretyzację czasu (stały, zmienny przyrost) oraz czy (ewentualnie jak) sprawdzano wiarygodność braku (nieistotności) wpływu wariacji wyższych rzędów?
8. Dla minimalizacji liczby danych opisujących przestrzenne charakterystyki obiektów (położenie i orientacja) stosowano parametryzację kwaternionową (lekką redundantną, ale znacznie lepszą niż np. kąty RPY czy inne parametryzacje trójelementowe grupy obrotów $\mathbb{SO}_+(3)$ wprowadzające osobliwości i złe uwarunkowanie numeryczne). Wiadomo, że także kwaternionowa parametryzacja jest niejednoznaczna - dla parametryzacji Rodrigueza, z której wprost wynika kwaternionowa, mamy $((\alpha, \mathbf{v}) \simeq (-\alpha, -\mathbf{v}))$. Czy numerycznie były(-by) kłopoty z tą niejednoznacznością?
9. Czy rozważana była inna reprezentacja sterowań niż punktowe wartości połączone z wielomianową interpolacją niskostopniową? Przykładowo, zapis sterowań jako kombinacja liniowa pewnej bazy funkcyjnej $\phi_i(\cdot)$ (fourierowska, wielomianowa)

$$u^k(t) = \sum_{i=1}^{N_k} c_i^k \cdot \phi_i(t), \quad k = 1, \dots, n.$$

Wtedy optymalizowanymi parametrami są współczynniki kombinacji $\mathbf{c} = (c^1, \dots, c^N)$. Co uzyskujemy: mamy pewnie mniejszą liczbę zmiennych (niezależną od dyskretyzacji horyzontu czasu i równą $\sum_{i=1}^n N_k$), różniczkowalność sterowań za darmo dowolnego stopnia.

10. Autor w sekcji 6.4 przedstawia ciekawe (i nieoczekiwane wyniki) dotyczące skalowalności wskazując, że za pewne anomalie odpowiadają zasoby pamięciowe. Podejrzewam, że przy mniejszych zasobach pamięciowych procesory nie są wykorzystywane najbardziej efektywnie, czekając na transfery danych. Stąd pytania: a) Czy chodzi tu o pamięć cache, czy RAM (czy obie)? b) Czy można (symulować) zmienność zasobów pamięciowych (tj. fizycznie nie, ale może logicznie?) przez konfigurację przed początkiem obliczeń? c) Ciekawym problemem poznawczym jest jak dobierać zasoby (pamięć, liczba procesorów (wątków)), by obliczenia były najbardziej efektywne, dla ustalonej liczby ciał (i ich topologii).
11. Może warto było podkreślić używając terminu optymalny, że stosując technikę gradientową uzyskujemy rozwiązanie lokalnie optymalne (czasem mocno zależne od warunków początkowych). Pominięcie tego rozróżnienia ma uzasadnienie jedynie w przypadku optymalizacji funkcji unimodalnych, ale to jest bardzo rzadki przypadek praktyczny.
12. Czy można oszacować złożoność pamięciową proponowanego podejścia? Jakościowo: jak zależy od topologii układu, dyskretyzacji horyzontu czasu, rozważanego zadania (opty-

malne sterowanie, optymalny dobór parametrów konstrukcyjnych), liczby wątków? A może jak ilościowo (najgorszy przypadek)?

13. Czy istnieją heurezy ułatwiające dobór parametrów początkowych (b^{init}) dla układów rozważanych w pracy? Czy podczas badań symulacyjnych stosowano technikę multistartu, by (sprawdzić, czy uwiarygodnić), że rozwiązanie uzyskane jest globalnie optymalne?

Przedstawione uwagi, pytania czy komentarze mają charakter głównie polemiczny czy poznawczy i nie wpływają ani na prawdziwość tezy dysertacji (czy elementów składowych jej uzasadnienia) ani jej rangę naukową.

5. Podsumowanie.

Przydatność dysertacji dla nauk inżynierskich jest wieloraka. Przyspieszenie obliczeń układów wieloczłonowych pozwala analizować i projektować układy o większej liczbie stopni swobody. Obniżenie stopnia różniczkowego (wynikające ze stosowania formalizmu hamiltonowskiego) ogranicza błędy w wyższych pochodnych oraz pozwala na uniknięcie stosowania dodatkowych normalizacji, gdy niezmienniki (np. faktyczna wartość dla ograniczeń nominalnie zerowych) nie są zachowywane idealnie. Wydaje się, że gotowe wzory dla tego sposobu modelowania układu mogą być praktycznie 1:1 dokomponowane do bibliotek oprogramowania wspierającego obliczenia równoległe dla tego typu układów (a także w tym kierunku idzie informatyka, by ułatwić, przyspieszyć implementację własnych pomysłów - czego bardzo dobrym przykładem jest język Python z licznym zestawem bibliotek).

Tezę pracy przekonująco uzasadniono, zadania służące do jej weryfikacji zaimplementowano algorytmicznie i przeprowadzono ich badania symulacyjne, a także na jednym praktycznym obiekcie.

6. Konkluzja.

Po zapoznaniu się z rozprawą doktorską mgra inż. Pawła Maciąga wyrażam opinię, że jej Autor wniósł oryginalny wkład do dyscypliny Inżynieria Mechaniczna. Problemy naukowe sformułowane w dysertacji w sposób jasny i klarowny, a są one aktualne i ważne praktycznie. Autor wykazał, że opanował odpowiednio warsztat naukowy, zna literaturę przedmiotu, posiada wiedzę i sprawność zarówno na poziomie formułowania jak i rozwiązywania zadań optymalizacji sterowania czy parametrów dla układów wieloczłonowych. Wykazał kompetencję na etapie implementacji zaproponowanego paradygmatu i jej weryfikację praktyczną bazującą na właściwym doborze miar jakości. Autor zna odpowiednie narzędzia ułatwiające obliczenia i potrafi je efektywnie wykorzystać. Udokumentowany w rozprawie doktorskiej wkład Autora w rozwój wiedzy (aspekt poznawczy) i możliwości jej potencjalnego zastosowania (aspekt praktyczny) należy uznać za istotny.

Stwierdzam zatem, że **przedłożona do recenzji rozprawa spełnia z dużym nadmiarem wymagania zawarte w ustawie o stopniach i tytule naukowym oraz stawiam wniosek o dopuszczenie rozprawy doktorskiej mgr. inż. Pawła Maciąga do publicznej obrony. Biorąc pod uwagę przytoczone powyżej argumenty oraz doceniając jakość i liczbę publikacji, stawiam formalny wniosek o wyróżnienie rozprawy.**