

Recenzja pracy doktorskiej Mgra Bartosza Ambrożego Grenia

**pod tytułem „Topological analysis of biomolecules”
przygotowanej pod kierunkiem dr hab. Joanny Idy Sułkowskiej, prof. UW,
Centrum Nowych Technologii, Uniwersytet Warszawski**

1. Ogólna charakterystyka rozprawy

Rozprawa ma formę cyklu dziewięciu prac: ośmiu opublikowanych i jednego preprintu. Poprzedza je 48-stronicowe wprowadzenie, które porządkuje aparat teorii węzłów i referuje stan wiedzy o topologicznie złożonych biomolekułach. Praca łączy zagadnienia z dziedzin takich jak topologia i biologia strukturalna do opisu krajobrazu energetycznego biomolekuł.

Przedmiotem badań są biopolimery (białka i RNA) analizowane metodami topologicznymi: teorią węzłów i supłów. Kandydat nie tylko stosuje te metody, ale też je rozwija. Wykrywa i klasyfikuje nietrywialną topologię w strukturach biomolekularnych (węzły, supły, kajdanki typu Brunna) oraz buduje nowy aparat formalny – klasyfikację supłów algebraicznych. Aparat ten, choć motywowany zastosowaniami biologicznymi, ma wartość matematyczną niezależną od nich.

Tematyka wymyka się klasycznym granicom dyscyplin. Struktura przestrzenna białek i RNA zazwyczaj jest analizowana z punktu widzenia biologii strukturalnej, chemii i fizyki polimerów, co pozwala na zrozumienie krajobrazu energii swobodnej opisującej związki między strukturą a dynamiką biopolimeru, determinujący funkcję biologiczną. Nowym podejściem jest zastosowanie metody topologicznych, które pozwalają na powiązanie: cech topologicznych (obecność węzła, typ supła) na ograniczenia dynamiki (bariery energetyczne zwijania/rozwijania, stabilność kinetyczna oraz termodynamiczna, flustracja energetyczna).

Matematycznie najistotniejszą częścią dysertacji jest praca „Classification of algebraic tangles” (Greń, Sułkowska, Gabrovšek, arXiv:2504.06901, złożona do recenzji), w której Autor obliczeniowo rozwiązuje problem Conwaya z 1970 roku. Pozostałe osiem prac jest interdyscyplinarnych: wykorzystują aparat teorii węzłów jako narzędzie, nie jako przedmiot rozwoju matematyki. Poniższa recenzja koncentruje się na warstwie matematycznej, z uwzględnieniem wkładu w zastosowania biofizyczne w zakresie potrzebnym do całościowej oceny.

2. Ocena głównego wkładu matematycznego: klasyfikacja supłów algebraicznych

2.1. Kontekst i postawiony problem

Conway (1970) wprowadził algebrę supłów (ang. tangles) do systematycznej klasyfikacji węzłów i splotów. Sklasyfikował w pełni supły wymierne (ang. rational tangles) za pomocą ułamka łańcuchowego. Klasyfikacja klasy ogólniejszej, tzn. supłów algebraicznych (ang. algebraic tangles), pozostawała otwarta. Jedyną wcześniejszą próbą była praca Moriuchiego (2008). Praca ta obejmowała supły do 7 skrzyżowań i jak wykazuje Autor była niekompletna: pominęła cztery supły-mutanty.

Autor:

1. definiuje kanoniczną reprezentację drzewiastą (binarne drzewo wyrażeniowe) dla pierwszych supłów algebraicznych (ang. prime algebraic tangles) (Definicja 21),
2. podaje algorytm (Algorytm 1) sprowadzający dowolny supeł do tej postaci,
3. dowodzi (Lematy 7–9, Twierdzenie 5), że algorytm jest niezmienniczy względem czterech typów ruchów zachowujących izotopię: ruchów elementarnych, twist, flype i ring,
4. na tej podstawie tabuluje supły algebraiczne do 14 skrzyżowań, rozszerzając wcześniejszą klasyfikację z 201 do ponad 2 milionów reprezentantów orbit (a licząc klasy izotopii – do ponad 25 milionów, Tabela 2),
5. klasyfikuje grupy symetrii supłów jako podgrupy grupy $D_8 \times Z_2$ (Twierdzenie 6 i Tabele 4–8).

2.2. Ocena jakości matematycznej

Warstwa formalna jest staranna i poprawna. Definicje (supła, sumy i iloczynu supłów, ruchów zachowujących izotopię, postaci kanonicznej) są precyzyjne. Dowody lematów 1–9 (notacja potęgowa A^n , zachowanie ułamka uogólnionego względem obrotów i ruchów zachowujących izotopię, niezmienniczość algorytmu) są rzetelne i pełne. Dowód Lematu 8 dla ruchu „flype” proponuje cztery przypadki położenia węzła w drzewie i podaje kilka podprzypadków.

Reprezentacja supła jako drzewa binarnego, z lewymi liśćmi (supły wymierne) i prawymi (supły integralne en. integral tangles), jest elegancka i nowa (o ile mi wiadomo). Ma przewagę nad klasyczną notacją Conwaya (przecinkową) i pozwala wykryć błąd w klasyfikacji Moriuchiego (mutanty $3(2(20))$ i $2(3(20))$, Tabela 3), jest to konkretny, sprawdzalny rezultat.

Twierdzenie 6 wyklucza cztery generatory grupy $D_8 \times Z_2$ jako możliwe grupy symetrii supła, przez prostą obserwację o niezmienniczości uogólnionego ułamka. To ładny wynik strukturalny, uzupełniający klasyfikację kombinatoryczną.

2.3. Zastrzeżenie zasadnicze: status Hipotezy 1 a kompletność klasyfikacji

Główna wątpliwość matematyczna dotyczy relacji między Twierdzeniem 5 a Hipotezą 1 (podrozdział 3 preprintu): „The canonical form of a prime algebraic tangle is unique up to twist, flype, ring, and elementary moves”. Twierdzenie 5 dowodzi, że algorytm daje ten sam wynik dla supłów powiązanych tymi ruchami, a postać kanoniczna jest ich niezmiennikiem. To nie to samo, co dowód, że postać kanoniczna jest pełnym niezmiennikiem izotopii. Gdyby istniały pary supłów izotopijnych, ale niepowiązanych żadnym skończonym ciągiem tych czterech ruchów (czyli gdyby Hipoteza 1 była fałszywa) klasyfikacja przeszacowywałaby

liczbę klas izotopii: dwa supły o różnych postaciach kanonicznych mogłyby być tym samym supłem.

Dla supłów wymiernych analogiczne twierdzenie jest udowodnione (Kauffman–Lambropoulou, Twierdzenia 2–4). Dla klasy supłów algebraicznych (klasy ogólniejszej), pozostaje przypuszczeniem. To luka logiczna w opisie wyniku („I extended the old classification from 201 tangles to over 2 million.”): rezultat jest uzasadniony jako klasyfikacja „do rozważanego zbioru ruchów zachowujących izotopię”, nie jako ostateczna klasyfikacja topologiczna.

Nie umniejsza to wartości pracy. Rozwiązanie takiej hipotezy zwykle wymaga osobnych narzędzi (triangulacje kanoniczne, niezmienniki hiperboliczne – o nich Autor pisze we wstępie). Samo precyzyjne sformułowanie hipotezy jest wkładem naukowym. Podczas obrony proszę Autora o:

- doprecyzowanie, w jakim sensie należy rozumieć słowo „klasyfikacja” w podsumowaniu pracy (2.1, 2.2) wobec statusu Hipotezy 1,
- wskazanie częściowych argumentów (np. dla supłów o ograniczonej liczbie skrzyżowań, weryfikowalnych niezależnym niezmiennikiem, wielomianem Kauffmana, objętością hiperboliczną) za prawdziwością hipotezy w zakresie do 14 skrzyżowań.

3. Ocena pozostałych osiągnięć (zastosowania biofizyczne)

Pozostałe osiem stanowi spójny cykl prac mający na celu pokazanie zastosowania topologii do opisu biomolekuł, przekraczający typowe wymagania dla rozpraw doktorskich w Polsce. Nie są to prace stricte matematyczne.

Wszystkie osiem zostało opublikowanych w czasopismach recenzowanych, znajdujących się na liście filadelfijskiej (JCR) – w tym dwukrotnie w *Nucleic Acids Research*, czasopiśmie o szczególnie wysokim prestiżu, które też są bardzo dobrze cytowane, co dodatkowo podnosi wartość tego dorobku. Dwie prace pierwszoautorskie omawiam szerzej, w nich wkład metodologiczny Kandydata jest najłatwiej zweryfikować. Pozostałe sześć prac współautorskich streszczam krótko.

3.1. Prace, w których Kandydat jest pierwszym autorem

Poza preprintem matematycznym (rozdz. 2) Kandydat jest pierwszym autorem dwóch prac. Nie jest to moja tematyka, ale precyzja definicji i rygor wnioskowania łączą narzędzia matematyczne z dobrze prowadzoną nauką empiryczną – to oceniam.

Greń, Dąbrowski-Tumański, Niemyska, Sułkowska (2021), „Lasso proteins – unifying cysteine knots and miniproteins”, *Polymers* 13(22), 3988.

Pytanie badawcze: czy topologia lassa wiąże się z funkcją białka? Lasso jest to motyw topologiczny obserwowany w białkach, powstaje z zamkniętej pętli na której rozpięta jest minimalna powierzchnia, która to jest przecina przez końce białka. W zależności od ilości przecięć oraz sposobu (super coiling) definiuje się topologię danej struktury. W pracy została opracowana wydajna numerycznie i poprawna matematycznie procedura detekcji topologii lassa w białku.

Metodologia zastosowana w pracy jest przejrzysta i odtwarzalna: Kandydat przeanalizował wszystkie struktury PDB z wiązaniem kowalencyjnym poza łańcuchem głównym białka (rekord LINK), potraktował każde jako zamknięcie pętli laso, zredukował nadmiarowość zbioru przez ClustalΩ w 734 klastrach homologii ($\geq 40\%$ identyczności), zawężając struktury $\geq 95\%$ identyczne do jednego reprezentanta. Jasne progi i jednoznaczna, powtarzalna procedura – dobra praktyka empiryczna, spójna z dbałością o jednoznaczność klasyfikacji z pracy matematycznej.

Wyniki są prezentowane z ostrożnością interpretacyjną. 55% klastrów (405/734) ma jednego reprezentanta – złożone topologie laso są rzadsze, im bardziej skomplikowane; wynik zgodny z intuicją, ale sformułowany jako obserwacja statystyczna, nie twierdzenie. Preferencja strony przekłucia pętli słabnie z odległością od mostka – zgodnie z hipotezą o chiralności łańcucha jako przyczynie efektu, klasyfikowaną i potwierdzoną danymi. Zachowanie cystein w klastrach niejednoelementowych (329/734) wynosiła $\geq 80\%$ tylko w 54% przypadków – nie odbiegała od typowych białek. Kandydat jasno prezentuje ten wynik jako neutralny, bez nadinterpretacji na korzyść hipotezy funkcjonalnej (znaczenia dla funkcji biologicznej). Analiza 275 domen PFAM (laso złożone 35%, trywialne 24%, brak laso 41%) prowadzi do wyważonego wniosku: złożona topologia laso może być funkcjonalna (np. w białkach antymikrobowych), ale nie jest to reguła. Ten wynik jest zgodny z wcześniejszymi wynikami dla laso peptydów.

Greń, Antczak, Zok, Sułkowska, Szachniuk (2024), „Knotted artifacts in predicted 3D RNA structures”, PLOS Computational Biology 20(6).

Najbardziej polemiczna praca cyklu: krytyka jakości metod predykcji struktury RNA. Kandydat (osobiście odpowiedzialny za analizę węzłów) przeanalizował wszystkie modele z konkursu CASP15, wyodrębnił szkielet łańcucha (P-O5'-C5'-C4'-C3'-O3') i policzył wielomian Alexandera po zamknięciu łańcucha, jest to solidne zastosowanie klasycznego niezmiennika jako narzędzia diagnostycznego. Argument jest elegancki: naturalne RNA są praktycznie zawsze niewęzłowe (jeden wyjątek: trójlistnik w RydC, odkryty przez Kandydata w Article 6), więc węzeł w modelu komputerowym jest najpewniej artefaktem, nie sygnałem biologicznym.

Wyniki: metody ML generowały błędną topologię siedmiokrotnie częściej niż metody klasyczne (8.2% vs 1.2%) a splątania czterokrotnie częściej (16.2% vs 3.9%); tylko modele ML generowały węzły nieklasyfikowalne (2.5% vs 0%). Cele sztuczne miały dziesięciokrotnie więcej węzłów i pięciokrotnie więcej splątań niż naturalne. Węzły w naturalnych RNA to wyłącznie trójlistniki, a w sztucznych RNA są w połowie przypadków bardziej złożone. Czysty, systematyczny sygnał statystyczny, dobrze łączący się z głównym wątkiem matematycznym rozprawy: klasyczny niezmiennik węzłowy jako narzędzie kontroli jakości w predykcji struktury RNA metodami AI. Pytanie na obronę: jak często metoda probabilistycznego i bezpośredniego zamknięcia łańcucha dawały rozbieżne wyniki i czy przeprowadzono analizę czułości na wybór procedury zamknięcia?

Jak autor skomentuje wynik biorąc pod uwagę, iż kolejne prace pokazują iż istnieją bardziej złożone węzły w RNA, np. 5.2.

W obu pracach, jak w preprincie matematycznym, powtarza się ten sam rys: precyzyjne definicje, systematyczna redukcja i klasyfikacja danych, uczciwe raportowanie wyników

negatywnych. Ta konsekwencja metodologiczna między matematyką, statystyką strukturalną białek i walidacją metod obliczeniowych wzmacnia ocenę krytyki naukowej.

3.2. Pozostałe prace współautorskie

- Odkrycie pierwszego węzła w naturalnie występującej cząsteczce RNA (trójkąt 3₁ w RydC RNA), opublikowane w Journal of Molecular Biology; wynik podważający wcześniejsze przekonanie o braku węzłów w RNA.
- Narzędzia AlphaKnot i AlphaLasso, opublikowane w Nucleic Acids Research.
- Odkrycie węzła 6₃ (niebędącego „twist knotem”) w ludzkim proteomie oraz nieopublikowanej topologii typu Brunna (kajdaneq h₂₁#₃ h₂₁) w rodzinie białek ENPP.
- Konserwacja topologii węzłów i supłów w transporterach błonowych (rodziny LeuT, DAACS, CPA, NCX), opublikowana w Biophysical Journal. Kandydat obliczył wielomiany Alexandera dla wszystkich podłańcuchów w najbardziej złożonej topologicznie grupie białek (rodzina LeuT, supł S₃₄₃₁), algorytmicznie zidentyfikował rdzenie węzłów i pętle supłów, oraz sklasyfikował typy węzłów w strukturach z domeną PFAM PF01699 (nadrodzina NCX) przewidzianych przez AlphaFold.
- Pakiet Topoly, biblioteka Pythona łącząca metody obliczeniowe teorii węzłów (wielomiany Alexandera, HOMFLY-PT, Yamady, GLN) w spójne narzędzie dla społeczności naukowej.

We wszystkich pracach wieloautorskich Autor precyzyjnie opisuje własny wkład (Rozdział 3), co ułatwia ocenę zgodną z wymogami ustawowymi dla rozpraw w formie cyklu publikacji.

4. Uwagi redakcyjne i formalne

- Drobne usterki w prepryncie: „sums abd products” zamiast „and” (Definicja 5), literówka „Ljubjana” zamiast „Ljubljana” w afiliacji, „handeled” zamiast „handled” (dowód Lematu 9). Do poprawienia przed publikacją.
- Struktura rozprawy (wprowadzenie, cel i wnioski, streszczenia, pełne teksty) jest przejrzysta i ułatwia lekturę prac o zróżnicowanym charakterze.

5. Ocena spełnienia ustawowych przesłanek rozprawy doktorskiej (art. 187 ustawy)

Art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce wymaga, by rozprawa doktorska: (1) prezentowała ogólną wiedzę teoretyczną kandydata w dyscyplinie, (2) wykazywała umiejętność samodzielnego prowadzenia pracy naukowej, (3) stanowiła oryginalne rozwiązanie problemu naukowego. Odnoszę się do każdej przesłanki osobno.

5.1. Ogólna wiedza teoretyczna w dyscyplinie nauki fizyczne

Ocena: pozytywna. Wprowadzenie (rozdział 1) prezentuje szeroką wiedzę teoretyczną: aparat matematyczny teorii węzłów i supłów (algebra Conwaya, niezmienniki węzłowe –

wielomiany Alexandera, Jonesa, HOMFLY-PT, Yamady, GLN, klasyfikacja supłów wymiernych) oraz kontekst fizyczny i biologiczny (fizyka polimerów, biologia strukturalna białek i RNA, hipoteza węzłowego atomu Kelvina–Thomsona). Autor swobodnie łączy te rejestry, poprawnie osadza własne wyniki na tle stanu wiedzy (klasyfikacja Moriuchiego, twierdzenia Kauffmana–Lambropoulou) i rozumie zarówno dowody matematyczne, jak i metodologię obliczeniową biologii strukturalnej (predykcja struktur, bazy PDB/AlphaFold). To dojrzała, interdyscyplinarna wiedza teoretyczna – w warstwie matematycznej istotnie przekraczająca standardowe wymagania dla rozpraw w dyscyplinie nauki fizyczne.

5.2. Umiejętność samodzielnego prowadzenia pracy naukowej

Ocena: pozytywna. Trzy fakty potwierdzają ten wniosek. (a) Autor jest pierwszym autorem preprintu „Classification of algebraic tangles”, głównego wyniku matematycznego rozprawy, co wskazuje na wiodącą rolę w sformułowaniu problemu, konstrukcji aparatu dowodowego i klasyfikacji obliczeniowej. (b) W pozostałych ośmiu pracach Autor jasno opisuje własny wkład: w dwóch (p. 3.1) jako pierwszy autor, w sześciu jako współautor z wyodrębnioną rolą (analizy obliczeniowe, rozwój narzędzi Topoly, AlphaKnot, AlphaLasso, interpretacja wyników biologicznych) – to pozwala zweryfikować samodzielność mimo współautorstwa zespołowego. (c) Dorobek publikacyjny (8 artykułów w czasopismach z listy filadelfijskiej, w tym dwukrotnie w Nucleic Acids Research, plus preprint) rozłożony na kilka lat i kilka niezależnych linii badawczych (białka lassowe, węzły w transporterach (transbłonowe) błonowych, węzły w RNA, klasyfikacja supłów) świadczy o zdolności do samodzielnego inicjowania i prowadzenia badań naukowych. Nowe narzędzie formalne (reprezentacja drzewiasta supłów) i jego użycie do skorygowania błędu w klasyfikacji Moriuchiego (2008) jest przekonującym dowodem samodzielności naukowej w zagadnieniach matematycznych, które bardzo często są podstawowym aparatem w badaniach biofizycznych.

5.3. Oryginalne rozwiązanie problemu naukowego

Ocena: pozytywna, z zastrzeżeniem opisanym w p. 2.3. Za oryginalne rozwiązanie problemu naukowego uznaję przede wszystkim rozszerzenie klasyfikacji supłów algebraicznych (problem otwarty od Conwaya, 1970) z 201 do ponad 2 mln reprezentantów orbit do 14 skrzyżowań, wraz z nowym narzędziem formalnym (kanoniczna reprezentacja drzewiasta) i klasyfikacją grup symetrii supłów (Twierdzenie 6); jest to wynik nowy, sprawdzalny, częściowo zweryfikowany przez wykrycie i korektę błędu w jedynej wcześniejszej klasyfikacji tej klasy obiektów. Za dodatkowe oryginalne rozwiązania problemów naukowych uznaję: odkrycie pierwszego węzła w naturalnie występującej cząsteczce RNA (RydC), odkrycie nietrywialnej topologii typu Brunna w białkach ENPP oraz narzędzia obliczeniowe (Topoly, AlphaKnot, AlphaLasso) rozwiązujące praktyczny problem detekcji i klasyfikacji topologii biomolekuł w erze masowego przewidywania struktur (AlphaFold). Zastrzeżenie dotyczy wyłącznie precyzji sformułowania głównego wyniku matematycznego: teza o pełnej klasyfikacji „do izotopii” jest warunkowa względem nieudowodnionej Hipotezy 1. To nie podważa oryginalności rozwiązania, wymaga jednak doprecyzowania jego dokładnego zasięgu.

6. Konkluzja i wniosek końcowy

Rozprawa rozszerza teorię supłów Conwaya poza przypadek wymierny, z konkretnym, sprawdzalnym rezultatem (korekta klasyfikacji Moriuchiego, nowa klasyfikacja grup symetrii). Dorobek w zastosowaniach do biologii strukturalnej jest bogaty i dobrze udokumentowany, publikowany w renomowanych czasopismach z listy filadelfijskiej. Autor opanował warsztat matematyczny do samodzielnych badań na styku topologii i nauk przyrodniczych oraz umie jasno wyodrębnić własny wkład w pracach zespołowych.

Jedynе poważniejsze zastrzeżenie (p. 2.3): tekst powinien precyzyjniej rozgraniczać „klasyfikację do zbioru znanych ruchów zachowujących izotopię” (udowodnioną) od „klasyfikacji do izotopii” (zależną od nierozstrzygniętej Hipotezy 1). Kwestia do wyjaśnienia podczas obrony – nie podważa wartości ani oryginalności rozprawy jako całości.

Rekomendacja: rozprawa spełnia wymogi art. 187 ustawy Prawo o szkolnictwie wyższym i nauce. Wnoszę o dopuszczenie jej do publicznej obrony.

Jozef H.Przytycki



Washington, 12 sierpnia 2026

George Washington University

e-mail: przytyck@gwu.edu