

UNIVERSIDADE FEDERAL DO PARANÁ

THIAGO HENRIQUE DE ARAÚJO LEMOS

COMPARAÇÃO DE INVARIANTES DA TEORIA DOS GRAFOS NA
PREVISÃO DA ESTABILIDADE DE FULERENOS

CURITIBA

2015

THIAGO HENRIQUE DE ARAÚJO LEMOS

COMPARAÇÃO DE INVARIANTES DA TEORIA DOS GRAFOS NA
PREVISÃO DA ESTABILIDADE DE FULERENOS

Dissertação apresentada como requisito parcial
à obtenção do grau de Mestre. Programa
de Pós-Graduação em Informática, Setor de
Ciências Exatas, Universidade Federal do
Paraná.

Orientador: Prof. Dr. André Luiz Pires
Guedes

CURITIBA

2015

L556c

Lemos, Thiago Henrique de Araújo
Comparação de invariantes da teoria dos grafos na previsão da
estabilidade de fulerenos/ Thiago Henrique de Araújo Lemos. – Curitiba,
2015.
114 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas,
Programa de Pós-graduação em Informática, 2015.

Orientador: André Luiz Pires Guedes .
Bibliografia: p. 107-114.

1. Teoria dos grafos. 2. Fulerenos. 3. Moléculas - Estabilidade. 4.
Invariantes. I. Universidade Federal do Paraná. II. Guedes, André Luiz Pires.
III. Título.

CDD: 511.5



Ministério da Educação
Universidade Federal do Paraná
Programa de Pós-Graduação em Informática

PARECER

Nós, abaixo assinados, membros da Banca Examinadora da defesa de Dissertação de Mestrado em Informática, do aluno Thiago Henrique de Araújo Lemos, avaliamos o trabalho intitulado, “*Comparação de invariantes da teoria dos grafos na previsão da estabilidade de fulerenos*”, cuja defesa foi realizada no dia 10 de abril de 2015, às 10:00 horas, no Departamento de Informática do Setor de Ciências Exatas da Universidade Federal do Paraná. Após a avaliação, decidimos pela:

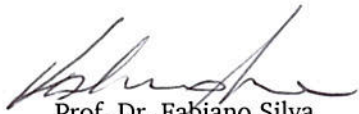
☒ **aprovação** do candidato. ☐ **reprovação** do candidato.

Curitiba, 10 de abril de 2015.


Prof. Dr. André Luiz Pires Guedes
PPGInf/UFPR – Orientador


Prof. Dr. Luerbio Faria
UERJ - Membro Externo


Prof. Dr. Marina Groshaus
UBA – Membro Externo


Prof. Dr. Fabiano Silva
PPGInf/UFPR – Membro Interno



AGRADECIMENTOS

Gostaria de agradecer, em primeiro lugar, à minha companheira, Bianca, por me ajudar nos momentos difíceis e por me entender como nenhuma outra pessoa seria capaz. Agradeço também ao orientador André Luiz Pires Guedes, que me apresentou a um tema tão interessante e me deu uma mão nas partes mais complicadas dessa empreitada que é a pesquisa científica. Agradeço aos meus pais, Ricardo e Marileide, e a Deus, por tudo que sou e fui capaz de alcançar. E também aos meus segundos pais, João e Luciana, que me acolheram tão bem e me fizeram sentir como se ainda estivesse em casa. Às minhas avós: Alzira, Zildete e Zélia, que sempre rezaram muito por mim. A vovô Afranio por ter atizado em mim o gosto pela leitura, e a vovô Delfino (in memoriam) pelas partidas de dominó. Agradeço a Nathália pelos pratos do Congo, e a Giovana pelos filmes do Lovecraft. Aos meus tios e tias, primos e primas, por todo o apoio que me deram nessa nova etapa da minha vida. Por fim, não posso esquecer dos amigos, meus irmãos que não são de sangue. Tanto os mais antigos, que me acompanham desde a época do Neves, do HK, da UFRN, e das aulas de Mihoko Sensei, quanto quem está embarcando agora comigo na lcoz. Muito obrigado por não terem me esquecido mesmo estando longe!

The beginning of wisdom is the statement 'I do not know.'
The person who cannot make that statement is one who
will never learn anything. And I have prided myself on
my ability to learn.

Thrall, Warchief of the Horde

RESUMO

Este texto descreve um projeto de mestrado que consiste em uma comparação de algumas invariantes da teoria dos grafos no contexto da previsão da estabilidade de moléculas de fulerenos. As invariantes investigadas incluem, entre outras, o critério de Fowler-Manolopoulos, o diâmetro, o índice de Wiener, a frustração bipartida de arestas, o número de independência, o número de emparelhamentos perfeitos, o número de Fries, e o número de Taylor. O objetivo principal aqui é computar seus valores para todos os isômeros de fulereno com até 130 vértices, e para todos os isômeros IPR com até 160 vértices. Até onde se sabe, ainda não foi feita nenhuma comparação experimental sobre a eficácia relativa dessas invariantes na previsão da estabilidade de fulerenos.

Palavras-chave: Grafos, fulerenos, estabilidade, invariantes.

ABSTRACT

This text describes a master's degree project that consists of a comparison of some graph theoretic invariants in the context of predicting the stability of fullerene molecules. Investigated invariants include, among others, the Fowler-Manolopoulos criterion, the diameter, the Wiener index, the bipartite edge frustration, the independence number, the number of perfect matchings, the Fries number, and the Taylor number. The main objective here is to compute their values for each fullerene isomer with up to 130 vertices, and each IPR isomer with up to 160 vertices. As far as is known, no experimental comparison has yet been made about the relative effectiveness of these invariants in predicting the stability of fullerenes.

Keywords: Graphs, fullerenes, stability, invariants.

LISTA DE FIGURAS

1.1	C_{60} Buckminsterfullereno (STRÖCK, 2006).	14
1.2	C_{540} , C_{70} e segmento de nanotudo de carbono (STRÖCK, 2006).	15
2.1	Exemplos de grafos completos (K_4 e K_5) e um bipartido completo ($K_{3,3}$).	21
2.2	Uma subdivisão do K_4	25
2.3	Exemplo de emparelhamento.	26
2.4	Um grafo planar G e um grafo plano H que é um desenho de G	27
2.5	Grafo correspondente Buckminsterfullereno (NONENMACHER, 2008).	29
2.6	Uma tentativa de construção do C_{22} que leva ao C_{24}	31
2.7	Espiral facial canônica do C_{60} Buckminsterfullereno.	32
3.1	Hexágono com índice de vizinhança 2. Pentágonos adjacentes em vermelho.	38
3.2	Grafo do 1,1,3-trimetil-ciclobutano, C_7H_{14} , adaptado de (GUTMAN et al., 1993).	40
3.3	Dois passos do algoritmo FKT. Árvore geradora T em vermelho, e H em azul. .	53
3.4	Exemplo de hexágono benzenóide (arestas do emparelhamento em destaque). ..	53
3.5	Hexágonos adjacentes s e t compartilham a variável X_{34}	55
4.1	C_{24} com vértices rotulados e a lista de vizinhos correspondente a cada vértice. .	65
4.2	Exemplo de pares ordenados e faces do programa <i>planar2dual</i>	67
4.3	Código para calcular o Critério de Fowler-Manolopoulos.	70
4.4	Código para calcular o Índice de Wiener e o Diâmetro.	70
4.5	Trecho do código que calcula a Frustração Bipartida de Arestas.	72
4.6	Código para calcular o Número de Independência.	73
4.7	Trecho de código para encontrar uma árvore geradora, T , de G	75
4.8	Trecho de código para calcular a árvore H	76
4.9	Fim da construção da orientação Pfaffiana e cálculo da invariante.	76
4.10	Primeira parte do código para calcular o número de Fries.	78
4.11	Segunda parte do código para calcular o número de Fries.	79
4.12	Terceira parte do código para calcular o número de Fries.	80
4.13	Criação das novas variáveis e restrições no cálculo do número de Taylor.	82
4.14	Última etapa do cálculo do número de Taylor.	83
5.1	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Diâmetro.	95
5.2	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Índice de Wiener.	95

5.3	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e a Frustração Bipartida de Arestas.....	95
5.4	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Independência*.....	95
5.5	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Estruturas de Kekulé.....	96
5.6	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Fries*.....	96
5.7	Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Taylor*.....	96
5.8	Coeficiente de Spearman entre o Diâmetro e o Índice de Wiener.....	96
5.9	Coeficiente de Spearman entre o Diâmetro e a Frustração Bipartida de Arestas.....	97
5.10	Coeficiente de Spearman entre o Diâmetro e o Número de Independência*.....	97
5.11	Coeficiente de Spearman entre o Diâmetro e o Número de Estruturas de Kekulé.....	97
5.12	Coeficiente de Spearman entre o Diâmetro e o Número de Fries*.....	97
5.13	Coeficiente de Spearman entre o Diâmetro e o Número de Taylor*.....	98
5.14	Coeficiente de Spearman entre o Índice de Wiener e a Frustração Bipartida de Arestas.....	98
5.15	Coeficiente de Spearman entre o Índice de Wiener e o Número de Independência*.....	98
5.16	Coeficiente de Spearman entre o Índice de Wiener e o Número de Estruturas de Kekulé.....	98
5.17	Coeficiente de Spearman entre o Índice de Wiener e o Número de Fries*.....	99
5.18	Coeficiente de Spearman entre o Índice de Wiener e o Número de Taylor*.....	99
5.19	Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Independência*.....	99
5.20	Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Estruturas de Kekulé.....	99
5.21	Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Fries*.....	100
5.22	Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Taylor*.....	100
5.23	Coeficiente de Spearman entre o Número de Independência* e o Número de Estruturas de Kekulé.....	100
5.24	Coeficiente de Spearman entre o Número de Independência e o Número de Fries*.....	100
5.25	Coeficiente de Spearman entre o Número de Independência e o Número de Taylor*.....	101

5.26	Coeficiente de Spearman entre o Número de Estruturas de Kekulé e o Número de Fries*.	101
5.27	Coeficiente de Spearman entre o Número de Estruturas de Kekulé e o Número de Taylor*.	101
5.28	Coeficiente de Spearman entre o Número de Fries e o Número de Taylor*.	101

LISTA DE TABELAS

1.1	Número de isômeros de fulereno teoricamente possíveis para alguns valores de n .	16
3.1	Índice de vizinhança (pentagonal) e valores de N_p para alguns isômeros.	36
3.2	Número de isômeros de fulereno teoricamente possíveis para alguns valores de n .	37
5.1	Critério de Fowler-Manolopoulos – isômeros estáveis.	90
5.2	Diâmetro – isômeros estáveis.	90
5.3	Índice de Wiener – isômeros estáveis.	90
5.4	Frustração Bipartida de Arestas – isômeros estáveis.	91
5.5	Número de Independência – isômeros estáveis.	91
5.6	Número de Estruturas de Kekulé – isômeros estáveis.	91
5.7	Número de Fries – isômeros estáveis.	92
5.8	Número de Taylor – isômeros estáveis.	92

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Contextualização	13
1.2	Objetivo e Contribuições	15
1.3	Organização	18
2	PRELIMINARES	19
2.1	Terminologia e Notação	19
2.2	Grafos de Fulereo	28
3	REVISÃO BIBLIOGRÁFICA	34
3.1	Regra dos Pentágonos Isolados (IPR).....	35
3.2	Critério de Fowler-Manolopoulos.....	38
3.3	Diametro e Índice de Wiener	39
3.4	Frustração Bipartida de Arestas.....	42
3.5	Número de Independência	45
3.6	Número de Independência <i>Closed-Shell</i>	47
3.7	Número de Estruturas de Kekulé	50
3.8	Número de Fries.....	53
3.9	Número de Taylor.....	56
3.10	Constante de Cheeger.....	59
4	PROJETO E IMPLEMENTAÇÃO	61
4.1	Escopo	61
4.2	Código Planar.....	64
4.3	Implementação	65
4.3.1	Critério de Fowler-Manolopoulos	66
4.3.2	Diâmetro e Índice de Wiener.....	68
4.3.3	Frustração Bipartida de Arestas	69
4.3.4	Número de Independência	71
4.3.5	Número de Estruturas de Kekulé.....	73
4.3.6	Número de Fries	75
4.3.7	Número de Taylor	81

5 EXPERIMENTOS E RESULTADOS	84
5.1 Experimentos	84
5.2 Resultados.....	85
5.2.1 Comparação com Valores da Literatura	86
5.2.2 Isômeros Estáveis.....	87
5.2.3 Análise Estatística	89
6 CONCLUSÃO	102
6.1 Sobre o trabalho desenvolvido	102
6.2 Dificuldades Encontradas.....	104
6.3 Oportunidades Futuras.....	105
REFERÊNCIAS	107

1 INTRODUÇÃO

1.1 Contextualização

Durante a maior parte do século XX as únicas substâncias conhecidas constituídas inteiramente por átomos de carbono eram o diamante, o grafite, e as formas de carbono amorfo. Mas, já em 1966, o químico e escritor David E. H. Jones (mais conhecido por seu personagem, o inventor fictício *Daedalus*) menciona em sua coluna de curiosidades científicas uma molécula de grafite em formato de “concha vazia” (JONES, 1966), e durante as décadas seguintes houve muita discussão sobre a possibilidade de uma estrutura molecular formada por 60 átomos de carbono baseada no icosaedro truncado (BOCHVAR e GAL'PERN, 1973; DAVIDSON, 1981; OSAWA, 1970).

O primeiro indício confirmando essas especulações teóricas surgiu num artigo de 1985, publicado por Kroto, Heath, O'Brien, Curl e Smalley, no qual é relatado um experimento de vaporização de grafite com um efeito colateral interessante: a produção, em quantidade significativa, de um agregado com exatamente 60 átomos de carbono (KROTO et al., 1985). Porém, naquele momento o grupo de Kroto não tinha ainda evidências suficientes de que esse agregado possuía a estrutura hipotética, e foi só em 1990 que Krätschmer, Lamb, Fostiropoulos, e Huffman apresentaram prova espectroscópica definitiva de sua organização em forma de icosaedro truncado (KRÄTSCHMER et al., 1990). Batizada de C_{60} : *Buckminsterfullereno*, em homenagem ao arquiteto Richard Buckminster Fuller (pela semelhança com suas cúpulas geodésicas), a estrutura foi declarada “Molécula do Ano” pela revista Science (KOSHLAND JR., 1991) e rendeu, em 1996, o Prêmio Nobel de Química a Kroto, Curl e Smalley. A Figura 1.1 a seguir mostra um modelo do Buckminsterfullereno.

Posteriormente, descobriu-se também que o Buckminsterfullereno ocorre naturalmente na fuligem (DAUGHERTY, 2009), acompanhado por outras estruturas semelhantes, mas com números atômicos diferentes. O termo simplificado *Fulereno* passou então a designar qualquer molécula composta exclusivamente por átomos terciários de carbono, ligados de maneira a formar uma superfície (semelhante a uma esfera, tubo ou elipsóide, exemplificados na Figura 1.2)

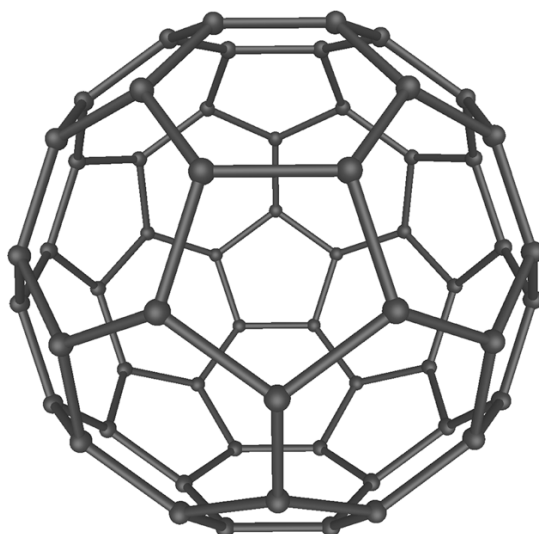


Figura 1.1: C_{60} Buckminsterfulereno (STRÖCK, 2006).

na qual apenas faces pentagonais e hexagonais podem aparecer. Uma abordagem abrangente sobre as propriedades teóricas dos fulerenos em geral pode ser encontrada em (FOWLER e MANOLOPOULOS, 2006). No caso particular das superfícies cilíndricas, as moléculas recebem a denominação especial de *nanotubos de carbono*, e são alvo de um campo de pesquisa próprio (DRESSELHAUS e AVOURIS, 2001).

Desde então, fulerenos foram encontrados em locais tão variados quanto rochas do período Pré-Cambriano (MOSSMAN et al., 2003) e nebulosas planetárias no espaço sideral (CAMI et al., 2010), um forte indicativo de que os fulerenos podem se formar de maneira eficiente até mesmo nas condições mais extremas. As propriedades químicas e físicas extraordinariamente peculiares das moléculas dessa nova família atraíram o interesse de pesquisadores das mais diversas áreas. Muitas aplicações práticas notáveis foram e continuam sendo encontradas. A lista abaixo enumera algumas delas (não sendo, porém, exaustiva):

- Armazenamento eficiente de hidrogênio em células de combustível (DILLON et al., 1997; ZHAO et al., 2005);
- Construção de células solares orgânicas (THOMPSON e FRÉCHET, 2008);
- Fabricação de semicondutores com precisão nanométrica (GIBBONS et al., 2008);

- Bloqueio de radicais livres e inflamações causadas por reações alérgicas (GOHO, 2007);
- Purificação de água e atividade antimicrobial (LI et al., 2008; TEGOS et al., 2005);

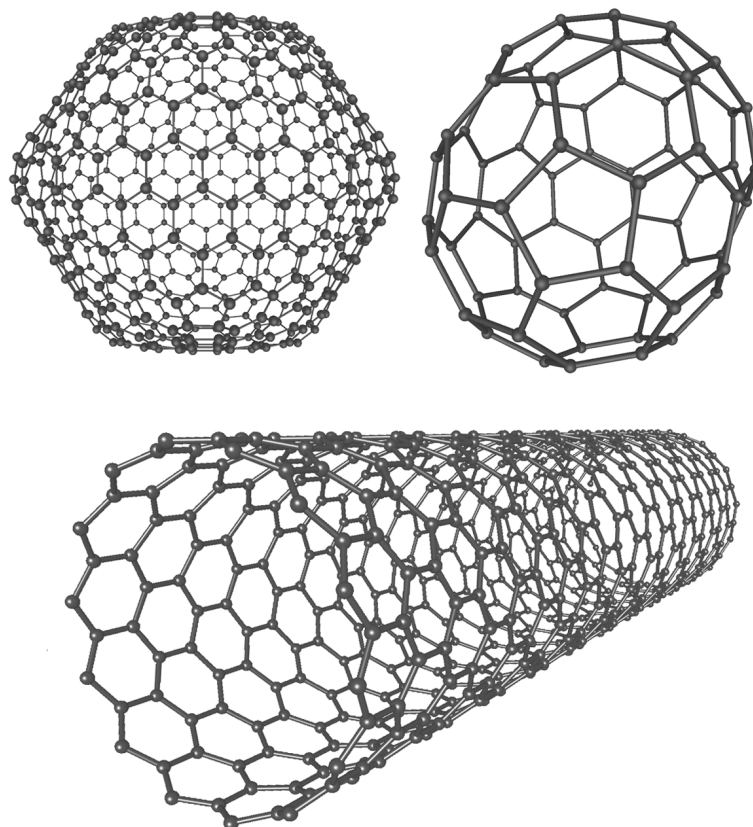


Figura 1.2: C_{540} , C_{70} e segmento de nanotudo de carbono (STRÖCK, 2006).

1.2 Objetivo e Contribuições

Um dos principais problemas encontrados pelos químicos quando vão lidar com fulerenos é o fato de a fórmula C_n não servir como identificador único para um fulereno com n átomos (ou seja, podem existir diversas moléculas matematicamente viáveis se encaixando na definição de C_n para um mesmo valor de n). Por exemplo, embora o Buckminsterfulereno seja denotado por C_{60} , sua estrutura é apenas uma entre as 1812 maneiras combinatoriamente distintas de dispor 60 átomos de carbono sem violar a definição de fulereno – as demais levam a *isômeros* do Buckminsterfulereno (isto é, moléculas que possuem a mesma fórmula química,

mas estruturas diferentes). Na verdade, certos resultados da área da matemática que estuda superfícies poliédricas (abstrações matemáticas úteis para modelar a estrutura dos fulerenos) levam a crer que existem fulerenos C_n para cada valor par de n maior ou igual a 20 (com a única exceção do 22 – um detalhe que será melhor explicado na Seção 2.2) (GRÜNBAUM e MOTZKIN, 1963), e o número de isômeros de C_n está em $\Theta(n^9)$ (DAUGHERTY, 2009; SAH, 1994). Essa taxa de crescimento surpreendentemente alta é ilustrada na Tabela 1.1 (dados retirados de (DAUGHERTY, 2009)).

Tabela 1.1: Número de isômeros de fulereno teoricamente possíveis para alguns valores de n .

Número de átomos, n	Número de isômeros de C_n
20	1
30	3
40	40
50	271
60	1.812
70	8.149
80	31.924
90	99.918
100	285.914
110	713.319
120	1.674.171
130	3.580.637
140	7.341.204
150	14.059.173
160	26.142.839
170	46.088.148
180	79.538.725
190	131.561.725
200	214.127.713

Entretanto, embora o número de isômeros possa ser bastante elevado para algum C_n em particular, apenas uma fração dessas estruturas é estável o suficiente para ser isolada experimentalmente com sucesso. Isto é, o fato de uma configuração atômica ser matematicamente possível não significa que ela leve a uma molécula quimicamente viável – no sentido de existir por tempo suficiente para ser detectada. A dificuldade surge, então, em saber *quais* isômeros são mais relevantes (ou seja, *estáveis*) entre a massa de possibilidades para cada C_n . Dito de

outra maneira: suponha que certo fulereno com n átomos é produzido, é possível saber com antecedência a qual isômero ele corresponde (i.e., qual é sua estrutura)? Uma das maneiras encontradas pelos pesquisadores de responder a essa pergunta se baseia na observação de que as propriedades químicas de uma molécula (e, conseqüentemente, sua estabilidade esperada) possuem uma relação próxima com certas invariantes no grafo molecular que a representa – um grafo no qual os vértices correspondem aos átomos e as arestas às ligações entre os átomos.

Essa técnica de avaliação de invariantes se provou particularmente adequada à previsão da estabilidade de fulerenos, já que a presença exclusiva de átomos de carbono simplifica a representação via grafo molecular. Isso motivou a investigação de diversas invariantes na literatura até o momento, mas é difícil determinar o quão eficazes elas realmente são. Geralmente, quando um grupo de autores propõe alguma nova invariante para prever a estabilidade de fulerenos, há pouca ou nenhuma preocupação em apresentar resultados experimentais comparando o desempenho do novo método com o de outras invariantes. E, mesmo nos casos em que são realizadas comparações, essas tendem a levar em conta apenas uma ou outra invariante mais tradicional. Logo, a principal contribuição deste trabalho consiste em apresentar, num único texto, alguma evidência da consistência (ou da discrepância) entre as estabilidades esperadas de acordo com cada uma das diversas invariantes selecionadas da literatura. Ou seja, investigar se essas invariantes estão “de acordo” quanto a quais são os isômeros mais estáveis. Para isso as invariantes foram implementadas (seus valores calculados em um conjunto de dados) e comparadas duas a duas conforme descrito na Subseção 5.2.3.

Além disso, os experimentos documentados na literatura costumam se restringir a valores não muito elevados de n (ficando em torno de 100 ou 120), visto que o cálculo da invariante – que pode consumir uma quantidade significativa de tempo dependendo do algoritmo usado – deve ser feito para *cada* isômero. Logo, outro objetivo deste trabalho é elevar o limite de maneira a abranger uniformemente as invariantes estudadas. Em particular, o cálculo é feito para isômeros com n até 160 (com algumas ressalvas – os detalhes são explicados no Capítulo 4). Por fim, o trabalho tem também o benefício de apresentar ao autor oportunidades de aprofundar-se em diversos tópicos da teoria dos grafos, de ganhar experiência com um tema

de pesquisa, e de estudar um problema bastante atual que fica na interseção entre áreas tão diversas quanto química, matemática e computação.

1.3 Organização

O Capítulo 1, que acaba de introduzir a proposta de trabalho, é o capítulo presente. A Seção 1.1 contextualizou o problema-alvo da proposta. A Seção 1.2 apresentou os objetivos e contribuições do trabalho, e esta Seção, 1.3, descreve como o restante do texto está organizado. O Capítulo 2 apresenta as definições necessárias para o entendimento do problema-alvo. O Capítulo 3 revisa os principais trabalhos que motivaram esta dissertação, enumerando algumas invariantes da teoria dos grafos e os algoritmos para calculá-las. O Capítulo 4 declara formalmente o problema-alvo da proposta e seu escopo, apresentando detalhes do projeto e mostrando trechos importantes do código construído. O Capítulo 5 documenta os experimentos realizados e os resultados obtidos, extraíndo algumas conclusões. Finalmente, o Capítulo 6 sintetiza os pontos mais importantes do texto, além de discutir sobre dificuldades encontradas e possíveis continuações deste trabalho.

2 PRELIMINARES

Neste capítulo são apresentadas a terminologia e a notação utilizadas ao longo do texto, além de uma série de definições necessárias para o pleno entendimento do problema-alvo do trabalho proposto. Em particular, a Seção 2.1 introduz conceitos elementares de Teoria dos Grafos relevantes para este trabalho, objetivando uniformizar a maneira como eles serão abordados. Já a Seção 2.2 trata de fulerenos, enfatizando a maneira como essas moléculas podem ser modeladas através de grafos e caracterizando o tipo específico de grafo com o qual esta proposta lida, além de discutir alguns resultados importantes relacionados com grafos de fulerenos. O conteúdo da Seção 2.1 é uma compilação adaptada do material encontrado em (BONDY e MURTY, 1976; DIESTEL, 2000; WEST, 2002; ZEEMAN, 1975).

2.1 Terminologia e Notação

Esta proposta trata da comparação de invariantes da teoria dos grafos quanto ao potencial para prever a estabilidade de uma classe específica de moléculas, denominadas fulerenos, formadas exclusivamente por átomos de carbono. Portanto, antes de entrar em maiores detalhes sobre fulerenos e invariantes, é necessário apresentar definições básicas relacionadas a grafos, tais como isomorfismo, grafos cúbicos, caminhos, k -conectividade, emparelhamentos, e planaridade. Este é o objetivo desta seção.

Definição 2.1.1. Um *grafo* é uma dupla ordenada $G = (V, E)$, onde V é um conjunto não vazio de *vértices* (ou *nós*) e $E \subseteq \binom{V}{2}$ é um conjunto de *arestas* disjunto de V . Note que não se presume nenhuma ordem entre os vértices de cada aresta (ou seja, para efeito de comparação $\{u, v\} = \{v, u\}$) e que essa definição não leva em conta arestas que estão associadas a um par de vértices iguais (i.e., $\{v, v\}$, também chamadas de *laços*). Podem existir vértices que não fazem parte de nenhuma aresta (denominados *vértices isolados*).

Definição 2.1.2. Um *grafo* $G = (V, E)$ é dito *finito* se, e somente se, V e E são ambos conjuntos finitos. Todos os demais grafos são considerados *infinitos*. A *ordem* e o *tamanho* de

um grafo G correspondem, respectivamente, aos valores $v(G) = |V|$ e $e(G) = |E|$ (ou seja, a cardinalidade do conjunto de vértices e a cardinalidade do conjunto de arestas de G).

Esta proposta lida apenas com grafos finitos. Então, de agora em diante, assume-se que o termo “grafo” é sempre empregado com o sentido de “grafo finito” e que os símbolos $v(G)$ e $e(G)$ denotam, respectivamente, o número de vértices e o número de arestas de um grafo G . Além disso, quando apenas um grafo estiver em discussão ou quando for óbvio a qual grafo o texto está se referindo, a dupla (V, E) será omitida e o grafo será tratado apenas por G .

Definição 2.1.3. Dado um grafo G , para cada aresta $e \in E$, tal que $e = \{u, v\}$, diz-se que:

- (i) A aresta e conecta u a v ;
- (ii) Os vértices u e v são os *extremos* de e ;
- (iii) Os vértices u e v são *adjacentes* ou *vizinhos*;
- (iv) Os vértices u e v são *incidentes* em e ;
- (v) A aresta e é *incidente* nos vértices u e v .

Opcionalmente, cada aresta $e = \{u, v\}$ de um grafo G pode estar associada a um número não negativo denominado *peso de e* e denotado pelo símbolo $w(e)$. Nesse caso, o grafo G (juntamente com os pesos nas arestas) é denominado *grafo ponderado* ou *grafo valorado*. Dependendo da aplicação, os pesos podem ser interpretados de diversas maneiras diferentes. Algumas das interpretações mais comuns são como a distância, o custo, ou a intensidade da ligação entre os dois vértices da aresta.

Definição 2.1.4. Dado um grafo G e um vértice $v \in V$, denota-se por $N_G(v)$ o conjunto de todos os vizinhos de v em G .

Definição 2.1.5. Dado um grafo G e duas arestas quaisquer $e, e' \in E$, diz-se que as arestas e e e' são *adjacentes* se, e somente se, $e \cap e' \neq \emptyset$ (i.e., se elas incidem sobre ao menos um vértice em comum).

Note que a definição utilizada para E não permite que duas ou mais arestas diferentes sejam incidentes sobre os mesmos dois vértices, caso em que as arestas envolvidas seriam denominadas *paralelas*. Arestas que não são paralelas são denominadas *simples*.

Definição 2.1.6. Um grafo G é dito ser *simples* se, e somente se, o grafo G não possui laços nem arestas paralelas, ou seja, se, e somente se, toda aresta de G for uma aresta simples.

Definição 2.1.7. Um grafo G é dito ser *completo* se, e somente se, todos os seus vértices são adjacentes dois-a-dois (i.e., para cada par de vértices $u, v \in V$, u é adjacente a v). O grafo completo com n vértices é denotado por K_n . O grafo com um único vértice e nenhuma aresta é considerado trivialmente completo e é denotado por K_1 .

Definição 2.1.8. Um *grafo bipartido*, G , é um grafo cujo conjunto de vértices, V , pode ser particionado em dois subconjuntos disjuntos X e Y , tal que cada aresta $e \in E$ possua um extremo em X e o outro em Y . Essa partição (X, Y) é denominada uma *bipartição* do grafo G . Se, além disso, cada vértice de X estiver ligado a cada vértice de Y por uma aresta $e \in E$, o grafo G é denominado um *grafo bipartido completo* e é denotado por $K_{m,n}$ (sendo $m = |X|$ e $n = |Y|$). A Figura 2.1 mostra alguns grafos completos e um bipartido completo.

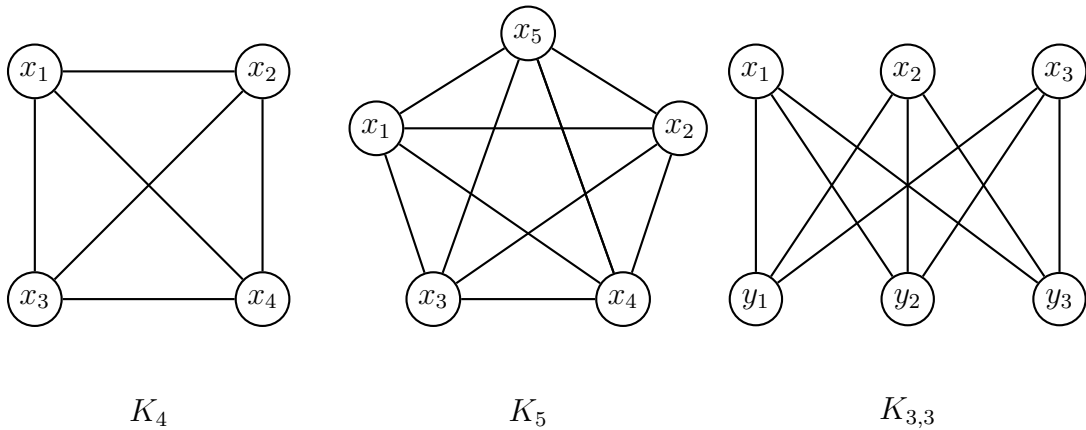


Figura 2.1: Exemplos de grafos completos (K_4 e K_5) e um bipartido completo ($K_{3,3}$).

Definição 2.1.9. Dois grafos $G = (V, E)$ e $H = (V_H, E_H)$ são ditos *isomorfos* (relação denotada por $G \simeq H$) se, e somente se, existe uma bijeção $\theta : V \rightarrow V_H$ tal que $\{u, v\} \in E$

se, e somente se, $\{\theta(u), \theta(v)\} \in E_H$. A bijeção θ é também denominada um *isomorfismo* entre G e H . Intuitivamente, o isomorfismo pode ser visto como um mapeamento (ou uma “renomeação”) entre os vértices dos dois grafos de modo que as arestas sejam preservadas.

Definição 2.1.10. Dados um grafo G e um vértice $v \in V$, o grafo obtido através da remoção do vértice v e de todas as arestas nele incidentes é denotado por

$$G - v = (V - \{v\}, E_v),$$

onde $E_v = \{e \in E \mid v \notin e\}$. De forma análoga, dados um grafo G e uma aresta $e \in E$, denota-se por

$$G - e = (V, E - \{e\})$$

Sendo que os vértices e outras arestas são preservados. Essas notações podem ser estendidas para considerar a remoção de um conjunto de vértices $X \subset V$ ou de um conjunto de arestas $Y \subseteq E$ de G da maneira a seguir:

$$G - X = (V - X, E_X),$$

onde $E_X = \{e \in E \mid v \notin e \forall v \in X\}$.

$$G - Y = (V, E - Y)$$

onde, novamente, os vértices e outras arestas (as fora de Y) são preservados.

Definição 2.1.11. Seja G um grafo. Então, para cada $u \in V$, o *grau* de u (também conhecido como *valência* do vértice u) é o número de arestas que incidem em u e é denotado por

$$\deg_G(u) = |\{e \in E \mid u \in e\}|.$$

Um vértice com grau zero é denominado um *vértice isolado*.

Definição 2.1.12. Um grafo G é dito k -regular se, e somente se, $\deg_G(v) = k$ para todo $v \in V$. Ou seja, quando todos os vértices possuem o mesmo grau, k . Um grafo é *regular* quando ele é k -regular para algum valor de k . Um grafo 3-regular é também chamado de *cúbico*.

Definição 2.1.13. Um *passeio* em um grafo G é uma sequência finita,

$$P = v_0 e_1 v_1 e_2 v_2 \cdots v_{n-1} e_n v_n,$$

tal que $n \in \mathbb{N}$ e onde os termos são alternadamente vértices e arestas de G . Ou, de maneira mais formal, onde $v_i \in V$, para todo $i \in \{0, \dots, n\}$ e onde $e_j \in E$ tal que $e_j = \{v_{j-1}, v_j\}$, para todo $j \in \{1, \dots, n\}$. Nesse caso, P é dito um passeio de v_0 para v_n ou, simplesmente, um *passeio- v_0-v_n* . Os vértices v_0 e v_n são denominados, respectivamente, a *origem* e o *destino* de P . O *comprimento* do passeio P , que corresponde ao valor n , é denotado por $|P|$. Quando $n = 0$, tem-se um *passeio nulo* de v_0 para v_n . Se $v_0 = v_n$ (e $n > 0$), então o passeio P é um *passeio fechado*; caso contrário, ele é um *passeio aberto*.

Finalmente, quando todas as arestas do passeio são distintas, o passeio é denominado uma *trilha*. Se, além disso, todos os vértices do passeio também forem distintos (com a possível exceção de v_0 , se o passeio for fechado) ele é denominado um *caminho*.

Trivialmente, considera-se o passeio nulo como um caminho nulo.

Definição 2.1.14. Dado um grafo G e dois vértices $u, v \in V$, a *distância* entre os vértices u e v , denotada por $d_G(u, v)$, é o comprimento do menor caminho- $u-v$ em G . De maneira mais formal:

$$d_G(u, v) = \text{mín. } \{|P| \mid P \text{ é um caminho-}u-v \text{ em } G\}$$

Se não existir nenhum caminho entre u e v , considera-se $d_G(u, v) = \infty$.

Definição 2.1.15. Seja G um grafo. Um *ciclo* em G é um passeio fechado em G no qual todas as arestas são distintas (ou seja, uma trilha fechada). O ciclo é *simplex* se, e somente se, ele for um caminho fechado (i.e., os vértices também são distintos). Um ciclo de comprimento

k , isto é, um ciclo com k arestas, é dito um k -ciclo. Um k -ciclo é considerado *par* ou *ímpar* de acordo com a paridade de k .

Definição 2.1.16. Dois vértices, u e v , em um grafo G são ditos *conectados* se, e somente se, existe um caminho entre u e v em G . Seja C_G uma relação binária sobre V tal que, para cada par de vértices, u e v , do grafo G , uC_Gv se, e somente se, u e v estão *conectados*. Note que C_G é uma relação de equivalência (reflexiva, simétrica e transitiva), e suas classes de equivalência recebem a denominação especial de *componentes conexas* de G . Um grafo G é *conexo* se, e somente se, ele possui uma única componente conexa (ou, de modo menos formal, se existe um caminho entre quaisquer dois vértices de G). Caso contrário ele é *desconexo*.

Definição 2.1.17. Um grafo G é dito k -conexo (para algum valor $k \in \mathbb{N}$) se, e somente se, $v(G) > k$ e $G - X$ é conexo para cada conjunto $X \subset V$ tal que $|X| < k$. Ou, de maneira menos formal, se é possível remover até $k - 1$ vértices (quaisquer que eles sejam) sem aumentar o número de componentes conexas de G e sem deixá-lo isomorfo ao K_1 . Dessa forma, todo grafo k -conexo é também, por definição, $(k - 1)$ -conexo, $(k - 2)$ -conexo, $\dots$, e assim por diante, até 0-conexo. Note que todo grafo não vazio é trivialmente 0-conexo, e que um grafo G é 1-conexo se, e somente se, o grafo G é conexo (i.e., seu número de componentes conexas é igual a 1) e G possui pelo menos dois vértices.

Definição 2.1.18. A *conectividade (de vértices)* de um grafo G , denotada por $\kappa(G)$, é o maior número tal que G seja k -conexo, $k \in \mathbb{N}$.

Definição 2.1.19. Seja G um grafo e $e = \{u, v\}$ uma aresta de G . Diz-se que e é *subdividida* quando ela é removida de G e substituída por um caminho P de comprimento 2 ligando u a v (sendo o vértice interno de P , x , um novo vértice que é inserido em G nesse processo). Uma *subdivisão* de G é um grafo que possa ser obtido de G a partir de uma sequência de subdivisões de aresta.

Algumas das invariantes que serão analisadas mais adiante neste trabalho envolvem a busca por emparelhamentos em grafos. Como o nome sugere, um *emparelhamento* envolve a

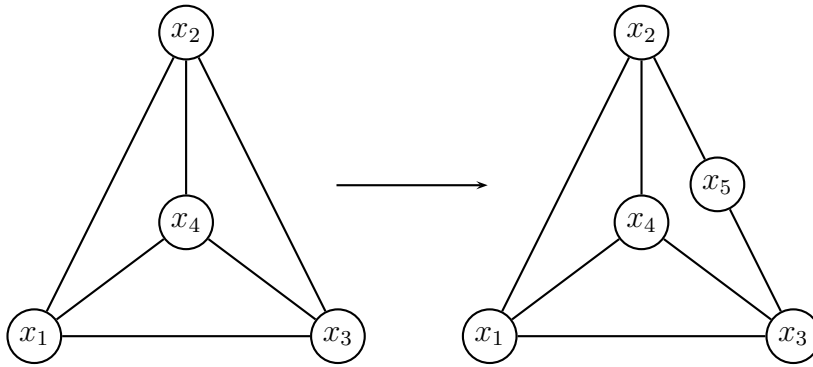


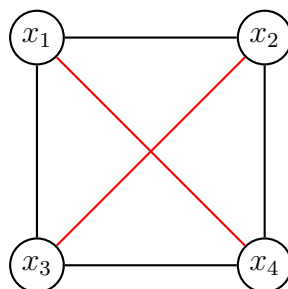
Figura 2.2: Uma subdivisão do K_4 .

formação de pares com os vértices do grafo, sendo que nenhum vértice pode fazer parte de mais de um par ao mesmo tempo. As definições a seguir explicam de maneira mais formal os conceitos relacionados.

Definição 2.1.20. Dado um grafo G , um *emparelhamento* $M \subseteq E$ em G é um subconjunto de arestas de E tal que cada vértice $v \in V$ seja incidente em, no máximo, uma aresta em M , ou, de maneira equivalente, tal que quaisquer duas arestas distintas em M não possuam extremos em comum. Vértices incidentes em alguma aresta de M são ditos *emparelhados* (ou *saturados com respeito a M*), e os demais são ditos *não emparelhados* (*não saturados* ou *livres*). A Figura 2.3 mostra um exemplo de emparelhamento (arestas destacadas em vermelho).

Como M é um conjunto de arestas, sua *cardinalidade*, $|M|$, corresponde ao número de arestas em M . Logo, uma maneira de construir um emparelhamento é escolhendo, iterativamente, arestas cujos extremos ainda não estejam saturados, até que nenhuma outra aresta possa ser adicionada. Um emparelhamento construído dessa maneira é definido por (WEST, 2002) como *maximal*, não sendo necessariamente de *cardinalidade máxima*, conforme a definição abaixo:

Definição 2.1.21. Seja G um grafo, seja $\mathcal{M}$ o conjunto de todos os emparelhamentos em G . Um emparelhamento M em $\mathcal{M}$ é de *cardinalidade máxima* se, e somente se, $|M| \geq |M'|$, para todo $M' \in \mathcal{M}$.



K_4

Figura 2.3: Exemplo de emparelhamento.

Definição 2.1.22. Seja $G = (V, E, st)$ um grafo. Um emparelhamento M em G é um *emparelhamento perfeito* se todos os vértices de G estão saturados (emparelhados) com respeito a M . O emparelhamento da Figura 2.3 é perfeito.

Observe que um emparelhamento ser perfeito implica diretamente que ele também seja um emparelhamento de cardinalidade máxima, mas o inverso nem sempre é verdade. Isto é, pode existir um emparelhamento M de cardinalidade máxima que não seja perfeito e, quando isso ocorre, o grafo em questão *não admite* um emparelhamento perfeito. Caso contrário, haveria um emparelhamento (o perfeito) de cardinalidade maior que $|M|$, o que contradiria o fato da cardinalidade de M ser máxima.

Note também que, no caso em que G é um grafo ponderado e M é um emparelhamento em G , define-se o *peso* (ou *custo*) de M como sendo a soma do peso de todas as arestas que fazem parte de M :

$$w(M) = \sum_{e \in M} w(e)$$

Para encerrar esta seção, são revisadas algumas das definições mais relevantes sobre grafos planares. Esses grafos foram caracterizados pela primeira vez em 1930 pelo matemático polonês Kazimierz Kuratowski (no que ficou conhecido como o teorema de Kuratowski), partindo de dois resultados intermediários sobre a não planaridade dos grafos completos K_5 e $K_{3,3}$ (cuas provas podem ser encontradas em (BONDY e MURTY, 1976)).

Intuitivamente, um *grafo plano* (ou *imerso no plano*) é um grafo G no qual cada vértice corresponde a um ponto em $\mathbb{R}^2$, cada aresta é um arco entre seus dois extremos, e o interior de uma aresta não contém nenhum vértice ou ponto de qualquer outra aresta (as arestas só se tocam nos extremos que possuem em comum). As regiões de $\mathbb{R}^2 \setminus G$ são denominadas as *faces* de G , e são subconjuntos abertos de $\mathbb{R}^2$ cujas fronteiras estão em G . Exatamente uma das faces de G é denominada de *face externa*: a face que não é delimitada. Isto é, dado um disco D suficientemente grande, que contenha todos os pontos de G , a face externa é a face que contém $\mathbb{R}^2 \setminus D$. Todas as outras faces são denominadas *faces internas* de G .

Definição 2.1.23. Um grafo G é denominado *planar* se, e somente se, G é isomorfo a algum grafo plano H . Nesse caso, o isomorfismo é chamado *imersão de G no plano* (ou *imersão planar de G*) e H recebe a denominação especial de *desenho de G* . Veja um exemplo na Figura 2.4.

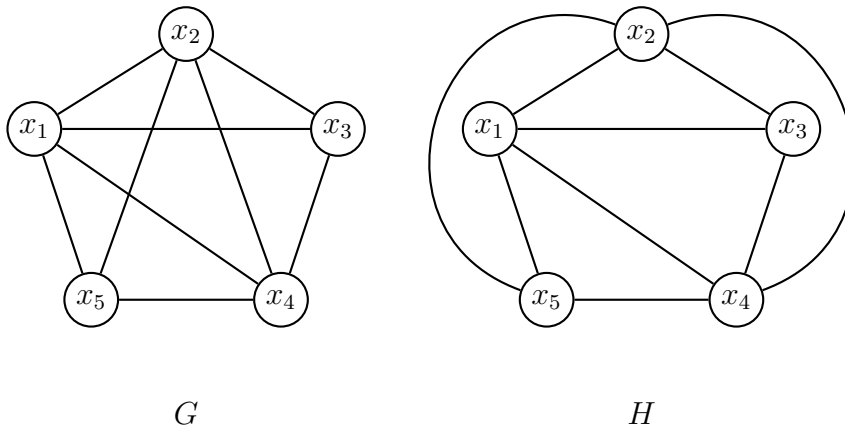


Figura 2.4: Um grafo planar G e um grafo plano H que é um desenho de G .

Definição 2.1.24. Dado um grafo plano G , é possível criar um novo grafo D como descrito a seguir: para cada face f de G , insira um vértice v em D ; se duas faces f_1 e f_2 são adjacentes em G , insira uma aresta e em D ligando os dois vértices correspondentes. Um grafo D obtido dessa forma é denominado *dual* de G .

Teorema 2.1.25 (Teorema de Kuratowski, 1930). Um grafo é planar se, e somente se, ele não contém nenhuma subdivisão do K_5 ou do $K_{3,3}$.

Pode-se encontrar uma prova para o teorema de Kuratowski em (BONDY e MURTY, 1976). Algumas conclusões relacionadas com essa prova são especialmente dignas de nota, em particular os dois lemas abaixo e o fato de os grafos K_5 e $K_{3,3}$ (ou grafos isomorfos a eles) serem os menores grafos não planares possíveis.

Lema 2.1.26. Se um grafo G não é planar, então cada subdivisão de G também não é planar.

Lema 2.1.27. Se um grafo G é planar, então cada subgrafo de G também é planar.

2.2 Grafos de Fulereno

Conforme mencionado na Seção 1.1, fulerenos são moléculas compostas exclusivamente por átomos terciários de carbono ligados de maneira a formar uma superfície fechada na qual apenas faces hexagonais e pentagonais podem aparecer. Essa superfície pode ser representada visualmente como um poliedro, no qual os vértices correspondem aos átomos de carbono, as arestas às ligações, e as faces aos anéis da molécula. De maneira equivalente, o poliedro correspondente a um fulereno pode ser representado através de um grafo planar com características específicas, o que leva à seguinte definição de grafo de fulereno (DAUGHERTY, 2009):

Definição 2.2.1. Um *grafo de fulereno*, denotado por C_n , é um grafo 3-regular planar, no qual $v(C_n) = n$ e cada face é formada por 5 ou 6 arestas. Note que, devido à 3-regularidade, C_n deve conter exatamente $3n/2$ arestas. A Figura 2.5 mostra o grafo de fulereno correspondente ao Buckminsterfulereno.

Novamente de acordo com (DAUGHERTY, 2009), o fato de os grafos de fulerenos serem 3-conexos implica que cada grafo possui uma imersão única no plano (preservando-se as fronteiras das faces). Partindo dessa definição, é possível recorrer a uma fórmula amplamente conhecida que relaciona o número de vértices, arestas, e faces em um grafo plano e conexo para obter alguns resultados interessantes. Estabelecida em 1758 pelo matemático e físico suíço Leonhard Euler (EULER, 1758), a fórmula se aplica aos grafos planares definidos pelos vértices e arestas de um poliedro convexo, como é o caso dos grafos de fulereno, e é enunciada a seguir (BONDY e MURTY, 1976).

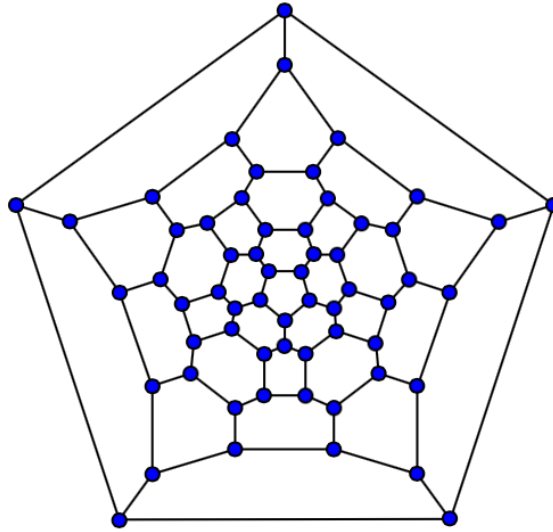


Figura 2.5: Grafo correspondente Buckminsterfulereno (NONENMACHER, 2008).

Teorema 2.2.2 (Fórmula de Euler). Seja G um grafo plano e conexo, e sejam os números de vértices, arestas e faces de G denotados, respectivamente, por $v(G)$, $e(G)$, e $\phi(G)$. Então:

$$v(G) - e(G) + \phi(G) = 2$$

Com a ajuda da fórmula acima é fácil observar que o número de faces em um grafo de fulereno, C_n , é $\phi(C_n) = v(C_n)/2 + 2$, que é o mesmo para qualquer grafo plano, conexo, e 3-regular (FOWLER e MANOLOPOULOS, 2006). Entretanto, é possível refinar ainda mais esse resultado ao considerar separadamente o número de faces pentagonais, $p(C_n)$, e o número de faces hexagonais, $h(C_n)$. Obviamente, a soma desses dois números é igual ao número total de faces, e, portanto, $p(C_n) + h(C_n) = v(C_n)/2 + 2$. Reservando essa equação, considere agora a soma de todas as arestas em cada face de C_n (5 para cada face pentagonal e 6 para cada face hexagonal). Como cada aresta faz parte de exatamente duas faces, essa soma será igual a duas vezes o número total de arestas, ou seja:

$$5p(C_n) + 6h(C_n) = 2e(C_n)$$

Novamente valendo-se da fórmula de Euler, é fácil chegar ao resultado $v(C_n) = 2e(C_n)/3$,

no qual é possível fazer uma substituição, obtendo $v(C_n) = (5p(C_n) + 6h(C_n))/3$. Juntando essa equação com a reservada anteriormente, têm-se o seguinte sistema linear:

$$p(C_n) + h(C_n) = v(C_n)/2 + 2 \quad (2.1)$$

$$v(C_n) = (5p(C_n) + 6h(C_n))/3 \quad (2.2)$$

Cuja solução é $p(C_n) = 12$ e $h(C_n) = v(C_n)/2 - 10$. Portanto, cada grafo de fulereno C_n possui exatamente 12 faces pentagonais e $n/2 - 10$ faces hexagonais (FOWLER e MANOLOPOULOS, 2006). Essa característica os diferencia de outros poliedros trivalentes e os coloca num subconjunto dos poliedros mediais de Goldberg (GOLDBERG, 1935, 1937). Porém, de modo similar a todos os outros poliedros trivalentes (nos quais o número de arestas deve ser um inteiro e é definido pela relação $e(C_n) = 3n/2$), *não* existem grafos de fulereno com número ímpar de vértices: n deve obrigatoriamente ser par. Além disso, o menor grafo de fulereno que é possível construir respeitando as equações acima é o dodecaedro, que possui 20 vértices, 12 faces pentagonais, e nenhuma face hexagonal. Colocando as duas afirmações juntas, conclui-se que grafos de fulereno, C_n , só podem existir para valores pares de n maiores ou iguais a 20.

Na verdade, existe ao menos um grafo de fulereno para cada um desses possíveis valores, excetuando-se $n = 22$ (FOWLER e MANOLOPOULOS, 2006; GRÜNBAUM e MOTZKIN, 1963). Embora essa impossibilidade pareça não fazer muito sentido de imediato, pelas equações acima o hipotético C_{22} deveria ter 12 faces pentagonais (como todos os outros grafos de fulereno) e uma única face hexagonal. Considere uma imersão do C_{22} no plano de modo que essa face hexagonal fique no centro e as outras 12 faces estejam agrupadas ao seu redor. Como é fácil enxergar na Figura 2.6, a única maneira de “fechar” as três últimas faces pentagonais seria acrescentando dois vértices ao grafo, causando o surgimento de mais uma face hexagonal (a externa) e resultando no grafo de fulereno C_{24} .

É preciso também apresentar um esquema que permita caracterizar de maneira única cada isômero de fulereno e seu grafo correspondente, algo útil para evitar confusões quando grafos

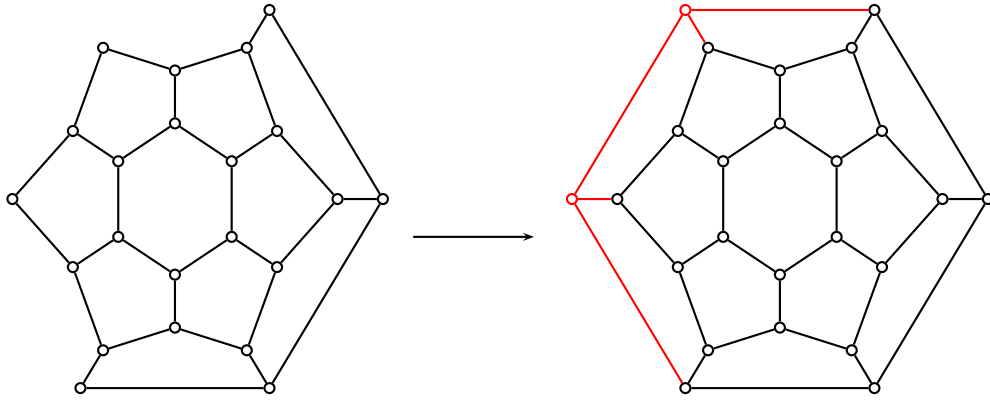


Figura 2.6: Uma tentativa de construção do C_{22} que leva ao C_{24} .

específicos estão sendo considerados. Conforme visto na Seção 1.1, o número de isômeros para um fulereno C_n está em $\Theta(n^9)$ (DAUGHERTY, 2009; SAH, 1994), uma taxa de crescimento considerável que é exemplificada na Tabela 1.1. A notação C_n serve como esquema de identificação para grafos de fulereno apenas nos três casos mais simples (aqueles com n igual a 20, 24, ou 26), nos quais o número de isômeros é igual a 1. Em todos os outros casos essa notação é ambígua, pois existem múltiplos isômeros e, para cada um deles, há um grafo de fulereno distinto. A notação mais popular atualmente (e que fornece uma caracterização única para cada fulereno) depende de um conceito denominado *espiral facial* (do inglês *face spiral*), definido a seguir (adaptado de (DAUGHERTY, 2009; FOWLER e MANOLOPOULOS, 2006)):

Definição 2.2.3. Seja G um grafo de fulereno e H um desenho de G no plano. Uma *espiral facial* em H é uma ordenação das faces tal que cada face na espiral compartilha uma aresta com seu predecessor e sucessor imediatos, e cada face na espiral depois da segunda compartilha uma aresta com: i) seu predecessor imediato na espiral e ii) com a primeira face na espiral precedente que ainda possui uma aresta não compartilhada. Uma *sequência de espiral facial* (ou *sequência espiral*) de G é uma lista com doze valores inteiros distintos que correspondem às posições dos doze pentágonos na espiral facial (começando a contagem na posição zero). A Figura 2.7

mostra um exemplo de espiral facial com sequência $(0, 6, 8, 10, 12, 14, 17, 19, 21, 23, 25, 31)$.

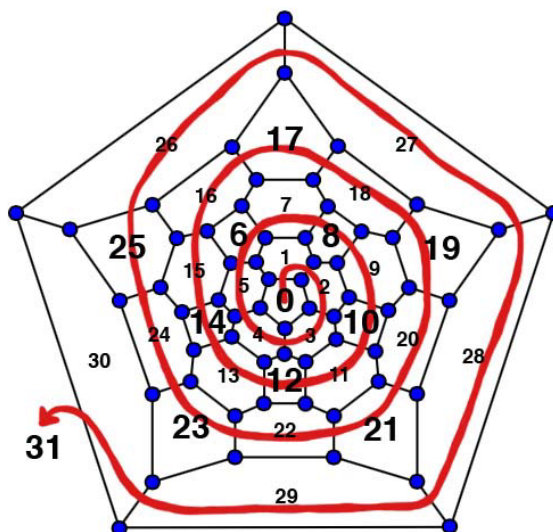


Figura 2.7: Espiral facial canônica do C_{60} Buckminsterfullereno.

Note que a espiral facial de um grafo G depende de seu desenho H (ou seja, da maneira como G está imerso no plano). Diferentes imersões vão levar a espirais faciais distintas, cada uma com sua própria sequência espiral. Para resolver este problema de multiplicidade, basta considerar como a *espiral canônica* do isômero aquela cuja sequência é a lexicograficamente menor. Essa espiral canônica e a correspondente sequência espiral servem para identificar de maneira única um isômero de fullereno C_n , de forma que é possível ordenar os isômeros (lexicograficamente) com base em suas sequências espirais (canônicas) e atribuir a cada um deles um número com base em sua posição na ordenação.

A notação para identificar um isômero específico pode ser finalmente definida como $C_n:k$, onde C_n é a fórmula química de um fullereno com n vértices e k é sua posição na ordenação descrita acima, começando a contagem em 1. Por exemplo, o Buckminsterfullereno pode ser identificado como $C_{60}:1812$, pois sua sequência espiral canônica é a última (lexicograficamente) entre os 1812 isômeros.

Embora inicialmente tenha se conjecturado que *todos* os fullerenos possuíam uma espiral facial, já se sabe que esse não é o caso nem para os fullerenos nem para a classe mais geral de poliedros cúbicos (FOWLER, JOOYANDEH e BRINKMANN). Entretanto, como o

menor fulereno conhecido que não possui uma espiral facial tem 380 vértices (FOWLER e MANOLOPOULOS, 2006), a notação apresentada acima é seguramente aplicável na prática.

Resta agora fazer um breve comentário sobre o uso da palavra *estabilidade* quando se estiver caracterizando um isômero de fulereno neste trabalho (i.e., “isômero mais estável”, “isômero menos estável”). Na literatura sobre o assunto, geralmente esse termo é empregado de forma ambígua, algumas vezes sendo usado com o sentido de não apresentar reatividade muito grande com o ambiente no qual se encontra, em outras com o sentido de estabilidade termodinâmica (significando que um sistema está em seu estado mais baixo de energia), e em outras ainda com o sentido de estabilidade cinética.

Embora esses sentidos estejam de certa forma relacionados, aqui a expressão *isômero estável* vai ser usada para designar especificamente os isômeros que foram ou que *teoricamente* podem ser produzidos em quantidade suficiente para ser observados. Esse “*teoricamente*” é o fator mais importante na questão, já que garantias teóricas a respeito da estabilidade podem, um dia, servir como guia para a síntese artificial de moléculas através de diagramas (FAJTOLOWICZ e LARSON, 2003). Logo, um isômero é *mais* (respectivamente, *menos*) estável do que outro se ele é (teoricamente ou na prática) *mais* (respectivamente, *menos*) facilmente observável em laboratório.

3 REVISÃO BIBLIOGRÁFICA

Este capítulo revisa alguns dos principais trabalhos que motivaram a presente dissertação de mestrado. Cada um desses trabalhos investiga uma ou mais invariantes da teoria dos grafos quanto à capacidade de prever a estabilidade de moléculas de fulereno. A análise feita pelos autores geralmente consiste em: i) mostrar que existe uma relação entre a invariante e alguma propriedade química desejável ou não desejável na molécula (casos em que se procura, respectivamente, maximizar ou minimizar o valor calculado), e ii) explicar como essa propriedade pode impactar a estabilidade esperada do fulereno.

Note que, com a intenção de tornar a discussão a seguir mais clara e focada no aspecto computacional, a maioria dos detalhes e termos da Química será omitida e o estudo será centralizado nas invariantes correspondentes. Porém, breves menções não podem deixar de ser feitas em alguns casos, e ao leitor mais curioso e interessado no lado químico da questão recomenda-se seguir as referências fornecidas em cada seção. O objetivo aqui é apenas fornecer um panorama geral, documentando (em uma ordem que favoreça o entendimento) quais foram as invariantes investigadas.

Em particular, a Seção 3.1 trata do primeiro critério estrutural utilizado na previsão da estabilidade de fulerenos, a regra dos pentágonos isolados. Um aprimoramento sobre essa regra leva à invariante que, ainda hoje, é considerada como uma das mais relevantes e precisas, se não a mais relevante de todas: o critério de Fowler-Manolopoulos, analisado na Seção 3.2. Já a Seção 3.3 lida com duas invariantes bastante relacionadas entre si e que dependem das distâncias entre os vértices do grafo de fulereno: o diâmetro e o índice de Wiener. Por sua vez, a Seção 3.4 estuda a frustração bipartida de arestas, uma invariante que mede o quão próximo de bipartido é o grafo de fulereno.

A Seção 3.5 revisa uma invariante já bastante estudada na teoria dos grafos, o número de independência. Isso serve como preparação para o assunto da Seção 3.6, que revisa não apenas uma, mas duas invariantes: o número de independência *closed-shell* e o menor autovalor da matriz de adjacência. As Seções 3.7, 3.8, e 3.9, apresentam, respectivamente, três

invariantes intimamente relacionadas com os emparelhamentos em um fulereno: o número de emparelhamentos perfeitos (também conhecido como contagem de estruturas de Kekulé), o número de Fries (uma constante calculada sobre os emparelhamentos perfeitos), e o número de Taylor (uma versão mais restrita do número de Fries). Por fim, a Seção 3.10 lida com uma invariante ainda não muito conhecida na literatura sobre fulerenos, denominada constante de Cheeger.

3.1 Regra dos Pentágonos Isolados (IPR)

Proposta por Kroto e Schmalz (KROTO, 1987; SCHMALZ et al., 1988), a *Regra dos Pentágonos Isolados* (do inglês *IPR – Isolated Pentagon Rule*) foi o primeiro critério de sucesso em determinar a estabilidade de fulerenos. Em termos químicos, essa regra se baseia na observação de que adjacências entre faces pentagonais causam um aumento no que é conhecido como *tensão estérica* da molécula (ou *tensão de Van der Waals*). Intuitivamente, isso implica que um isômero com menos pentágonos adjacentes é mais estável do que um com mais, ou seja, um isômero no qual essas adjacências sejam minimizadas deve, teoricamente, ser o mais estável. Para representar de maneira mais formal esse conceito, Raghavachari (RAGHAVACHARI, 1992) propôs um esquema de indexação, o *índice de vizinhança* (pentagonal), que permite quantificar o número de adjacências pentagonais em um isômero.

Aqui será apresentada a versão ligeiramente modificada (e mais intuitiva) desse índice, proposta por Fowler e Manolopoulos (FOWLER e MANOLOPOULOS, 2006), na qual o índice de vizinhança de um pentágono corresponde ao número de outros pentágonos aos quais ele é adjacente. Dessa forma, é possível atribuir a cada isômero uma assinatura na forma $\langle p_0, p_1, p_2, p_3, p_4, p_5 \rangle$, onde p_i é o número de pentágonos com índice de vizinhança i . Por exemplo, todos os pentágonos no C_{20} possuem índice de vizinhança igual a 5, o que resulta na assinatura $\langle 0, 0, 0, 0, 0, 12 \rangle$. Já o C_{60} Buckminsterfulereno possui assinatura $\langle 12, 0, 0, 0, 0, 0 \rangle$, já que todos os pentágonos possuem índice de vizinhança igual a zero (i.e., estão isolados).

Como o número de pentágonos é sempre igual a 12, para qualquer fulereno é óbvio que

$$\sum_{i=0}^5 p_i \leq 12$$

Entretanto, isso não significa que todos os 12 pentágonos em um fulereno tenham que ter o mesmo índice sempre. Na verdade, as assinaturas podem variar bastante de isômero para isômero, como pode ser visto na Tabela 3.1 (dados retirados de (FOWLER e MANOLOPOULOS, 2006)), que mostra as assinaturas de alguns isômeros – os mais estáveis para o valor mencionado de n . Note que a tabela também mostra um número, N_p , que corresponde ao número total de adjacências pentagonais em cada isômero e que pode ser considerado como o valor da invariante (se a regra dos pentágonos isolados for tomada como uma invariante no grafo do fulereno). Para calculá-lo, basta fazer a seguinte soma:

$$N_p = \sum_{i=0}^5 i p_i / 2$$

Tabela 3.1: Índice de vizinhança (pentagonal) e valores de N_p para alguns isômeros.

n : isômero	Índice de Vizinhança	N_p
20:1	0 0 0 0 0 12	30
24:1	0 0 0 0 12 0	24
26:1	0 0 0 6 6 0	21
28:2	0 0 0 12 0 0	18
30:3	0 0 2 10 0 0	17
36:14	0 0 12 0 0 0	12
40:38	0 4 8 0 0 0	10
42:45	0 6 6 0 0 0	9
50:271	2 10 0 0 0 0	5
54:540	6 4 2 0 0 0	4
56:916	4 8 0 0 0 0	4
60:1812	12 0 0 0 0 0	0
62:2378	6 6 0 0 0 0	3
68:6328	8 4 0 0 0 0	2
70:8149	12 0 0 0 0 0	0

Dessa forma, o valor de N_p pode ser utilizado como um critério bastante preciso para filtrar os isômeros mais estáveis (isto é, os que minimizam o valor de N_p), tendo se mostrado consistente com as observações experimentais sobre estabilidade. Por exemplo, para $n < 70$, o único fulereno com $N_p = 0$ é o Buckminsterfulereno, que é, conforme mencionado anteriormente, o mais abundante na natureza. Para $n \geq 70$, todos os fulerenos possuem ao menos um isômero com o índice ideal $\langle 12, 0, 0, 0, 0, 0 \rangle$ e $N_p = 0$, e cada um deles recebe a denominação especial de *isômero IPR*. A Tabela 3.2 (dados retirados de (DAUGHERTY, 2009)), semelhante à Tabela 1.1, compara o número de isômeros IPR com o número total de isômeros C_n para alguns valores de n . Infelizmente, como se pode ver na tabela, a partir de certo ponto até mesmo o número de isômeros IPR se torna grande demais para ser gerenciável. Obviamente, esse problema implica a necessidade de um novo esquema de avaliação, pois a regra dos pentágonos isolados não possui a capacidade de distinguir entre isômeros IPR.

Tabela 3.2: Número de isômeros de fulereno teoricamente possíveis para alguns valores de n .

Número de átomos, n	Número de isômeros de C_n	Número de isômeros IPR
20	1	0
30	3	0
40	40	0
50	271	0
60	1.812	1
70	8.149	1
80	31.924	7
90	99.918	46
100	713.319	450
110	713.319	2.355
120	1.674.171	10.774
130	3.580.637	39.393
140	7.341.204	121.354
150	14.059.173	335.569
160	26.142.839	836.497
170	46.088.148	1.902.265
180	79.538.725	4.071.832
190	131.561.725	8.187.581
200	214.127.713	15.655.672

3.2 Critério de Fowler-Manolopoulos

No mesmo artigo que descreve o índice de vizinhança para pentágonos, Raghavachari também propõe uma adaptação que permite lidar com fulerenos nos quais há múltiplos isômeros IPR. De acordo com (RAGHAVACHARI, 1992), deve-se considerar não apenas se duas faces pentagonais estão separadas, mas o *quão* separadas elas estão. Teoricamente, a tensão estérica (que interfere na estabilidade) é minimizada quando as curvaturas da regiões pentagonais se distribuem da maneira mais uniforme possível pela superfície da molécula (FOWLER e MANOLOPOULOS, 2006).

O primeiro passo para quantificar essa distribuição é redefinir o índice de vizinhança para lidar com hexágonos, ao invés de pentágonos (um passo necessário, visto que todos os isômeros IPR possuem a assinatura pentagonal ideal $\langle 12, 0, 0, 0, 0, 0 \rangle$ e $N_p = 0$). Para cada hexágono, o índice passa a ser o número de outros hexágonos aos quais ele é adjacente. A assinatura característica de cada isômero passa então a ser na forma $\langle h_0, h_1, h_2, h_3, h_4, h_5, h_6 \rangle$, onde h_k é o número de hexágonos com índice de vizinhança k . No caso particular dos isômeros IPR, a principal motivação dessa nova definição, cada hexágono é adjacente a, no mínimo, três outros. Caso contrário, dois dos pentágonos adjacentes ao hexágono em questão estariam se tocando – a Figura 3.1 ilustra os possíveis casos para um hexágono adjacente a apenas dois outros hexágonos (em todos não há como evitar pentágonos adjacentes). Portanto, a assinatura de um isômero IPR pode ser representada de modo abreviado por $\langle h_3, h_4, h_5, h_6 \rangle$, descartando h_0, h_1, h_2 (que valem sempre zero nesse tipo de isômero).

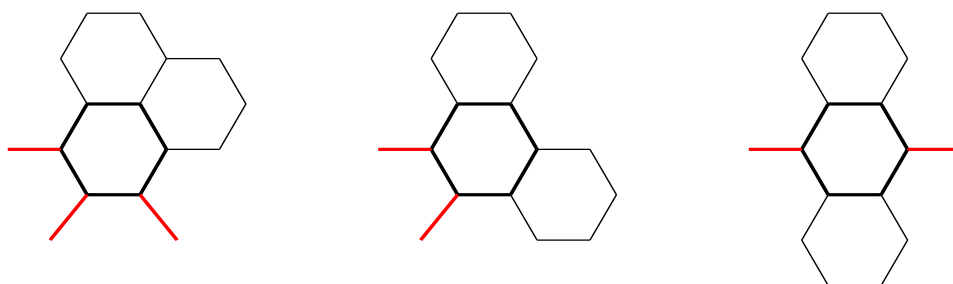


Figura 3.1: Hexágono com índice de vizinhança 2. Pentágonos adjacentes em vermelho.

No intuito de apresentar uma forma mais compacta da assinatura hexagonal, semelhante ao N_p para a assinatura pentagonal, Fowler e Manolopoulos definiram o desvio padrão, σ_h , da distribuição do índice de vizinhança hexagonal como um quantificador abreviado da tensão estérica (e, conseqüentemente, da estabilidade esperada) (FOWLER e MANOLOPOULOS, 2006):

$$\sigma_h = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$$

Onde (no caso geral)

$$\langle k \rangle = \frac{\sum_{k=0}^6 k h_k}{\sum_{k=0}^6 h_k} \quad \text{e} \quad \langle k^2 \rangle = \frac{\sum_{k=0}^6 k^2 h_k}{\sum_{k=0}^6 h_k}$$

Ou (no caso IPR)

$$\langle k \rangle = \frac{\sum_{k=3}^6 k h_k}{\sum_{k=3}^6 h_k} \quad \text{e} \quad \langle k^2 \rangle = \frac{\sum_{k=3}^6 k^2 h_k}{\sum_{k=3}^6 h_k}$$

Cabe aqui uma pequena observação. Como o C_{20} (apenas 1 isômero) é o único fulereno que não possui hexágonos, $\langle k \rangle$ e $\langle k^2 \rangle$ são divisões por zero e o valor da invariante para esse fulereno em particular é indefinido.

3.3 Diametro e Índice de Wiener

O *índice (ou número) de Wiener*, denotado por W , é um índice topológico apresentado em 1947 pelo químico norte-americano Harry Wiener (WIENER, 1947) que possui uma relação muito próxima com um grande número de propriedades físicas e químicas dos alcanos (também conhecidos como parafinas, são um tipo de hidrocarboneto). Embora Wiener tenha definido W apenas para os alcanos, em (HOSOYA, 1971) o químico japonês Haruo Hosoya foi o primeiro a propor uma relação entre W e as distâncias em um grafo molecular geral.

Hosoya mostrou que o valor de W em uma molécula corresponde a exatamente metade da soma de todos os elementos da matriz de distâncias de seu respectivo grafo molecular (a matriz M que contém uma linha e uma coluna para cada vértice no grafo, e na qual a posição $M_{i,j}$ corresponde à distância entre os vértices identificados por i e j). Essa divisão por 2 se justifica, visto que cada distância seria considerada duas vezes nessa soma: uma em $M_{i,j}$ e a outra em $M_{j,i}$. Dito de outra maneira, para um grafo $G = (V, E)$, o índice $W(G)$ é apenas a soma das distâncias entre todos os pares de vértices de G (ROUVRAY, 2002). Numa notação mais próxima da utilizada na Definição 2.1.14, $W(G)$ poderia ser expresso da seguinte forma:

$$W(G) = \sum_{u,v \in V} d_G(u,v)/2$$

A divisão pode ser removida se a distância for considerada apenas uma vez para cada par. Isto é, ao invés de somar $d_G(u,v)$ e $d_G(v,u)$, só uma delas seria levada em conta. Por convenção, supondo que os vértices no grafo sejam identificados através de números inteiros, a preferência é pela distância na qual o vértice com menor identificador aparece primeiro. Formalmente:

$$W(G) = \sum_{u,v \in V \mid u < v} d_G(u,v)$$

Para esclarecer melhor esse conceito, considere o exemplo a seguir. A Figura 3.2 mostra o grafo molecular correspondente ao 1,1,3-trimetil-ciclobutano (adaptado do grafo apresentado em (GUTMAN et al., 1993)). Neste caso:

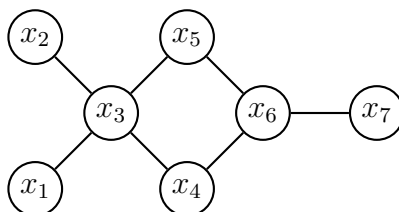


Figura 3.2: Grafo do 1,1,3-trimetil-ciclobutano, C_7H_{14} , adaptado de (GUTMAN et al., 1993).

$$\begin{aligned}
W(G) = & d_G(x_1, x_2) + d_G(x_1, x_3) + d_G(x_1, x_4) + d_G(x_1, x_5) + d_G(x_1, x_6) + d_G(x_1, x_7) \\
& + d_G(x_2, x_3) + d_G(x_2, x_4) + d_G(x_2, x_5) + d_G(x_2, x_6) + d_G(x_2, x_7) \\
& + d_G(x_3, x_4) + d_G(x_3, x_5) + d_G(x_3, x_6) + d_G(x_3, x_7) \\
& + d_G(x_4, x_5) + d_G(x_4, x_6) + d_G(x_4, x_7) \\
& + d_G(x_5, x_6) + d_G(x_5, x_7) \\
& + d_G(x_6, x_7)
\end{aligned}$$

$$\begin{aligned}
W(G) = & 2 + 1 + 2 + 2 + 3 + 4 \\
& + 1 + 2 + 2 + 3 + 4 \\
& + 1 + 1 + 2 + 3 \\
& + 2 + 1 + 2 \\
& + 1 + 2 \\
& + 1 \\
W(G) = & 42
\end{aligned}$$

Resta explicar onde essa definição se encaixa no estudo dos fulerenos. Na verdade, o índice de Wiener é uma forma de medir a ramificação do arcabouço formado pelos átomos de carbono na molécula e, conseqüentemente, de medir o quão compacta ela realmente é (BONCHEV e TRINAJSTIĆ, 1977; ROUVRAY, 1987). Isso significa que W representa a razão entre superfície e volume em uma molécula formada por átomos de carbono, e serve como indício da atuação de certas forças intermoleculares (GUTMAN et al., 1993) que, por sua vez, afetam a estabilidade. A relação esperada é que, conforme W aumente, a estabilidade diminua, mas já há estudos comentando (para alguns fulerenos) a pouca precisão dessa relação,

como (SLANINA et al., 2001).

Para finalizar esta seção, será apresentada uma invariante não muito conhecida na literatura de fulerenos, mas conceitualmente bastante próxima ao índice de Wiener. O *diâmetro* de um grafo G é a maior distância entre dois vértices de G considerando todos os possíveis pares de vértices. Ou seja,

$$\text{diam}(G) = \max\{d_G(u, v) \mid u, v \in V\}$$

Embora essa invariante não seja considerada particularmente relevante pela literatura na previsão da estabilidade de fulerenos, a utilização do diâmetro com esse propósito não é sem justificção. Em (ANDOVA et al., 2012), Andova e colegas definiram uma cota superior para o diâmetro de um grafo de fulereno, G , como sendo $\text{diam}(G) \leq \frac{1}{5}v(G) + 1$. No mesmo artigo é ressaltado que, mesmo que o diâmetro possa ser relativamente pequeno em grafos de fulereno apresentando simetria semelhante ao Buckminsterfulereno (ou seja, icosaédrica), o diâmetro dos grafos que modelam nanotubos de carbono é linear no número de vértices. Além disso, este trabalho apresenta uma boa oportunidade de comparar o diâmetro (que pode ser obtido “de graça” quando o índice de Wiener é calculado) com as demais invariantes.

3.4 Frustração Bipartida de Arestas

Uma das principais características dos grafos bipartidos é a ausência de ciclos de comprimento ímpar, que podem ser vistos como obstáculos à “bipartividade” de um grafo. Logo, uma maneira óbvia de quantificar essa bipartividade (ou seja, o quão próximo de bipartido o grafo se encontra) seria contar o número de ciclos de comprimento ímpar presentes no grafo (DOŠLIĆ, 2005; HOLME et al., 2003). Outra técnica semelhante, exposta em (ESTRADA e RODRÍGUEZ-VELÁZQUEZ, 2005), é a contagem de passeios fechados. Porém, uma medida muito mais intuitiva da bipartividade pode ser definida através da contagem do número de arestas que violam a característica mais marcante de um grafo bipartido: o fato de os dois extremos em cada aresta não fazerem parte da mesma classe da bipartição. Essa invariante, discutida a seguir, foi introduzida primeiramente no estudo de redes complexas (HOLME

et al., 2003) e, embora não seja computável de modo eficiente no caso geral, pode ser calculada em tempo polinomial no caso específico dos grafos de fulereno (DOŠLIĆ e VUKČEVIĆ, 2007).

Para prosseguir com a explicação é preciso primeiro definir de maneira mais clara o que significa uma aresta *violar* determinada bipartição (V_1, V_2) do conjunto de vértices de um grafo $G = (V, E)$. Seja $e \in E$ uma aresta do grafo G , e sejam $u, v \in V$ seus dois extremos. A aresta e é dita *frustrada* com respeito à (V_1, V_2) se, e somente se, $u, v \in V_1$ ou $u, v \in V_2$ (isto é, se seus dois extremos se encontram na mesma classe da bipartição). Para cada bipartição (X, Y) de V , denota-se por F_{XY} o conjunto formado por todas as arestas que violam (frustram) a bipartição (X, Y) . Finalmente, a *frustração bipartida de arestas* de um grafo G , denotada por $\varphi(G)$, é a cardinalidade do menor F_{XY} considerando todas as possíveis bipartições (X, Y) de V . De modo mais formal:

$$\varphi(G) = \text{mín} \{|F_{XY}| \mid (X, Y) \text{ é bipartição de } V\}$$

Equivalentemente, $\varphi(G)$ pode ser definido como o tamanho do menor conjunto de arestas que precisam ser removidas para se obter um subgrafo gerador bipartido. Decorrente dessa definição, é imediato que $\varphi(G) = 0$ se G for um grafo bipartido, e $\varphi(G) > 0$ caso contrário. Mais especificamente, para um grafo simples G (DOŠLIĆ e VUKČEVIĆ, 2007):

$$0 \leq \varphi(G) \leq |E|/2$$

No caso particular dos grafos de fulereno, esse limite é ligeiramente diferente. Todo grafo de fulereno possui exatamente 12 faces pentagonais, cada uma correspondendo a um ciclo de tamanho ímpar (embora, certamente, existam outros), o que torna fácil perceber que grafos de fulereno não são bipartidos. Portanto, $\varphi(G) > 0$ se G for um grafo de fulereno. Entretanto, é possível ir mais além considerando os pentágonos de G . Como no mínimo uma aresta deve ser removida em cada uma das 12 faces pentagonais, e como a remoção de uma única aresta é capaz de destruir simultaneamente apenas 2 faces pentagonais (se elas eram adjacentes),

obtêm-se:

$$12 \leq \varphi(G) \leq |E|/2, \text{ se } G \text{ corresponder a um isômero IPR}$$

$$6 \leq \varphi(G) \leq |E|/2, \text{ caso contrário}$$

Por fim, resta explicar como a frustração bipartida de arestas pode ser computada em um grafo de fulereno, G . Došlić e Vukčević mostram, em (DOŠLIĆ e VUKČEVIĆ, 2007), a relação existente entre $\varphi(G)$ e um conjunto particular de arestas no grafo dual do fulereno. Esse conjunto, denotado por H e denominado *obstáculo*, é o conjunto formado pelas arestas que precisam ser removidas de G' (o dual de G) para obter um subgrafo gerador que não possua vértices de grau ímpar. Para explicar de modo mais preciso, considere a distância entre cada par de vértices de grau 5 em G' (ou seja, a distância entre as faces pentagonais de G).

Seja $K_{12}(G)$ o grafo completo que possui um vértice para cada face pentagonal de G e no qual cada aresta está associada a um peso que corresponde à distância entre as faces representadas por seus extremos (esse grafo recebe a denominação especial de *grafo de distância pentagonal de G*). Novamente de acordo com Došlić e Vukčević, cada obstáculo de cardinalidade mínima vai estar mapeado diretamente em um emparelhamento perfeito de peso mínimo em $K_{12}(G)$. O peso deste emparelhamento (isto é, a soma dos pesos de suas arestas) corresponde ao valor da frustração bipartida de arestas em G .

Visto que o número de arestas em um grafo planar G está em $\mathcal{O}(v(G))$ e que o número de pentágonos em um grafo de fulereno é constante, é possível calcular as distâncias entre todos os pares de vértices de grau 5 em G' em tempo linear usando, por exemplo, um algoritmo de busca em largura (PAPADIMITRIOU e STEIGLITZ, 1982). O restante da computação (i.e., encontrar um emparelhamento perfeito em $K_{12}(G)$) pode ser feito em tempo constante – novamente devido ao fato de o número de faces pentagonais ser constante – e esse tempo é o mesmo independente do valor de $v(G)$, visto que todos os fulerenos possuem sempre o mesmo número de pentágonos. Logo, para qualquer fulereno, $\varphi(G)$ pode ser calculado em tempo linear no número de vértices (DOŠLIĆ e VUKČEVIĆ, 2007).

Note que a frustração bipartida de arestas pode ser calculada de modo diferente daquele proposto em (DOŠLIĆ e VUKČEVIĆ, 2007). Resumidamente: dado um grafo de fulereno, G , basta encontrar um subgrafo bipartido com o maior número possível de arestas para, em seguida, contar quantas das arestas de G não fazem parte desse subgrafo (pela definição dada anteriormente, esse número deve ser justamente a frustração bipartida de arestas). Entretanto, calcular $\varphi(G)$ dessa forma pode não ser muito eficiente, visto que o problema de encontrar um *subgrafo bipartido máximo* é equivalente ao problema de encontrar um *corte máximo* (NEWMAN, 2008), sabidamente NP-Difícil no caso geral (LIERS e PARDELLA, 2008). Certamente, existem soluções específicas para alguns casos especiais, como o algoritmo proposto em (LIERS e PARDELLA, 2008) para a classe dos grafos planares (que inclui os grafos de fulereno) e que apresenta complexidade $\mathcal{O}(v(G)^{3/2} \log v(G))$. Porém, pelo menos nesse artigo em particular, a complexidade do algoritmo apresentado é pior do que o método “direto” proposto por Došlić e Vukčević.

3.5 Número de Independência

A aplicação da invariante documentada nesta seção como um critério para a previsão da estabilidade de fulerenos foi inicialmente motivada por uma das conjecturas feitas de forma automática pelo programa *Graffiti* (FAJTLOWICZ, 1988, 1995). Embora partindo de um conjunto de dados extremamente limitado, a observação fortuita do programa, de que os isômeros estáveis possuíam número de independência menor do que os demais, parecia ser um critério promissor na determinação dos isômeros mais relevantes (pelo menos na pequena faixa de valores de n considerada na época).

Para definir formalmente o número de independência, é preciso primeiro revisar a noção de *conjunto independente*. Um conjunto independente, $S \subseteq V$, de um grafo $G = (V, E)$ é um subconjunto (maximal) dos vértices de G tal que o subgrafo induzido por S , $G[S]$, não possui arestas. Dito de maneira equivalente, é um conjunto de vértices tal que nenhum par entre eles é adjacente no grafo. Um conjunto independente que possua cardinalidade máxima

é denominado *conjunto independente máximo*, e seu tamanho corresponde ao *número de independência* do grafo G , denotado por $\alpha(G)$. De modo mais formal (DAUGHERTY, 2009):

$$\alpha(G) = \max\{|S| \mid S \text{ é um conjunto independente de } G\}$$

Uma das primeiras aparições do número de independência num contexto relacionado à química aconteceu através dos estudos de Taylor sobre os átomos de bromo agregados ao Buckminsterfulereno no $C_{60}Br_{24}$ (TAYLOR, 1995). As posições onde esses átomos se encaixam na estrutura de carbono definem um conjunto independente máximo no C_{60} . No caso mais geral, supondo que se deseja posicionar adendos de algum tipo X em um fullereno C_n para obter um composto na forma C_nX_q , e supondo também que esses adendos sejam grandes demais para se ligarem a átomos de carbono vizinhos no fullereno, as posições finais dos adendos vão corresponder (ignorando outras restrições químicas sobre a adição) a um conjunto independente máximo no grafo do fullereno, e q corresponderá ao número de independência (DAUGHERTY, MYRVOLD e FOWLER).

No caso dos grafos cúbicos e planares (uma classe que contém os fullerenos), já se provou que o problema de encontrar o número de independência é NP-Difícil (GAREY e JOHSON, 1990). Heckman e Thomas definiram uma cota inferior de $\alpha(G) \geq \frac{3}{8}v(G)$ quando G é um grafo cúbico, planar, e sem triângulos (HECKMAN e THOMAS, 2006), novamente uma classe da qual os fullerenos fazem parte. Já uma cota superior para $\alpha(G)$ foi definida em (FOWLER et al., 1999) como $\frac{v(G)}{2} - 2$. Há também uma relação muito significativa entre o número de independência e o diâmetro de um grafo de fullereno, G , expressa em (ANDOVA et al., 2012) como $\alpha(G) \geq 2(\text{diam}(G) - 1)$.

Embora a observação estatística feita pelo programa *Graffiti* tenha motivado a proposição do número de independência como um critério para prever a estabilidade de fullerenos em (FAJ-TOLOWICZ e LARSON, 2003), não havia uma argumentação química coerente que justificasse essa proposição. Na verdade, no caso específico de nanotubos de carbono, o número de independência é uma invariante especialmente imprecisa, visto que esse tipo de fullereno possui

um número de independência muito maior do que outros isômeros com o mesmo número de átomos (o que não significa necessariamente que eles sejam menos estáveis) (FAJTOLOWICZ e LARSON, 2003). De acordo com estudos mais atuais (FOWLER, DAUGHERTY e MYRVOLD; FOWLER e MYRVOLD, 2005), parece não haver muita relação entre a estabilidade de fulerenos e o número de independência, mas o estudo do problema não deixa de ser útil, por exemplo, por ter motivado a invariante descrita na Seção 3.6 a seguir, que é uma versão modificada do número de independência.

3.6 Número de Independência Closed-Shell

O número de independência por si só, embora seja indicado na literatura como uma invariante na previsão da estabilidade de fulerenos e tenha sido bem investigado, ignora certas restrições da teoria química quanto aos átomos que estão *fora* do conjunto independente máximo. Em linhas gerais, embora possam formar um subgrafo desconexo, esses átomos devem possuir o que se conhece como uma distribuição estável de elétrons π (DAUGHERTY, MYRVOLD e FOWLER). Esse é um tópico que merece uma pequena digressão para explicar como os conceitos da química se relacionam com os conceitos em teoria dos grafos.

Moléculas de fulereno são compostas, como já foi mencionado várias vezes, inteiramente por átomos de carbono. Cada átomo de carbono possui 4 elétrons na camada de valência, sendo que três desses elétrons formam ligações com os três vizinhos do átomo no fulereno (relembrando que fulerenos são 3-regulares). O quarto elétron que “sobra” em cada átomo é usado para compor o que se denomina *sistema π insaturado* (ou *sistema conjugado*) na superfície da molécula (DAUGHERTY, 2009).

No caso dos derivados de um fulereno C_n formados pelo acréscimo de átomos de algum tipo X , quando um adendo X se liga a um átomo de carbono, o quarto elétron daquele átomo deixa de fazer parte do sistema π do fulereno e passa a formar uma ligação com X . Ligando um número suficiente de adendos ao fulereno, o sistema π é quebrado em várias partes disjuntas que correspondem aos subgrafos induzidos pelos átomos de carbono que não estão

ligados a nenhum adendo. Ou seja, aos vértices que não fazem parte do conjunto independente definido pelas posições onde os adendos X estão ligados. De acordo com a teoria de Hückel, a molécula de fulereno com adendos será estável se cada um desses subgrafos for estável.

Como, então, testar essa estabilidade? Voltando agora a falar do número de independência, uma das maneiras encontradas para resolver esse problema é adaptar a invariante para considerar as restrições impostas pelos modelos químicos de distribuição eletrônica. Felizmente, para o caso dos fulerenos (formados inteiramente por átomos de carbono), esses modelos podem ser simplificados e as restrições sobre os mesmos podem ser traduzidas diretamente em restrições sobre os autovalores na matriz de adjacência do grafo correspondente. Os autovalores são uma interpretação matemática dos níveis de energia orbital no modelo simplificado de estrutura eletrônica da teoria de Hückel (STREITWIESER JR., 1961).

Um grafo $G = (V, E)$ é dito ser *closed-shell* (numa tradução livre do jargão químico: “de camada fechada”) se possui um número par de vértices (condição atendida por todos os grafos de fulereno) e se exatamente metade dos autovalores da matriz de adjacência de G são positivos. Já um subconjunto, S , dos vértices de um grafo é dito ser um *conjunto independente closed-shell* se:

- S é um conjunto independente; e
- cada componente conexa de $G - S$ é *closed-shell*.

Como no número de independência, o tamanho do maior conjunto independente *closed-shell* em um grafo de fulereno G corresponde ao *número de independência closed-shell*, denotado por $\alpha^-(G)$. Definido mais formalmente:

$$\alpha^-(G) = \max\{|S| \mid S \text{ é um conjunto independente } closed-shell \text{ de } G\}$$

Como $\alpha^-(G)$ é uma versão mais restrita de $\alpha(G)$, é imediato que:

$$\alpha^-(G) \leq \alpha(G)$$

Um limitante menos óbvio é definido em (DAUGHERTY, MYRVOLD e FOWLER) como sendo (onde G é um grafo de fulereno e $2\lfloor v(G)/5 \rfloor$ denota o maior inteiro par menor ou igual a $2v(G)/5$):

$$\alpha^-(G) \leq 2\lfloor v(G)/5 \rfloor$$

Uma versão aperfeiçoada desse limitante é fornecida em (DAUGHERTY, 2009), sendo que a nova versão é tão boa quanto a anterior para $v(G) \geq 46$ e estritamente melhor do que a anterior quando $v(G) \geq 126$:

$$\alpha^-(G) \leq 2\lfloor 3v(G)/16 + 12/16 \rfloor$$

Dois problemas importantes – e, até onde se sabe, ainda em aberto – sobre o número de independência *closed-shell* são: i) o de determinar a complexidade de calculá-lo; difícil devido à necessidade de computar os autovalores da matriz de adjacência, e ii) o da existência de um conjunto independente *closed-shell* para todos os fulerenos. Quanto ao segundo problema, embora conjuntos independentes com essa propriedade tenham sido encontrados para um grande número de fulerenos (DAUGHERTY, 2009), ainda não há prova definitiva de que esse seja o caso para *qualquer* fulereno, sendo que um possível passo no caminho da resolução seria a investigação de fulerenos com mais autovalores negativos do que positivos. Infelizmente, o menor fulereno conhecido com essa propriedade possui 628 vértices (FOWLER, 1997), o que torna o cálculo desses valores algo bastante complexo.

Para encerrar esta seção seguindo o exemplo da Seção 3.3, que apresentou o diâmetro como uma invariante relacionada ao índice de Wiener, aqui será apresentada uma invariante relacionada ao número de independência *closed-shell*, e que pode ser obtida facilmente quando este é calculado. Dado um grafo de fulereno G , denota-se por $\lambda_n(G)$ o *menor autovalor da matriz de adjacência* de G . Esse valor, embora pareça ser apenas marginalmente relacionado com certas energias orbitais na molécula, foi, entretanto, investigado como uma invariante na previsão da estabilidade de fulerenos (ANDOVA et al., 2012; FOWLER, 2003) e se relaciona com o maior autovalor Laplaciano do grafo de acordo com a relação (GODSIL e ROYLE,

2001):

$$\lambda_n(G) = 3 - \mu_\infty(G)$$

Lembrando que o *Laplaciano* de um grafo com matriz de adjacência $A_{n \times n}$ é a matriz $C_{n \times n}$ onde $C_{i,j} = d(i)$ se $i = j$, e $C_{i,j} = -A_{i,j}$, caso contrário. Um *autovalor Laplaciano* de um grafo G é um autovalor de seu Laplaciano, sendo o maior autovalor Laplaciano denotado por $\mu_\infty(G)$ (FARIA, KLEIN e STEHLÍK).

Observações empíricas feitas por Fowler num conjunto limitado de fulerenos (os fulerenos C_n delimitados por $n = 20 + 2k$, $k = 0, \dots, 40$, $k \neq 1$) mostram uma tendência de crescimento de $\lambda_n(C_n)$ conforme n aumenta. Porém, os principais resultados sobre esse valor são uma conjectura, feita em (FOWLER, 2003), sobre o Buckminsterfullereno ser o fullereno com 60 vértices ou mais que maximiza $\lambda_n(G)$, e o seguinte limite superior (ANDOVA et al., 2012):

$$\lambda_n(G) \leq -3 + \frac{157,16}{\sqrt{v(G)}}$$

Posteriormente, em (FARIA, KLEIN e STEHLÍK), esse limite foi aperfeiçoado para

$$\lambda_n(G) \leq -3 + 8\sqrt{\frac{3}{5v(G)}}$$

3.7 Número de Estruturas de Kekulé

Saindo do assunto sobre derivados de fulerenos e voltando a falar exclusivamente sobre fulerenos C_n , esta seção (juntamente com as duas próximas seções) apresenta outra maneira de medir a estabilidade do sistema π insaturado em uma molécula de carbono. Modelando a ligação dos elétrons de valência do carbono como arestas no grafo de fullereno (3-regular), a ligação do quarto elétron em cada átomo (a que forma o sistema π) pode ser vista como definindo um *emparelhamento perfeito* no grafo.

Em química, o conceito equivalente a um emparelhamento perfeito num arcabouço de carbono é denominado *estrutura de Kekulé* (em homenagem ao químico alemão Friedrich

August Kekulé), e o número de estruturas desse tipo suportadas pela molécula (denotado aqui por $K(G)$ para um fullereno G) fornece uma medida de sua estabilidade eletrônica. Esse número, no caso específico dos grafos planares, pode ser calculado em tempo polinomial por um algoritmo conhecido como *FKT* – uma sigla formada pelas iniciais dos pesquisadores Kasteleyn (KASTELEYN, 1961), Temperley e Fisher (TEMPERLEY e FISHER, 1961), que descobriram independentemente um resultado (posteriormente generalizado por Kasteleyn para todos os grafos planares (KASTELEYN, 1963, 1967)) que levaria ao algoritmo.

Resumidamente, a ideia principal por trás do algoritmo *FKT* é a relação que existe entre o número de emparelhamentos perfeitos de um grafo planar e o determinante da matriz de adjacência de um grafo orientado obtido a partir do grafo planar em questão. Antes de explicar os detalhes desse algoritmo, é necessário primeiro definir dois conceitos importantes relacionados à matriz de adjacência de um grafo planar.

Em termos simples, um *Pfaffiano* é a raiz quadrada do determinante de uma matriz anti-simétrica (BRESSOUD e PROPP, 1999). Embora seja um conceito da álgebra linear, pode também ser aplicado no estudo dos grafos, pois estes podem ser representados por matrizes de adjacência. Note, porém, que eles se aplicam apenas a grafos orientados (isto é, grafos nos quais se atribui uma *direção* ou *sentido* a cada aresta), caso contrário as matrizes de adjacência correspondentes não seriam anti-simétricas e a definição não se enquadraria.

O segundo conceito a ser definido é o de *orientação Pfaffiana*. Dito de maneira informal, uma *orientação Pfaffiana* de um grafo planar G (imerso no plano) é uma atribuição de direção a cada aresta de G de forma que cada face de G tenha um número ímpar de arestas no sentido horário (KASTELEYN, 1963, 1967; VAZIRANI, 1989). O Pfaffiano da matriz de adjacência de um grafo planar orientado desta forma (preenchendo a matriz apenas com os valores 1, -1 , e 0) deve corresponder precisamente ao número de emparelhamentos perfeitos no grafo. Além disso, ainda de acordo com Kasteleyn, todo grafo planar possui uma orientação Pfaffiana e essa orientação pode ser encontrada em tempo polinomial.

Uma vez que esses dois conceitos foram definidos é possível dar continuação à explicação do algoritmo *FKT*, que consiste em duas partes principais: i) dado um grafo planar G , encontrar

uma orientação Pfaffiana de G (preenchendo a matriz de adjacência apenas com os valores 1, -1 , e 0), e ii) calcular o Pfaffiano da matriz de adjacência orientada de G . O pseudocódigo, adaptado de (VAZIRANI, 1989), pode ser visto no Algoritmo 1. A Figura 3.3 ilustra duas iterações do laço “para cada folha v em H ”.

Algoritmo 1 FKT(G)

Entrada: Um grafo planar, $G = (V, E)$, imerso no plano

Saída: Número de emparelhamentos perfeitos, $K(G)$, de G

Seja $T = (V, F)$ uma árvore geradora de G

para cada $e \in F$ **faça**

 Dê uma orientação arbitrária a e

fim para

Seja H uma árvore com um vértice para cada face em G (incluindo a externa) e uma aresta entre dois vértices se suas faces correspondentes em G dividem uma aresta que não está em T . A escolha da raiz é arbitrária, mas convencionou-se escolher o vértice correspondente à face externa.

para cada folha v em H com exceção da raiz **faça**

 Seja f a face de G correspondente a v

 Seja e a (única) aresta de G em f que ainda não possui orientação

 Oriente e de forma que o número de arestas no sentido horário em f seja ímpar

 Remova v de H

fim para

Seja M a matriz de adjacência de G preenchida apenas com valores 1, -1 , e 0

retorne $|\sqrt{\det(M)}|$

Para encerrar esta seção, é importante mencionar que, conforme constatado em (AUSTIN et al., 1994), a medida de estabilidade eletrônica fornecida pelo número de estruturas de Kekulé parece ser um pouco grosseira. Neste artigo, os autores fizeram uma análise da invariante nos isômeros do C_{60} , através de uma técnica baseada na teoria de Kasteleyn-Little para emparelhamentos perfeitos em grafos planares (KASTELEYN, 1963; LOVÁSZ e PLUMMER, 1986), e descobriram que, mesmo nesse caso limitado, a contagem de estruturas de Kekulé falha em isolar o Buckminsterfullereno como o isômero mais estável (SCHMALZ et al., 1986).

Ao menos nessa análise limitada, a invariante parece favorecer estruturas não esféricas com alta tensão estérica (provavelmente devido ao fato de um alto valor K estar associado com um maior número de adjacências pentagonais), candidatos improváveis para moléculas estáveis. Essa falta de correlação de K com a estabilidade eletrônica sugere que talvez alguns

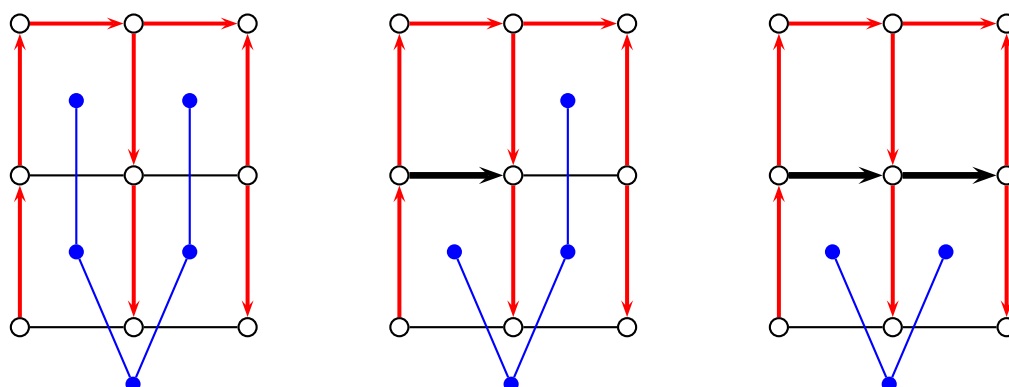


Figura 3.3: Dois passos do algoritmo FKT. Árvore geradora T em vermelho, e H em azul.

emparelhamentos em particular sejam mais importantes do que outros, ou do que o simples número total.

3.8 Número de Fries

Um dos primeiros trabalhos a levar essa noção em conta foi o de Fries, que propôs o número de *hexágonos benzenóides* como um critério de estabilidade (FRIES, 1927, WALTER e SCHILLING). Um *hexágono benzenóide* (com respeito a uma estrutura de Kekulé ou emparelhamento M) nada mais é do que uma face hexagonal tal que as ligações representadas por suas arestas sejam alternadamente simples (σ) e duplas (π), ou, equivalentemente, tal que suas arestas estejam alternadamente dentro e fora de M . A Figura 3.4 mostra um exemplo, destacando em vermelho as arestas que estão no emparelhamento.

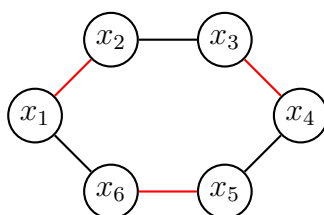


Figura 3.4: Exemplo de hexágono benzenóide (arestas do emparelhamento em destaque).

Posteriormente a teoria de Fries foi generalizada e sistematizada em duas variantes: uma proposta por Randić e colegas (RANDIĆ, 1961, 1976), e a outra por Herndon e colegas (HERNDON, 1973; HERNDON e ELLZEY, 1974). Experimentos limitados a alguns isômeros do C_{60} já mostram que as duas versões representam um avanço em relação ao número de estruturas de Kekulé puro (SCHMALZ et al., 1986). Neste trabalho serão utilizadas a terminologia e notação apresentadas em (AUSTIN et al., 1994). O *número de Fries* para um emparelhamento perfeito M , denotado por $n_F(M)$, é o número de hexágonos benzenóides em M . Já o *número de Fries* de um fulereno $G = (V, E)$, denotado por $F(G)$, é o valor máximo de $n_F(M)$ considerando todos os emparelhamentos perfeitos possíveis em G .

O problema de encontrar $F(G)$ pode ser modelado como um problema de programação inteira (HANSEN e ZHENG, 1994), conforme descrito a seguir. Em primeiro lugar, deve-se associar uma variável binária X_{vw} a cada aresta $e = \{v, w\} \in E$, para indicar se a aresta faz ou não parte do emparelhamento. Em seguida, para cada vértice $v \in V$, criam-se restrições na forma:

$$\sum_{w \in N_G(v)} X_{vw} = 1$$

Pela definição de emparelhamento, no máximo uma das arestas incidente em cada vértice pode fazer parte do emparelhamento, e, pela definição de emparelhamento perfeito, todos os vértices devem estar cobertos. Note que o grafo em questão admite emparelhamento perfeito se, e somente se, esse conjunto preliminar de equações possui uma solução. Além disso, no caso dos fulerenos (que são 3-regulares), cada restrição vai conter apenas três termos.

O próximo passo é criar, para cada face hexagonal s de G (ou para cada vértice s de grau 6 em seu dual), duas variáveis contínuas, Y_s e Z_s (limitadas ao intervalo $[0,1]$), e seis restrições sobre as variáveis binárias correspondentes às arestas da face s (supondo que essas arestas possam ser obtidas em sequência, e denotando-as a seguir por $X_{si}, 1 \leq i \leq 6$):

$$\begin{aligned}
Y_s &\leq X_{s1} & Z_s &\leq X_{s2} \\
Y_s &\leq X_{s3} & Z_s &\leq X_{s4} \\
Y_s &\leq X_{s5} & Z_s &\leq X_{s6}
\end{aligned}$$

Perceba que a notação X_{si} é utilizada apenas por facilidade de explicação, e a variável em si deve corresponder a alguma das variáveis X_{vw} definidas anteriormente (ou seja, não são criadas novas variáveis X). Obviamente, isso implica que restrições correspondentes a hexágonos adjacentes vão compartilhar variáveis, uma situação ilustrada na Figura 3.5.

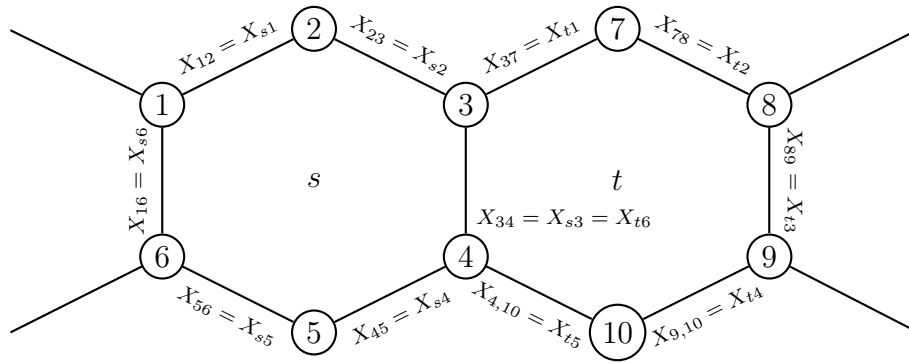


Figura 3.5: Hexágonos adjacentes s e t compartilham a variável X_{34} .

As variáveis Y_s e Z_s representam as duas únicas possíveis “configurações benzenóides” do hexágono s : ou a primeira, terceira, e quinta arestas fazem parte do emparelhamento (caso em que Y_s e Z_s assumirão, respectivamente, os valores 1 e 0), ou isso deve acontecer com a segunda, quarta, e sexta arestas (caso em que Z_s assumirá valor 1 e Y_s valor 0). Em todos os outros casos, as duas variáveis serão forçadas a assumir o valor 0, pois o hexágono não será benzenóide. Por fim, a função objetivo do problema pode ser definida como uma soma dessas variáveis em cada hexágono s :

$$F(G) = \text{maximize } z = \sum_s (Y_s + Z_s)$$

Perceba também que não importa qual aresta é rotulada como “primeira” do hexágono, pois as restrições criadas vão ser equivalentes – mesmo invertendo o significado de Y_s e Z_s para aquele hexágono, o somatório da função objetivo vai permanecer o mesmo – e não vão interferir nas restrições de outros hexágonos. Além disso, o fato de haver ou não arestas do emparelhamento ao longo de faces pentagonais não importa para o cálculo do número de Fries, e a formulação acima não impede a consideração de emparelhamentos que contenham algumas arestas em faces pentagonais (mesmo arestas entre duas faces pentagonais).

É interessante notar que o Buckminsterfulereno é o único entre todos os 1812 isômeros do C_{60} a possuir todos os seus hexágonos simultaneamente benzenóides, um bom indício da eficácia da invariante. Note também que, como um fulereno C_n possui $3v(C_n)/2$ arestas e $v(C_n)/2 - 10$ faces hexagonais, um emparelhamento perfeito terá cardinalidade $v(C_n)/2$ e o número máximo de hexágonos benzenóides é $v(C_n)/3$, um limitante superior para o valor de $F(C_n)$.

3.9 Número de Taylor

Embora o número de Fries seja um avanço em relação à contagem de estruturas de Kekulé pura, seguir à risca o critério de maximizar o número de hexágonos benzenóides pode não ser uma técnica muito “refinada”, especialmente nos isômeros IPR (que já se sabe serem os mais propícios à estabilidade). Em particular, os emparelhamentos perfeitos que possuem maior n_F acabam muitas vezes posicionando suas arestas ao longo das faces pentagonais do fulereno IPR, algo que deveria ser evitado, de acordo com Taylor (TAYLOR, 1992), com maior prioridade do que maximizar o número de hexágonos benzenóides.

Duas implementações do *número de Taylor* são propostas em (AUSTIN et al., 1994). As duas versões, denotadas por $T_a(G)$ e por $T_b(G)$, são definidas (de modo semelhante ao número de Fries $F(G)$) para um grafo de fulereno G como sendo o número máximo de hexágonos

benzenóides $n_F(M)$ entre os emparelhamentos perfeitos M de algum conjunto. Porém, ao invés de levar em conta o conjunto de *todos* os emparelhamentos perfeitos de G , é feita primeiro uma filtragem com base em dois critérios ligeiramente diferentes. Para um emparelhamento perfeito M em um fulereno G , seja $n_T(M)$ o número de arestas de M que fazem parte de faces pentagonais em G , e seja $n_{P_2}(M)$ o número de faces pentagonais contendo exatamente 2 arestas em M .

Então, $T_a(G)$ primeiro seleciona aqueles emparelhamentos perfeitos que *minimizam* o valor de $n_T(M)$, para em seguida selecionar algum que *maximize* $n_F(M)$. Já $T_b(G)$ seleciona os emparelhamentos perfeitos que minimizam $n_T(M)$ e $n_{P_2}(M)$, para depois selecionar algum que maximize $n_F(M)$. No geral, devido a essas restrições extras, vai ser válida a relação $F(G) \geq T_a(G) \geq T_b(G)$

Diferentemente do que ocorreu com o número de Fries, este autor não encontrou uma explicação detalhada de como modelar o cálculo do número de Taylor usando programação inteira, apenas uma afirmação em (AUSTIN et al., 1994) de que isso havia sido feito. Portanto, foi escolhida a versão $T_a(G)$ do número de Taylor e foi feita uma adaptação do método descrito em (HANSEN e ZHENG, 1994), com um cálculo em duas partes. Para facilitar a explicação, a partir de agora o valor mínimo de $n_T(M)$ entre todos os emparelhamentos M de um grafo G será denotado por $n_\Theta(G)$.

Da mesma forma que no cálculo de $F(G)$, a primeira parte envolve a criação das variáveis X_{vw} para cada aresta $\{v, w\} \in E$ e das restrições $\sum_{w \in N_G(v)} X_{vw} = 1$ para cada vértice $v \in V$. Porém, encontrar o valor de $n_\Theta(G)$ exige também a criação de duas novas variáveis e uma nova restrição para cada aresta: a variável P_{vw} , que indica se a aresta $\{v, w\}$ faz parte de um pentágono, a variável A_{vw} , que indica se a aresta faz simultaneamente parte do emparelhamento e de uma face pentagonal, e a restrição $A_{vw} = X_{vw} \wedge P_{vw}$. A função objetivo da primeira parte pode ser definida, simplesmente, como uma minimização da soma das variáveis A_{vw} :

$$n_{\Theta}(G) = \text{minimize } z = \sum_{\{v,w\} \in E} (A_{vw})$$

Uma vez que esse valor seja encontrado e armazenado, pode-se começar a tratar da segunda parte do problema: calcular um emparelhamento que maximize $n_F(M)$ mas que também atinja o valor mínimo $n_{\Theta}(G)$. Para resolvê-la, basta criar novamente as mesmas variáveis e restrições criadas no processo de calcular o número de Fries, acrescentar as variáveis e restrições mencionadas no parágrafo anterior (A_{vw} e P_{vw}), e acrescentar uma nova restrição na forma:

$$\sum_{\{v,w\} \in E} A_{vw} = n_{\Theta}(G)$$

A função objetivo também é a mesma usada no número de Fries, mas essa restrição extra garante que o valor calculado corresponda a um emparelhamento que minimize o número de arestas ao longo de faces pentagonais, como na definição de $T_a(G)$.

Como um pequeno adendo, resta apenas discutir mais uma invariante relacionada a emparelhamentos, mas não aos emparelhamentos perfeitos vistos até agora. O *número de saturação* para um grafo de fulereno G , denotado por $s(G)$, é a cardinalidade do *menor* (ou de algum dos menores, se houver mais de um com o menor tamanho) *emparelhamento maximal* de G (ANDOVA et al., 2012). Um limitante superior para o número de saturação foi definido em como sendo:

$$s(G) \leq \frac{v(G)}{2} - \frac{1}{4}(\text{diam}(G) - 2)$$

que pode ser simplificado para:

$$s(G) \leq \frac{v(G)}{2} - \frac{\sqrt{24v(G) - 15} - 15}{24}$$

O número de saturação não recebeu tanta atenção, até o momento, quanto outras invariantes utilizadas na previsão da estabilidade de fulerenos. Ele já foi, entretanto, estudado nesse contexto (DOŠLIĆ, 2002, 2008) e parece ser também digno de uma investigação empírica mais detalhada.

3.10 Constante de Cheeger

A invariante apresentada nesta seção busca encontrar grafos que sejam altamente conexos, mas que também sejam esparsos. Embora as duas propriedades sejam aparentemente contraditórias, grafos desse tipo são bastante úteis em diversas áreas (LUBOTZKY e PAK, 2000; NAOR e NAOR, 1993; PIPPENGER, 1987; SIPSER, 1988; SIPSER e SPIELMAN, 1996). Por exemplo, saindo momentaneamente do contexto de fulerenos, considere duas redes de transmissão de informação: a primeira correspondendo a um grafo completo K_n e a segunda a um ciclo C com o mesmo número de vértices.

Embora a rede representada por K_n possua altíssima conectividade, isso ocorre apenas devido às arestas em excesso, que podem acabar sendo muito custosas, potencialmente tornando inviável a construção da rede na prática. Por outro lado, a rede representada por C não sofre desse problema, já que o número de arestas é bem baixo. Porém, C ser muito esperso significa também que a rede é facilmente destruída (desconectada ou rompida), sendo necessário remover apenas 2 arestas para tanto. Uma maneira de encontrar um equilíbrio entre esses dois extremos é a invariante conhecida como *constante de Cheeger* (ou *constante isoperimétrica*, ou ainda *constante expansora*) (ALON, 1986; DAVIDOFF, SARNAK e VALETTE; LUBOTZKY, PHILLIPS e SARNAK; SARNAK, 2004), que fornece uma medida da conectividade de um grafo mais complexa do que o índice de Wiener e é particularmente aplicável no caso de grafos k -regulares.

O cálculo dessa constante em um grafo $G = (V, E)$ exige que se saiba, para cada subconjunto de vértices $S \subset V$ possível, a razão entre dois valores: 1) o número de arestas que vão do conjunto para seu complemento, e 2) o tamanho do menor entre os dois conjuntos (S e seu complemento). Dessa forma, em contraste com a definição normalmente utilizada de conectividade, o número de arestas que passam entre os dois conjuntos é considerado em relação ao tamanho desses conjuntos (JUSTUS, 2007). O problema de determinar essa constante em um grafo é também conhecido como o *problema do corte mais esperso* e já se sabe que ele é NP-Difícil (MOHAR, 1989).

De maneira mais formal, a constante de Cheeger em um grafo G , denotada por $ch(G)$ e derivada de um conceito similar proposto por Jeff Cheeger sobre variedades de Riemann (CHEEGER, 1970), pode ser definida como a seguir (JUSTUS, 2007):

$$ch(G) = \min_{\emptyset \subset U \subset V} \frac{|\delta U|}{\min(|U|, v(G) - |U|)}$$

onde

$$\delta U = \{\{u, v\} \mid u \in U, v \in V - U\}$$

é o conjunto de todas as arestas que conectam U a $V - U$. Quando $ch(G)$ é elevado, muitos subconjuntos grandes possuem muitas arestas ligando-os a seus complementos (o que torna difícil remover essas partes do grafo). Já se o valor de $ch(G)$ for pequeno, existe um “gargalo” no qual um grande subconjunto de vértices está conectado ao resto do grafo por relativamente poucas arestas.

No contexto específico da estabilidade de fulerenos, ainda não se sabe se a constante de Cheeger possui alguma capacidade de previsão significativa, mas ela pode servir como uma pista extra na identificação dos isômeros mais estáveis (FOWLER et al., 2000). Além disso, como grafos de fulereno são 3-regulares, o número de arestas de um fulereno cresce linearmente com o número de vértices. Logo, nesse caso a propriedade de ser esparsa pode ser considerada como “garantida” e o objetivo é encontrar os grafos com maior valor de $ch()$.

Por fim, existe também uma relação entre a constante de Cheeger e os autovalores da matriz de adjacência de um grafo (ALON e MILMAN, 1985; TANS et al., 1997). Seja G um grafo k -regular e A sua matriz de adjacência. Sabe-se que um dos autovalores de A é o valor k , e todos os outros estão no intervalo $[-k, k]$. Sendo $\lambda_1(A)$ o segundo maior autovalor de A , a desigualdade abaixo é válida:

$$\frac{k - \lambda_1(A)}{2} \leq ch(G) \leq \sqrt{2k(k - \lambda_1(A))}$$

4 PROJETO E IMPLEMENTAÇÃO

Neste capítulo há uma breve descrição dos aspectos práticos e dos detalhes de implementação do projeto. A Seção 4.1 define o escopo do trabalho, destacando seu objetivo principal e as etapas envolvidas na sua execução. Já a Seção 4.2 apresenta a representação escolhida para os grafos de fulereno neste trabalho, conhecida como *código planar* e voltada especificamente para o armazenamento eficiente de grafos planares. Por fim, a Seção 4.3 é dividida em várias subseções, uma para cada invariante incluída na comparação, mostrando decisões importantes de implementação e alguns trechos relevantes de código.

4.1 Escopo

O principal objetivo do trabalho ora proposto, como indicado pelo título, é fazer uma comparação de algumas invariantes da teoria dos grafos. Essa comparação é feita no contexto específico da previsão da estabilidade de moléculas de fulereno, moléculas que podem ter sua estrutura modelada através dos grafos de fulereno, explicados no Capítulo 2. Logo, para cumprir esse objetivo, é necessário: i) obter um conjunto de grafos de fulereno nos quais a análise possa ser feita, ii) definir quais invariantes serão incluídas na comparação, e iii) coletar dados sobre o valor de cada invariante em cada um dos grafos do conjunto de dados. Nos parágrafos a seguir serão discutidos os itens i) e ii) – o item iii) será analisado com detalhes no Capítulo 5, juntamente com os resultados da comparação.

Certamente, existem diversas maneiras possíveis de gerar um conjunto de grafos de fulerenos, mas as três principais das quais o proponente tem notícia são os programas *CaGe* (BRINKMAN et al., 2010)¹, *buckygen* (BRINKMAN, GOEDGEBEUR e MCKAY), e *fullgen* (BRINKMAN et al., 2005; BRINKMAN e MCKAY, 2005, 2007). Durante a elaboração da proposta desta dissertação, o autor utilizou o terceiro desses programas para gerar um conjunto preliminar de grafos de fulerenos (com até 150 vértices), um processo que levou cerca de dois dias para ser concluído em um Macbook Pro com processador Intel Core 2 Duo 2,4GHz, 4GB de

¹Acesse o elo <http://www.math.uni-bielefeld.de/~CaGe/>

memória RAM e sistema operacional Mac OS X Mountain Lion. Vale mencionar que, devido ao número elevado de isômeros, o conjunto total ocupava cerca de 60GB em disco, mesmo com uma representação (explicada mais abaixo) especificamente voltada para o armazenamento eficiente de grafos planares.

Entretanto, foram encontradas duas complicações que levaram esse conjunto inicial a ser descartado e substituído. A primeira foi o fato de os isômeros em cada arquivo não estarem ordenados lexicograficamente pela sequência espiral canônica (algo imprescindível para fazer uma análise correta, com o mesmo referencial que vários dos artigos citados por este trabalho). A segunda foi a lentidão inesperada dos programas para calcular algumas das invariantes (como o número de independência, o número de Fries, e o número de Taylor), que acabaria por inviabilizar este projeto se a decisão de calcular até $n = 150$ (o objetivo inicial durante a proposta) se mantivesse.

Felizmente, este autor teve sucesso em contatar por e-mail os pesquisadores Gunnar Brinkmann, Brendan McKay, e Jan Goedgebeur. Eles não só confirmaram que os programas *buckygen* e *fullgen* definem uma ordem própria para os isômeros, sem relação com as sequências espirais, como também tiveram a boa vontade de sugerir soluções alternativas. Uma dessas sugestões foi a base de dados *House of Graphs* (BRINKMAN et al., 2013), que acabou sendo escolhida por já apresentar todos os isômeros ordenados lexicograficamente por sequência espiral canônica (resolvendo de imediato um dos problemas).

A segunda dificuldade foi amenizada (embora não inteiramente resolvida, como será visto no Capítulo 5) com uma pequena modificação no objetivo inicial do projeto: calcular as invariantes para todos os isômeros com n de 20 até 130, e para todos os isômeros IPR com n de 132 até 160. Uma alteração que efetivamente elevou o limite além do inicialmente planejado e diminuiu drasticamente o número total de isômeros a considerar, ao mesmo tempo em que mantinha no conjunto os isômeros considerados mais promissores (IPR).

De posse dos dados de entrada, restava apenas definir quais invariantes seriam comparadas e em que ordem elas seriam consideradas. Essa filtragem, necessária devido ao tempo inerentemente limitado disponível em um projeto de mestrado, teve como propósito permitir um

equilíbrio entre documentar o maior número possível de invariantes e dar prioridade às mais importantes entre elas. A lista final com as invariantes que serão comparadas ficou composta por:

- (1) Critério de Fowler-Manolopoulos;
- (2) Diâmetro;
- (3) Índice de Wiener
- (4) Frustração Bipartida de Arestas;
- (5) Número de Independência;
- (6) Número de Emparelhamentos Perfeitos;
- (7) Número de Fries;
- (8) Número de Taylor;

Cabe mencionar aqui os principais motivos por trás da exclusão de algumas invariantes desse plano de trabalho. O primeiro é a suposta baixa eficácia da invariante em relação à outras mais importantes (de acordo com a literatura encontrada até o presente momento) na previsão da estabilidade de fulerenos, especialmente para valores de n acima de 100. Essa é a situação, por exemplo, do menor autovalor da matriz de adjacência – mencionado brevemente na Seção 3.6 – e da constante de Cheeger – alvo da Seção 3.10.

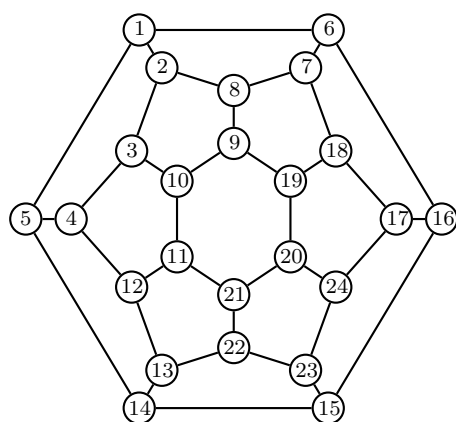
O segundo motivo, talvez mais importante, é a dificuldade associada à implementação dos algoritmos para calcular algumas das invariantes, que acabaria tomando uma parte excessiva do tempo total disponível para o desenvolvimento do projeto. Nesse caso em particular estão o número de independência *closed-shell* e (novamente) a constante de Cheeger. Há também motivos menos importantes, como no caso do número de saturação, excluído por falta de tempo hábil, e no da regra dos pentágonos isolados, que não precisa realmente ser calculada pois a base de dados já fornece separadamente os isômeros IPR.

4.2 Código Planar

Uma outra característica importante da base de dados *House of Graphs* é o formato escolhido para armazenar os grafos de fulereno em arquivos. Conhecido como *código planar* (do inglês *planar code*), esse formato específico para grafos planares não apenas faz uso eficiente do espaço em disco, como também fornece acesso direto a uma imersão do grafo. Essas duas vantagens são úteis por dois motivos: em primeiro lugar, devido aos grafos em questão serem, embora pequenos, em número proibitivamente alto (especialmente para valores elevados de n); e, em segundo lugar, por algumas das invariantes acima realmente necessitarem de uma imersão do grafo (ou de seu dual, calculado usando a imersão).

A versão do código planar utilizada neste trabalho representa o grafo (imerso no plano) através de uma sequência de bytes interpretados como caracteres sem sinal, numerando os vértices de 1 até um máximo de 255 – grafos com mais de 255 vértices precisam de um formato ligeiramente alterado, algo que não foi necessário aqui. Note que o conteúdo de um arquivo com essa representação, se aberto num editor de texto comum, vai acabar sendo visto como vários caracteres sem sentido. Para ajudar a evitar confusões desse tipo, e para sinalizar o começo do arquivo, todo arquivo em código planar deve começar com um cabeçalho formado pelos 15 caracteres legíveis `>>planar_code<<`.

O primeiro byte após o cabeçalho corresponde ao número de vértices, $v(G)$, do grafo. Em seguida, devem aparecer $v(G)$ seções, sendo que cada i -ésima seção ($1 \leq i \leq v(G)$) contém os vizinhos do vértice i listados em sentido horário, seguidos por um byte zero. A Figura 4.1 mostra, como um exemplo, uma versão rotulada do grafo do C_{24} e a lista de vizinhos correspondente a cada vértice. Não são utilizados caracteres de fim-de-linha em nenhum ponto do arquivo, e, se existir algum byte depois da última seção de um grafo, ele é o número de vértices do grafo seguinte (vários grafos podem estar armazenados em sequência).



Vértice	Vizinhos	Vértice	Vizinhos
1	5 6 2	13	12 22 14
2	1 8 3	14	13 15 5
3	2 10 4	15	14 23 16
4	3 12 5	16	15 17 6
5	4 14 1	17	16 24 18
6	1 16 7	18	17 19 7
7	6 18 8	19	18 20 9
8	7 9 2	20	19 24 21
9	8 19 10	21	20 22 11
10	9 11 3	22	21 23 13
11	10 21 12	23	22 24 15
12	11 13 4	24	23 20 17

Figura 4.1: C_{24} com vértices rotulados e a lista de vizinhos correspondente a cada vértice.

4.3 Implementação

Por maior familiaridade do proponente com a linguagem, todas as implementações foram realizadas em C++, e o planejado é disponibilizar (de forma pública e livre para atividades acadêmicas) todo o código produzido durante a elaboração do trabalho. Entretanto, conforme se verá a seguir, em certos pontos foi necessário utilizar algumas bibliotecas numéricas ou outras bibliotecas especiais, que podem estar sujeitas a licenças de uso específicas.

Outro ponto importante a ressaltar antes de prosseguir com a discussão é o fato de ter sido construído um arcabouço, bastante extenso e comum a todos os programas (replicado com pequenas alterações pontuais), responsável pela leitura e escrita de arquivos, inclusão de cabeçalhos, e outras pequenas coisas menos dignas de nota. Logo, com o objetivo de não cansar o leitor, nas subseções a seguir apenas os trechos de código mais relevantes serão exibidos, e os cortes serão realizados prezando sempre que possível pela unidade de um trecho auto-contido. Porém, em alguns lugares podem acabar aparecendo variáveis não declaradas, e nesses casos tentar-se-á fornecer uma explicação apropriada no texto.

4.3.1 Critério de Fowler-Manolopoulos

Com o objetivo de facilitar a tradução da fórmula em código, para esta primeira invariante foi escolhido um cálculo em três etapas. Na primeira etapa, criou-se um programa (denominado *planar2dual*) para obter um conjunto de grafos duais a partir do conjunto de grafos de fulereno. Na segunda etapa, são percorridos os vizinhos dos vértices de grau 6 no dual e são calculados os índices de vizinhança hexagonais. Por fim, na terceira etapa, aplica-se diretamente a fórmula exibida na Seção 3.2.

O programa *planar2dual* é uma implementação de um pequeno algoritmo adaptado de (GUEDES, 1995; GUIBAS e STOLFI, 1985) e descrito brevemente a seguir. Dado um grafo plano $G = (V, E)$, para cada aresta $\{u, v\} \in E$ crie dois pares ordenados: (u, v) e (v, u) . Cada um dos dois pares correspondentes a uma aresta pode ser visto como uma “meia-aresta” que aponta do primeiro elemento do par para o segundo. Seja A o conjunto de todos os pares ordenados criados dessa forma. Defina três funções de A em A :

- $symmetric(u, v)$: retorna (v, u) , o par simétrico ao fornecido;
- $nextV(u, v)$: retorna (u, x) , o par correspondente ao vértice, x , seguinte a v na lista de vizinhos de u . Pode ser interpretada como retornando a próxima aresta incidente ao vértice u depois de $\{u, v\}$, no sentido horário;
- $nextF(u, v)$: retorna $nextV(symmetric(u, v))$, que corresponde ao próximo par em torno da face esquerda (indo de u para v) no sentido anti-horário. Dito de outra forma, o par (v, x) que segue o contorno da face logo depois de (u, v) .

Em destaque na Figura 4.2 estão quatro pares, exemplificando cada um desses conceitos: i) o par $(3, 4)$, tomado como base e representado na cor vermelha, ii) o par $(4, 3)$, simétrico ao par $(3, 4)$ e representado em azul, iii) o par $(3, 2)$, correspondente a $nextV(3, 4)$ e representado em laranja, e iv) o par $(4, 10)$, correspondente a $nextF(3, 4)$ e representado em verde. Note que sucessivas chamadas à função $nextF()$ nesse caso retornariam os pares em torno da face f_2 no sentido anti-horário.

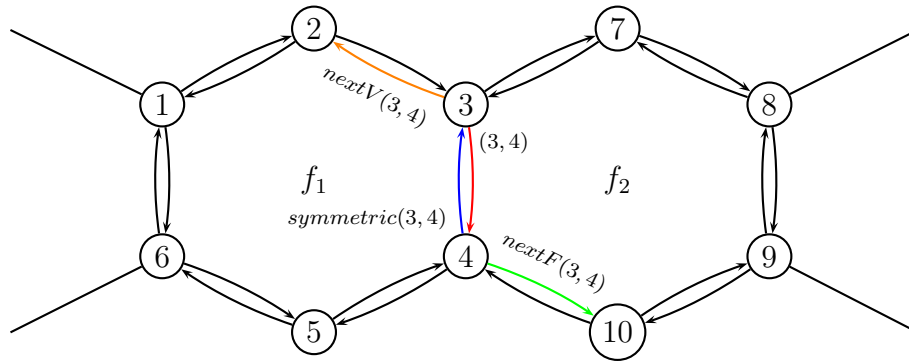


Figura 4.2: Exemplo de pares ordenados e faces do programa *planar2dual*.

Continuando com a explicação do algoritmo, o passo seguinte é associar a cada par ordenado de A uma das duas faces de G adjacentes à aresta que corresponde àquele par. Tomando como ponto de referência a orientação mencionada acima (de u para v num par (u, v)), convencionou-se associar a face situada à esquerda do par, de forma que a face direita fique associada ao elemento simétrico. Por exemplo, na Figura 4.2 o par $(3, 4)$ estaria associado à face f_2 , e o par $(4, 3)$ à face f_1 . Para fazer a associação, repetem-se os passos a seguir até que não haja mais elementos sem face associada:

- (1) Partindo de algum elemento $a \in A$ sem face associada, crie um rótulo, f , para a nova face (à esquerda de a);
- (2) Percorra o ciclo formado pelas arestas $a, nextF(a), nextF(nextF(a)), \dots$ (arestas ao redor da face f) associando cada uma delas à face f ;
- (3) Associe f a algum dos elementos do ciclo (convencionou-se associar f a a , o elemento selecionado anteriormente).

O último passo do algoritmo é criar os vértices e arestas do grafo dual, D , de G . Para cada face f rotulada da maneira descrita acima, crie um vértice v em D . Depois, novamente para cada face f , percorra os elementos de A associados a f . Para cada elemento encontrado

nesse percurso, obtenha a face, h , associada ao seu elemento simétrico e crie uma nova aresta entre os vértices de D correspondentes a f e h .

De posse dos duais dos grafos de fulereno, nos quais cada hexágono corresponde a um vértice de grau 6, a segunda etapa para encontrar o critério de Fowler-Manolopoulos é simplesmente percorrer os vizinhos de cada um desses vértices e calcular os índices de vizinhança hexagonal (isto é, contar quantos dos vizinhos também possuem grau 6). Uma vez calculados os índices, pode-se criar um arranjo representando a assinatura hexagonal do isômero e passá-lo como argumento para a função exibida na Figura 4.3, que corresponde à terceira e última etapa do cálculo da invariante.

4.3.2 Diâmetro e Índice de Wiener

Conforme explicado na Seção 3.3, o diâmetro e o índice de Wiener são duas invariantes diretamente relacionadas às distâncias entre os vértices de um grafo. Consequentemente, este autor preferiu escrever um único programa para calculá-las simultaneamente, ao invés de programas separados que acabariam sendo quase exatamente os mesmos.

O primeiro passo nessa direção é encontrar a *matriz de distâncias* do grafo de fulereno recebido como entrada, algo quase trivial e que torna ainda mais intuitivo o código para computar as duas invariantes. Como, por convenção, a distância sempre será um valor inteiro não negativo (a “distância unitária” de uma única aresta corresponde ao valor 1), a matriz de distâncias é representada aqui através de um arranjo bidimensional no qual cada posição (i, j) armazena um valor inteiro sem sinal.

Inicialmente, a matriz toda (menos as diagonais, preenchidas com 0) é preenchida com a constante `MAX_UNSIGNED`, que corresponde ao maior valor inteiro sem sinal e indica, simbolicamente, a distância infinita ∞ . Conforme o grafo de fulereno é lido do arquivo em *planar code*, as posições (i, j) são preenchidas com 1 se os vértices i e j possuírem uma aresta entre eles.

Por fim, a matriz parcialmente preenchida é passada – junto com seu tamanho, n , e uma variável que armazenará o diâmetro, *diameter* – para a função em destaque na Figura 4.4. Nela,

é possível ver uma implementação do algoritmo de Floyd-Warshall adaptada de (CORMEN et al., 2009), que termina de construir a matriz de distâncias, e o trecho de código que percorre a matriz somando os valores (no caso do índice de Wiener) e procurando o máximo (no caso do diâmetro).

4.3.3 Frustração Bipartida de Arestas

De forma semelhante ao programa do critério de Fowler-Manolopoulos, o programa correspondente à frustração bipartida de arestas não recebe diretamente como entrada os grafos de fulereno, mas sim seus duais. Partir do dual torna mais simples a tarefa de construir o grafo completo K_{12} , com pesos nas arestas representando as distâncias entre as doze faces pentagonais do fulereno e sobre o qual se calcula um emparelhamento perfeito de peso mínimo.

Com o objetivo de facilitar a manipulação dos duais, foi codificada uma classe *ad-hoc* para representar grafos, de maneira bem simples e direta, usando listas de adjacência. Conforme ilustrado na Figura 4.5, a função que realiza o cálculo da invariante recebe como argumento um objeto dessa classe, previamente inicializado com as informações do dual lido do arquivo em código planar. Veja também que ela recebe dois outros argumentos: os arranjos *level* (de inteiros sem sinal) e o arranjo *visited* (de booleanos), ambos alocados e destruídos fora da função em questão, mas usados dentro dela na busca em largura que calcula as distâncias entre os vértices de grau 5 do dual (e “zerados” antes de cada uso).

O grafo K_{12} em si é representado usando dois outros arranjos: um para armazenar o peso (*weights*), e outro para armazenar os extremos de cada aresta (*edges*). Há também um objeto *map<unsigned, int>* da biblioteca padrão C++, responsável por mapear os índices dos vértices de grau 5 no dual para a faixa $[0, 11]$ de índices usada no K_{12} . Note que, nesse caso, a escolha de arranjos e a representação usando inteiros com sinal é uma exigência da biblioteca, explicada mais abaixo, que calcula o emparelhamento perfeito.

Em linhas gerais, o mapeamento é feito como descrito a seguir. A lista de vértices do dual é percorrida e, partindo de cada vértice de grau 5 (uma face pentagonal no grafo de fulereno) que ainda não tenha sido contabilizado, é feita uma busca em largura até que todos os outros

```

double fowlerManolopoulosCriterion(unsigned* hexNeighbourhoodIndices) {
    unsigned summHk = 0;
    unsigned summKHk = 0;
    unsigned summK2Hk = 0;
    for (unsigned k = 0; k < 7; ++k) {
        summHk += hexNeighbourhoodIndices[k];
        summKHk += k * hexNeighbourhoodIndices[k];
        summK2Hk += k * k * hexNeighbourhoodIndices[k];
    }

    double k2 = ((double) summK2Hk) / ((double) summHk);
    double k = (((double) (summKHk * summKHk)) / ((double) (summHk * summHk)));

    return sqrt(k2 - k);
}

```

Figura 4.3: Código para calcular o Critério de Fowler-Manolopoulos.

```

unsigned wienerIndexAndDiameter(unsigned** distances, unsigned n, unsigned& diameter) {
    // Floyd-Warshall algorithm
    for (unsigned k = 0; k < n; ++k) {
        for (unsigned i = 0; i < n; ++i) {
            for (unsigned j = 0; j < n; ++j) {
                if (distances[i][k] != MAX_UNSIGNED
                    && distances[k][j] != MAX_UNSIGNED
                    && distances[i][j] > distances[i][k] + distances[k][j]) {
                    distances[i][j] = distances[i][k] + distances[k][j];
                }
            }
        }
    }

    unsigned wienerIndex = 0;
    diameter = 0;
    for (unsigned i = 0; i < n; ++i) {
        for (unsigned j = 0; j < n; ++j) {
            if (i < j) {
                wienerIndex += distances[i][j];
            }
            if (diameter < distances[i][j]) {
                diameter = distances[i][j];
            }
        }
    }

    return wienerIndex;
}

```

Figura 4.4: Código para calcular o Índice de Wiener e o Diâmetro.

vértices de grau 5 sejam encontrados. Novos vértices são inseridos no K_{12} conforme faces ainda não existentes são encontradas. O nível em que cada vértice é encontrado na busca corresponde justamente à distância entre a face pentagonal associada ao vértice “de origem” e a face associada ao vértice “de destino”, e também é o peso da aresta que liga os dois vértices correspondentes no K_{12} . Por fim, um emparelhamento perfeito de peso mínimo é calculado no K_{12} usando uma biblioteca especial que implementa o algoritmo Blossom V (KOLMOGOROV, 2009).

4.3.4 Número de Independência

Embora o número de independência seja uma invariante bastante conhecida no estudo dos grafos e tecnicamente de fácil implementação, e mesmo os grafos de fulereno sendo relativamente pequenos, o número elevadíssimo de grafos no conjunto de entrada e o fato de já se ter provado que o problema é *NP-Difícil* (GAREY e JOHSON, 1990) levantaram algumas preocupações sobre a melhor maneira de lidar com esta invariante no projeto.

Logo, ao invés de arriscar desperdiçar tempo escrevendo código que poderia acabar sendo *subpar* e prejudicasse os experimentos, optou-se por usar uma biblioteca especializada, como no caso da frustração bipartida de arestas acima. A biblioteca escolhida depende de uma definição alternativa do número de independência, equivalente à definição dada na Seção 3.5 e baseada no conceito de *clique*.

De forma semelhante a um conjunto independente, uma *clique*, $C \subseteq V$, de um grafo $G = (V, E)$ é um subconjunto (maximal) de vértices de G tal que o subgrafo induzido por C , $G[C]$, é completo. Denotando por $\overline{G}$ o complemento de G , o número de independência de G pode também ser definido como:

$$\alpha(G) = \text{máx}\{|C| \mid C \text{ é uma clique de } \overline{G}\}$$

Usando a definição acima e a biblioteca *mcqcd* (KONC e JANEŽIČ, 2007) para calcular cliques máximas, não foi difícil chegar no código exposto na Figura 4.6. O arranjo bidimensional de

```

int bipartiteEdgeFrustration(Graph& g, unsigned* level, bool* visited) {
    int nK12          = 12; // Number of K12 vertices
    int mK12          = 66; // Number of K12 edges
    int e              = 0;
    PerfectMatching *pm = new PerfectMatching(nK12, mK12);
    unsigned n         = g.getNumberOfVertices();
    int mapV           = 0;
    int weights [mK12]; int edges  [mK12 * 2];
    map<unsigned, int> mapping; // Maps pentagon vertices in g to the range [0..11]
    for (unsigned i = 1; i <= n; ++i) {
        if (g.getDegree(i) == 5) {
            if (mapV < nK12) {
                mapping[i] = mapV;
                ++mapV;
            }
            queue<unsigned> bfs;
            for (unsigned j = 0; j <= n; ++j) {
                level[j] = 0;
                visited[j] = false;
            }
            unsigned countPentagons = 0;
            bfs.push(i);
            visited[i] = true;
            while (!bfs.empty()) {
                unsigned v = bfs.front();
                bfs.pop();
                if (g.getDegree(v) == 5) {
                    if (i < v) {
                        map<unsigned, int>::iterator it = mapping.find(v);
                        if (it == mapping.end() && mapV < nK12) {
                            mapping[v] = mapV;
                            ++mapV;
                        }
                        weights[e] = (int)level[v];
                        edges[2 * e] = mapping[i];
                        edges[2 * e + 1] = mapping[v];
                        pm->AddEdge(edges[2 * e], edges[2 * e + 1], weights[e]);
                        ++e;
                    }
                    ++countPentagons;
                    if (countPentagons == 12) {
                        break;
                    }
                }
            }
            for (AdjacencyListIterator it = g.getNeighboursBegin(v);
                 it != g.getNeighboursEnd(v); ++it) {
                unsigned w = (*it);
                if (!visited[w]) {
                    bfs.push(w);
                    visited[w] = true;
                    level[w] = level[v] + 1;
                }
            }
        }
    }
}

struct PerfectMatching::Options options;
options.verbose = false;
pm->options = options;
pm->Solve();
int cost = -1;
int res = CheckPerfectMatchingOptimality(nK12, mK12, edges, weights, pm);
if (res == 0) {
    cost = (int)ComputePerfectMatchingCost(nK12, mK12, edges, weights, pm);
}
delete pm;
return cost;
}

```

Figura 4.5: Trecho do código que calcula a Frustração Bipartida de Arestas.

booleanos a e o inteiro n representam, respectivamente, a matriz de adjacência do complemento do grafo de fulereno e seu tamanho. O processo para criar o complemento é trivial: inicialmente preenchida com valores *true* (correspondendo a um grafo completo), as arestas vão sendo removidas da matriz (i.e., os elementos recebem o valor *false*) conforme elas são lidas do arquivo em código planar.

```
int independenceNumber(bool** a, int n) {
    Maxclique m(a, n);
    int* qmax;
    int qsize;
    m.mcqdyn(qmax, qsize);
    delete[] qmax;
    return qsize;
}
```

Figura 4.6: Código para calcular o Número de Independência.

4.3.5 Número de Estruturas de Kekulé

Conforme explicado anteriormente, o número de estruturas de Kekulé de um grafo G é obtido através do algoritmo FKT, cujo pseudocódigo pode ser visto no Algoritmo 1 e que pode ser descrito de forma resumida como dividido em quatro partes principais: i) a construção da árvore geradora T , com arestas orientadas, ii) a construção da árvore H com um vértice para cada face de G , iii) a orientação das arestas ainda não orientadas de G , e iv) o cálculo do valor absoluto da raiz quadrada do determinante da matriz de adjacência, M , de G .

Devido à necessidade do cálculo de determinantes, e da representação do grafo através de uma matriz de adjacência, durante a elaboração do programa este autor decidiu recorrer a uma biblioteca, bastante conhecida, para operações de álgebra linear: a biblioteca *Eigen* (GUENNEBAUD e al., 2010). O grafo é representado através de um objeto (*pfaffianOrientation*) da classe *Eigen::MatrixXd*, uma matriz quadrada de valores *double* que pode ter seu tamanho definido em tempo de execução.

Cada posição (i, j) da matriz é preenchida com uma de quatro constantes possíveis:

NO_EDGE (indicando que não há aresta entre os vértices i e j), *UNDIRECTED* (indicando que a aresta existe, mas ainda não foi orientada), *FROM_I_TO_J* (indicando que a aresta existe e aponta de i para j), e *FROM_J_TO_I* (indicando que existe e aponta de j para i). Inicialmente, todas as posições são preenchidas com a constante *NO_EDGE*. Conforme o grafo vai sendo lido do arquivo em código planar, as posições correspondentes às arestas são preenchidas com *UNDIRECTED*. Os outros dois valores são usados durante o algoritmo FKT em si.

Cabe aqui também um comentário sobre as árvores T e H e sobre a maneira escolhida para representá-las no código. A árvore T serve apenas como uma abstração para definir uma orientação inicial “válida” em algumas arestas de G (válida no sentido de poder ser transformada numa orientação Pfaffiana). De modo semelhante, a árvore H , que possui o mesmo conjunto de vértices que o dual de G e que é desmontada uma folha de cada vez, é usada apenas para fornecer uma ordem válida para processar as faces que ainda possuem alguma aresta não orientada (referindo o leitor novamente à Figura 3.3, note que, a cada passo, nas faces que correspondem a uma folha em H , resta apenas uma única aresta a orientar). Assim sendo, preferiu-se evitar a criação de classes específicas para manipular árvores e codificar T e H conforme descrito abaixo.

Como a árvore T possui o mesmo conjunto de vértices que G , ela pode ser representada simplesmente por seu conjunto de arestas, que, por sua vez, pode ser indicado através da orientação das arestas na matriz de adjacência de G . Usando uma simples busca em profundidade, as arestas que seriam inseridas num hipotético objeto árvore são, ao invés disso, orientadas para indicar que fazem parte de T . Note que a orientação neste estágio pode ser arbitrária, mas convencionou-se orientar as arestas da raiz para as folhas, conforme descrito no trecho de código da Figura 4.7.

No caso da árvore H , o processo é um pouco mais complexo, reutilizando parte do código responsável pela criação dos duais e que foi mencionado brevemente na Seção 4.3.1 acima. A árvore em si é representada através de uma lista na qual são armazenadas estruturas representando as faces de G (i.e., os vértices do dual), que armazenam o rótulo (*ID*) da face,

```

dfs.push(0);
visited[0] = true;
while (!dfs.empty()) {
    unsigned v = dfs.top();
    dfs.pop();
    for (unsigned w = 0; w < nCurrent; ++w) {
        if (pfaffianOrientation(v, w) == UNDIRECTED && !visited[w]) {
            dfs.push(w);
            visited[w] = true;
            // Give a direction to this edge, indicating that it is now part of T
            pfaffianOrientation(v, w) = FROM_I_TO_J;
            pfaffianOrientation(w, v) = FROM_J_TO_I;
        }
    }
}

```

Figura 4.7: Trecho de código para encontrar uma árvore geradora, T , de G .

um booleano indicando se já ela foi visitada, e um ponteiro para o par ordenado associado. Cada par, por sua vez, armazena os índices do primeiro e segundo vértices e a ID da face associada. A ideia-chave nesse ponto é justamente a ordem de inserção dos vértices. Se o objetivo, posteriormente, é desmontar a árvore uma folha de cada vez, processando a face correspondente, basta inserir os vértices de forma que o último elemento da lista seja sempre uma folha de H . Algo que pode ser feito com uma simples busca em profundidade no dual de G , partindo de algum vértice escolhido arbitrariamente. O critério para inserção, como pode ser visto na Figura 4.8 e como indicado pelo algoritmo FKT, é se a aresta entre as duas faces não foi ainda orientada.

Uma vez que a lista de faces seja preenchida, ela pode ser percorrida na ordem inversa, e cada face considerada (que pode ser vista como correspondendo a uma folha de H naquele momento) tem sua única aresta *UNDIRECTED* orientada de forma que o número total de arestas no sentido horário seja ímpar. Quando a última face é orientada, a matriz de adjacência corresponde a uma orientação Pfaffiana de G e resta apenas calcular o valor absoluto da raiz quadrada de seu determinante. A Figura 4.9 ilustra essa última etapa.

4.3.6 Número de Fries

Relembrando o que já foi dito na Seção 3.8, o cálculo número de Fries pode ser formulado como um problema de programação inteira. Portanto, é natural que uma das primeiras preo-

```

list<Face> hLeaves;
dfs.push(1);
faces[1].visited = true;
while (!dfs.empty()) {
    unsigned fID = dfs.top();
    dfs.pop();
    OrderedVertexPair* pair = faces[fID].pair;
    unsigned u = pair->first;
    unsigned v = pair->second;
    do {
        if (pfaffianOrientation(pair->first - 1, pair->second - 1) == UNDIRECTED) {
            OrderedVertexPair* other = symmetric((*pair), a);
            unsigned neighbouringFaceID = other->face;
            if (!faces[neighbouringFaceID].visited) {
                dfs.push(neighbouringFaceID);
                hLeaves.push_back(faces[neighbouringFaceID]);
                faces[neighbouringFaceID].visited = true;
            }
        }
        pair = nextF((*pair), a);
    } while (pair->first != u || pair->second != v);
}

```

Figura 4.8: Trecho de código para calcular a árvore H .

```

for (list<Face>::reverse_iterator rit = hLeaves.rbegin(); rit != hLeaves.rend(); ++rit) {
    // Take each face ("leaf" on H), and iterate through its bordering edges
    OrderedVertexPair* pair = (*rit).pair;
    unsigned u = pair->first;
    unsigned v = pair->second;
    unsigned countClockwise = 0;
    unsigned lastUndirectedU = 0;
    unsigned lastUndirectedV = 0;
    do {
        if (pfaffianOrientation(pair->first - 1, pair->second - 1) == FROM_J_TO_I) {
            ++countClockwise;
        } else if (pfaffianOrientation(pair->first - 1, pair->second - 1) == UNDIRECTED) {
            lastUndirectedU = pair->first;
            lastUndirectedV = pair->second;
        }
        pair = nextF((*pair), a);
    } while (pair->first != u || pair->second != v);

    // Now orient the last undirected edge around this face so that the number of
    // clockwise oriented edges is odd.
    if (countClockwise % 2 == 0) {
        // Orient it clockwise
        pfaffianOrientation(lastUndirectedU - 1, lastUndirectedV - 1) = FROM_J_TO_I;
        pfaffianOrientation(lastUndirectedV - 1, lastUndirectedU - 1) = FROM_I_TO_J;
    } else {
        // Orient it counterclockwise
        pfaffianOrientation(lastUndirectedU - 1, lastUndirectedV - 1) = FROM_I_TO_J;
        pfaffianOrientation(lastUndirectedV - 1, lastUndirectedU - 1) = FROM_J_TO_I;
    }
}

double kekule = abs(sqrt(pfaffianOrientation.determinant()));

```

Figura 4.9: Fim da construção da orientação Pfaffiana e cálculo da invariante.

cupações ao escrever o código fosse a escolha de uma biblioteca não apenas capaz de resolver esse tipo de problema, mas que também fosse de fácil utilização (levando em conta a falta de experiência deste autor com programação inteira).

Após alguns experimentos com as bibliotecas *GLPK*², *SCIP* (ACHTERBERG, 2009), e *lp_solve* (BERKELAAR, EIKLAND e NOTEBAERT), acabou-se escolhendo a *GLPK* principalmente pela facilidade de integração com o código já existente. As outras duas bibliotecas se provaram significativamente difíceis de compilar, além de bem mais complicadas de usar do que a *GLPK*. Infelizmente, um aspecto negativo decorrente da maneira escolhida para integrar a biblioteca foi a natureza extremamente prolixa do programa resultante, algo que ficará evidente logo adiante, quando serão mostrados os trechos de código correspondentes às principais etapas do algoritmo.

Na Figura 4.10 pode ser vista a criação do problema de programação inteira, assim como a criação de uma variável binária X_{vw} correspondendo a cada aresta (v, w) no grafo. Novamente, note que é usada uma versão alterada do esquema de pares ordenados mencionado anteriormente (o arranjo bidimensional a armazena os três pares que “saem” de cada vértice v), mas que apenas uma variável é criada por aresta (se um dos pares da aresta já foi processado, a mesma variável é atribuída ao par simétrico). Perceba também que os números de linhas $((4 * nCurrent) - 60)$ e de colunas $((5 * nCurrent)/2 - 20)$ da variável lp correspondem, respectivamente, ao número total de restrições do problema (uma para cada vértice e seis para cada hexágono) e ao número total de variáveis estruturais do problema (uma para cada aresta, duas para cada hexágono).

Já o trecho da Figura 4.11, bem mais extenso, é responsável pela criação das variáveis Y_s e Z_s e das restrições (denotadas na figura por “q”) associadas a cada hexágono, assim como das restrições associadas a cada vértice (denotadas por “p”). Os arranjos ia , ja e ar servem para armazenar os coeficientes de cada termo da restrição, e são passados, mais adiante e junto o ponteiro lp , para a função realmente responsável pelo cálculo. Veja que as restrições são criadas seguindo o padrão da biblioteca *GLPK*, que não permite a representação direta de

²Acesse o elo <https://www.gnu.org/software/glpk/>

desigualdades entre duas variáveis estruturais. Em particular, as restrições seguem o formato abaixo (onde j, k , e l seriam os únicos três vizinhos de um vértice i no grafo de fulereno e s denota um hexágono):

$$X_{ij} + X_{ik} + X_{il} = P \quad 1 \leq P \leq 1$$

$$Y_s - X_{s1} = Q_1 \leq 0 \quad Z_s - X_{s2} = Q_2 \leq 0$$

$$Y_s - X_{s3} = Q_3 \leq 0 \quad Z_s - X_{s4} = Q_4 \leq 0$$

$$Y_s - X_{s5} = Q_5 \leq 0 \quad Z_s - X_{s6} = Q_6 \leq 0$$

```
// Create a new GLPK problem instance
glp_prob *lp;
lp = glp_create_prob();
glp_term_out(GLP_OFF);
glp_set_prob_name(lp, "Fries_Number");
glp_set_obj_dir(lp, GLP_MAX);
glp_add_rows(lp, ((4 * nCurrent) - 60));
glp_add_cols(lp, ((5 * nCurrent) / 2) - 20);

unsigned variableID = 1;
for (unsigned v = 1; v <= numberOfVertices; ++v) {
    inputFile.get(byte);
    unsigned w = (unsigned)((unsigned char)byte);
    unsigned k = 0;
    while (w != 0) {
        a[v][k].second = w;
        if (v < w) {
            // Create a new structural binary variable
            glp_set_col_name(lp, variableID, "Xvw");
            glp_set_col_bnds(lp, variableID, GLP_DB, 0.0, 1.0);
            glp_set_obj_coef(lp, variableID, 0.0);
            glp_set_col_kind(lp, variableID, GLP_BV);
            a[v][k].varID = variableID;
            ++variableID;
        } else {
            // Variable already created, assign it to appropriate pair
            if (a[w][0].second == v) {
                a[v][k].varID = a[w][0].varID;
            } else if (a[w][1].second == v) {
                a[v][k].varID = a[w][1].varID;
            } else {
                a[v][k].varID = a[w][2].varID;
            }
        }
        ++k;
        inputFile.get(byte);
        w = (unsigned)((unsigned char)byte);
    }
}
```

Figura 4.10: Primeira parte do código para calcular o número de Fries.

```

unsigned faceID      = 1;
unsigned restrictionID = 1;
unsigned coefficientID = 0;
for (unsigned i = 1; i <= nCurrent; ++i) {
    for (unsigned j = 0; j < 3; ++j) {
        if (a[i][j].face == 0) {
            unsigned u = a[i][j].first;
            unsigned v = a[i][j].second;
            a[i][j].face = faceID;
            OrderedVertexPair* nextPair = nextF(a[i][j], a);
            unsigned sides = 1;
            while (nextPair->first != u || nextPair->second != v) {
                nextPair->face = faceID;
                nextPair = nextF(*nextPair, a);
                ++sides;
            }
            if (sides == 6) {
                glp_set_col_name(lp, variableID, "Ys");
                glp_set_col_bnds(lp, variableID, GLP_DB, 0.0, 1.0);
                glp_set_obj_coef(lp, variableID, 1.0);
                glp_set_col_kind(lp, variableID, GLP_BV);
                unsigned ys = variableID;
                ++variableID;
                glp_set_col_name(lp, variableID, "Zs");
                glp_set_col_bnds(lp, variableID, GLP_DB, 0.0, 1.0);
                glp_set_obj_coef(lp, variableID, 1.0);
                glp_set_col_kind(lp, variableID, GLP_BV);
                unsigned zs = variableID;
                ++variableID;

                glp_set_row_name(lp, restrictionID, "q");
                glp_set_row_bnds(lp, restrictionID, GLP_DB, 0.0, 1.0);
                ++coefficientID;
                ia[coefficientID] = restrictionID;
                ja[coefficientID] = ys;
                ar[coefficientID] = -1.0;
                ++coefficientID;
                ia[coefficientID] = restrictionID;
                ja[coefficientID] = a[i][j].varID;
                ar[coefficientID] = 1.0;
                ++restrictionID;

                bool alternate = true;
                nextPair = nextF(a[i][j], a);
                while (nextPair->first != u || nextPair->second != v) {
                    glp_set_row_name(lp, restrictionID, "q");
                    glp_set_row_bnds(lp, restrictionID, GLP_DB, 0.0, 1.0);
                    ++coefficientID;
                    ia[coefficientID] = restrictionID;
                    if (alternate) {
                        ja[coefficientID] = zs;
                        alternate = false;
                    } else {
                        ja[coefficientID] = ys;
                        alternate = true;
                    }
                    ar[coefficientID] = -1.0;
                    ++coefficientID;
                    ia[coefficientID] = restrictionID;
                    ja[coefficientID] = nextPair->varID;
                    ar[coefficientID] = 1.0;
                    ++restrictionID;
                    nextPair = nextF(*nextPair, a);
                }
            }
            ++faceID;
        }
    }
    glp_set_row_name(lp, restrictionID, "p");
    glp_set_row_bnds(lp, restrictionID, GLP_FX, 1.0, 1.0);
    ++coefficientID;
    ia[coefficientID] = restrictionID;
    ja[coefficientID] = a[i][0].varID;
    ar[coefficientID] = 1.0;
    ++coefficientID;
    ia[coefficientID] = restrictionID;
    ja[coefficientID] = a[i][1].varID;
    ar[coefficientID] = 1.0;
    ++coefficientID;
    ia[coefficientID] = restrictionID;
    ja[coefficientID] = a[i][2].varID;
    ar[coefficientID] = 1.0;
    // Increment restriction afterwards
    ++restrictionID;
}

```

Figura 4.11: Segunda parte do código para calcular o número de Fries.

Um outro ponto importante a lembrar é que a definição de quais arestas de uma face hexagonal s correspondem à variável Y_s e quais correspondem à variável Z_s não é relevante. Elas representam as duas únicas configurações benzenóides possíveis do hexágono s , e não importa para a função objetivo qual delas é igual a 1, apenas *se uma delas* é igual 1. Portanto, convencionou-se, no código, associar a primeira aresta encontrada de cada hexágono à Y_s , e alternar a partir daí conforme s é percorrido.

Finalmente, a Figura 4.12 mostra a terceira e última parte do programa para calcular o número de Fries. Este trecho corresponde, simplesmente, à definição de algumas opções de execução e à chamada do procedimento de resolução da biblioteca *GLPK*. Os valores dos parâmetros foram escolhidos empiricamente, avaliando (em instâncias menores do problema) quais opções levavam à uma conclusão mais rápida.

```
// Fine tune parameters
glp_smcp sparam;
glp_init_smcp(&sparam);
sparam.meth = GLP_PRIMAL;
sparam.pricing = GLP_PT_PSE;
sparam.r_test = GLP_RT_HAR;

glp_iocp param;
glp_init_iocp(&param);
param.br_tech = GLP_BR_FFV;
param.bt_tech = GLP_BT_BLB;
param.pp_tech = GLP_PP_NONE;

// Calculate fries number
glp_load_matrix(lp, coefficientID, ia, ja, ar);
glp_simplex(lp, &sparam);
glp_intopt(lp, &param);
double fries = glp_mip_obj_val(lp);

// GLPK cleanup
glp_delete_prob(lp);
glp_free_env();
```

Figura 4.12: Terceira parte do código para calcular o número de Fries.

4.3.7 Número de Taylor

Para encerrar este capítulo, resta apenas fazer alguns comentários a respeito do código para calcular o número de Taylor. Como essa invariante é apenas uma modificação do número de Fries, grande parte do código é exatamente igual, então o foco desta seção será nas diferenças entre as duas. Para começar, o problema de encontrar $T_a(G)$ (lembrando que essa foi a única variação implementada), para um grafo de fulereno G , pode ser dividido em dois subproblemas separados, ambos tratáveis como da mesma forma que o número de Fries – com programação inteira e a biblioteca GLPK.

No primeiro subproblema (encontrar $n_\Theta(G)$), conforme explicado na Seção 3.9, são criadas as mesmas variáveis binárias X_{vw} para cada aresta e restrições para cada vértice, além de duas variáveis novas: P_{vw} e A_{vw} . A Figura 4.13 ilustra a criação dos novos elementos. Note que embora todas as variáveis P_{vw} sejam criadas inicialmente com o valor 0 (indicando que a aresta correspondente não faz parte de uma face pentagonal), esse valor é modificado para 1 mais adiante se a aresta realmente faz parte de um pentágono (no trecho que seria equivalente ao da Figura 4.11). Perceba também que a restrição $A_{vw} = X_{vw} \wedge P_{vw}$ é reformulada da seguinte maneira, para seguir o padrão GLPK:

$$A_{vw} - X_{vw} = Q_1 \leq 0 \quad ; \quad A_{vw} - P_{vw} = Q_2 \leq 0 \quad ; \quad A_{vw} - X_{vw} - P_{vw} = Q_3 \geq -1$$

Uma vez que o valor da função objetivo da primeira etapa seja encontrado e armazenado, pode-se começar a tratar do segundo subproblema: calcular um emparelhamento que maximize $n_F(M)$ mas que também atinja o valor $n_\Theta(G)$. Para resolvê-lo, basta acrescentar as variáveis A_{vw} e P_{vw} (e restrições correspondentes) ao processo descrito acima no cálculo do número de Fries, além de uma restrição que obrigue o somatório de todas as variáveis A_{vw} a ter exatamente o valor $n_\Theta(G)$. A implementação dessa última etapa do processo pode ser vista na Figura 4.14, na qual se representa $n_\Theta(G)$ através da variável *numberOfMatchingEdgesOnPentagons*. Veja que, para evitar reler o grafo do arquivo de entrada, todas as variáveis e restrições do segundo problema foram criadas previamente (junto com as do primeiro) e armazenadas em objetos *vector* apropriados (*variablesLP2* e *restrictionsLP2*).

```

// Create two new binary variables, Pvw, Avw
// Assume this edge won't be part of a pentagon,
// so Pvw starts at 0 (if it is, it's changed later)
glp_set_col_name(lp, variableID, "Pvw");
glp_set_col_bnds(lp, variableID, GLP_FX, 0.0, 0.0);
glp_set_obj_coef(lp, variableID, 0.0);
a[v][k].edgePvwVarID = variableID;
++variableID;
glp_set_col_name(lp, variableID, "Avw");
glp_set_col_bnds(lp, variableID, GLP_DB, 0.0, 1.0);
glp_set_obj_coef(lp, variableID, 1.0);
glp_set_col_kind(lp, variableID, GLP_BV);
a[v][k].edgeAvwVarID = variableID;
++variableID;

// Create 3 new restrictions on Avw, Pvw, and Xvw
//  $A_{ij} \leq X_{ij} \rightarrow A_{ij} - X_{ij} \leq 0 \rightarrow A_{ij} - X_{ij} = Q \leq 0$ 
glp_set_row_name(lp, restrictionID, "Q");
glp_set_row_bnds(lp, restrictionID, GLP_UP, 0.0, 0.0);
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgeAvwVarID;
ar[coefficientID] = 1.0;
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgeXvwVarID;
ar[coefficientID] = -1.0;
++restrictionID;
//  $A_{ij} \leq P_{ij} \rightarrow A_{ij} - P_{ij} \leq 0 \rightarrow A_{ij} - P_{ij} = Q \leq 0$ 
glp_set_row_name(lp, restrictionID, "Q");
glp_set_row_bnds(lp, restrictionID, GLP_UP, 0.0, 0.0);
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgeAvwVarID;
ar[coefficientID] = 1.0;
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgePvwVarID;
ar[coefficientID] = -1.0;
++restrictionID;
//  $A_{ij} + 1 \geq X_{ij} + P_{ij} \rightarrow A_{ij} - X_{ij} - P_{ij} = Q \geq -1$ 
glp_set_row_name(lp, restrictionID, "Q");
glp_set_row_bnds(lp, restrictionID, GLP_LO, -1.0, 0.0);
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgeAvwVarID;
ar[coefficientID] = 1.0;
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgeXvwVarID;
ar[coefficientID] = -1.0;
++coefficientID;
ia[coefficientID] = restrictionID;
ja[coefficientID] = a[v][k].edgePvwVarID;
ar[coefficientID] = -1.0;
++restrictionID;

```

Figura 4.13: Criação das novas variáveis e restrições no cálculo do número de Taylor.

```

// Calculate the minimum value of n_T (number of matched edges on pentagons)
glp_load_matrix(lp, coefficientID, ia, ja, ar);
glp_simplex(lp, &sparam);
glp_intopt(lp, &param);
double numberOfMatchingEdgesOnPentagons = glp_mip_obj_val(lp);
glp_delete_prob(lp);
glp_free_env();

// Create a new GLPK problem instance
glp_prob *lp2;
lp2 = glp_create_prob();
glp_term_out(GLP_OFF);
glp_set_prob_name(lp2, "Taylor_Number");
glp_set_obj_dir(lp2, GLP_MAX);
glp_add_rows(lp2, ((17 * nCurrent) / 2) - 60 + 1);
glp_add_cols(lp2, ((11 * nCurrent) / 2) - 20);

for (vector<ProblemColumn>::iterator it = variablesLP2.begin();
     it != variablesLP2.end(); ++it) {
    ProblemColumn c = (*it);
    glp_set_col_name(lp2, c.id, "Var");
    glp_set_col_bnds(lp2, c.id, c.boundType, c.lowerBound, c.upperBound);
    glp_set_obj_coef(lp2, c.id, c.coefficient);
    glp_set_col_kind(lp2, c.id, c.kind);
}

for (vector<ProblemRow>::iterator it = restrictionsLP2.begin();
     it != restrictionsLP2.end(); ++it) {
    ProblemRow r = (*it);
    glp_set_row_name(lp2, r.id, "Rest");
    glp_set_row_bnds(lp2, r.id, r.boundType, r.lowerBound, r.upperBound);
}

// And use the previously calculated n_T to create the
// final restriction on the actual taylor number calculation
glp_set_row_name(lp2, restrictionID2, "Tp");
glp_set_row_bnds(lp2, restrictionID2, GLP_FX, numberOfMatchingEdgesOnPentagons,
                 numberOfMatchingEdgesOnPentagons);
for (vector<unsigned>::iterator it = pentagonEdgesVariablesLp2.begin();
     it != pentagonEdgesVariablesLp2.end(); ++it) {
    ++coefficientID2;
    ia2[coefficientID2] = restrictionID2;
    ja2[coefficientID2] = (*it);
    ar2[coefficientID2] = 1.0;
}
++restrictionID2;

glp_smcp sparam2;
glp_init_smcp(&sparam2);
sparam2.meth = GLP_PRIMAL;
sparam2.pricing = GLP_PT_PSE;
sparam2.r_test = GLP_RT_HAR;

glp_iocp param2;
glp_init_iocp(&param2);
param2.br_tech = GLP_BR_MFV;
param2.bt_tech = GLP_BT_BLB;
param2.pp_tech = GLP_PP_NONE;

glp_load_matrix(lp2, coefficientID2, ia2, ja2, ar2);
glp_simplex(lp2, &sparam2);
glp_intopt(lp2, &param2);
double taylor = glp_mip_obj_val(lp2);
glp_delete_prob(lp2);
glp_free_env();

```

Figura 4.14: Última etapa do cálculo do número de Taylor.

5 EXPERIMENTOS E RESULTADOS

O objetivo deste capítulo é documentar, em detalhes, os experimentos realizados com o código desenvolvido e a base de dados escolhida, além de analisar e extrair algumas conclusões sobre os dados obtidos. Esses tópicos serão tratados, respectivamente, nas Seções 5.1 e 5.2 a seguir.

5.1 Experimentos

Conforme mencionado anteriormente, um dos objetivos deste trabalho envolve o cálculo das invariantes para todos os isômeros com n entre 20 e 130, e todos os isômeros IPR com n entre 132 e 160. Portanto, os experimentos em cada invariante eram feitos nessa ordem, seguindo sequencialmente os valores de n , e geralmente o fim da sequência era atingido em não mais do que algumas horas ou poucos dias (mesmo com a quantidade de grafos a processar).

Apenas em três invariantes o tempo de execução se tornou tão longo que não foi possível cumprir o objetivo planejado: no número de independência, no número de Fries, e no número de Taylor. Uma estratégia que se tentou ao longo dos últimos meses para mitigar o problema foi uma alteração do código que permitisse execuções paralelas do programa para valores diferentes de n e, embora isso tenha surtido algum efeito, não foi capaz de evitar esse desfecho.

Em particular, o número de independência foi calculado para os valores de n entre 20 e 148, o número de Fries para os valores entre 20 e 124 e entre 132 e 142, e o número de Taylor para os valores entre 20 e 130. O caso do número de independência é particularmente digno de nota, pois foram feitas não menos do que três tentativas de calculá-lo nos 6 arquivos restantes. Em cada uma delas, após um período de *semanas* de execução nos computadores da universidade, os processos correspondentes “morriam” sem deixar nenhum traço, nenhuma indicativa de erro que pudesse ser a causa. A suspeita é que a execução tenha sido interrompida por uma reinicialização rotineira da máquina, uma queda de energia, ou a violação de algum limite na conta de usuário.

Com a dupla intenção de não atrasar a escrita deste capítulo e de não desperdiçar o

que já havia sido calculado para as outras invariantes, tomou-se a decisão de fazer a análise considerando esses três casos como exceções. A partir daqui, todos os trechos onde apareceria alguma das informações referente aos fulerenos não contabilizados serão apropriadamente sinalizados como incompletos por esse motivo.

Um último ponto importante a destacar antes de prosseguir é que os experimentos foram executados em diversas máquinas diferentes. Parte deles em um Macbook Pro mencionado anteriormente – com processador Intel Core 2 Duo 2,4GHz, 4GB de memória RAM e sistema operacional Mac OS X Mountain Lion – e parte em três computadores do Departamento de Informática da UFPR reservados para uso do programa de pós-graduação, todos com sistema operacional Linux Mint 17 Qiana. Como essas últimas máquinas eram compartilhadas e acessadas via *SSH*, em cada experimento foram usados os comandos *nohup* e *nice*, e os resultados foram gravados para conferência posterior.

5.2 Resultados

A saída gerada por todos os experimentos foi um conjunto de aproximadamente 500 arquivos *.csv*, ocupando mais de 4GB em disco. Há um arquivo para cada invariante e cada valor de n (com exceção do diâmetro e do índice de Wiener, que são calculados juntos e gravados nos mesmos arquivos). Dentro dos arquivos há uma linha para cada isômero, contendo o identificador daquele isômero e o valor da invariante, separados por vírgula.

De posse de todos esses dados, restava definir uma maneira apropriada de analisá-los e de comparar as invariantes entre si, algo que permitisse a extração de conclusões válidas e úteis. Nesse sentido, foram escolhidas três abordagens distintas: a comparação com valores da literatura (discutida na Subseção 5.2.1), a verificação do desempenho de isômeros conhecida-mente estáveis (alvo da Subseção 5.2.2), e a análise estatística (Subseção 5.2.3). Comparar os valores calculados neste trabalho com os calculados em outros fornece uma medida da correteza dos programas escritos e dá certo grau de confiança aos valores “inéditos” (os que não seja possível comparar por não terem ainda sido calculados em outros trabalhos ou por

não terem sido encontrados).

Por outro lado, em (FAJTOLOWICZ e LARSON, 2003) há uma lista de isômeros conhecidos estáveis (isto é, observados experimentalmente em laboratório) e que, embora limitada, pode ser usada para dar uma ideia da eficácia de cada invariante. Se os isômeros da lista estão entre os poucos a atingir os valores máximo ou mínimo (ou algo muito próximo) para determinada invariante, isso pode ser considerado um indício de que a invariante em questão tem um bom potencial para prever a estabilidade de fulerenos.

Já a análise estatística dos valores calculados, independente da literatura, permite medir o quanto as diferentes invariantes “concordam” ou “discordam” entre si. Em particular, o método escolhido para a análise foi o *coeficiente de correlação de postos de Spearman* (HOGG, MCKEAN e CRAIG), calculado para cada par de invariantes e que leva em consideração a *ordem* dos dados ao invés de seu valor intrínseco.

5.2.1 Comparação com Valores da Literatura

Com o intuito de fornecer uma maneira independente de verificar a validade dos valores calculados neste trabalho, a seguir são listadas algumas referências que contêm dados sobre a maioria das invariantes que foram implementadas (apenas no caso do diâmetro não foi possível encontrar nenhuma referência útil). É importante mencionar que, devido à própria natureza dos valores *double* e das bibliotecas utilizadas em certos casos, o resultado de algumas computações podem estar sujeitos a pequenos erros de arredondamento, que serão ignorados nesta comparação (i.e., será considerado o valor arredondado).

No caso do critério de Fowler-Manolopoulos, as Tabelas 4.3, 4.4, e 4.5 do capítulo 4 de (FOWLER e MANOLOPOULOS, 2006) mostram o valor da invariante para os isômeros IPR com 76, 78, e 84 átomos. Todos os valores exibidos conferem com os calculados aqui. Já no caso do número de Independência, a Tabela 3 de (FAJTOLOWICZ e LARSON, 2003) mostra os valores máximo e mínimo da invariante para os valores de n correspondentes aos isômeros conhecidos estáveis (veja mais abaixo), além do número de independência desses isômeros em particular. Novamente, todos os valores conferem.

Para a frustração bipartida de arestas, a Tabela 1 de (DOŠLIĆ e VUKČEVIĆ, 2007) mostra o valor máximo de $\varphi(G)$ entre todos os isômeros IPR com n na faixa de 94 a 120 ou com n igual a 140 ou 180. Embora neste trabalho a faixa considerada seja diferente (e leve em conta todos os isômeros, não apenas os IPR), os valores em comum também conferem.

Além disso, o texto faz duas outras afirmações, igualmente corroboradas pelos experimentos deste trabalho. A primeira é de que o valor máximo de $\varphi(G)$ para $n = 40$ é atingido apenas por 2 isômeros. A segunda é que todos os isômeros IPR com n entre 60 e 92 possuem $\varphi(G) = 12$. Note que os valores 62, 64, 66, e 68 (com $\varphi(G)$ respectivamente igual a 10, 11, 11, e 10) são uma exceção pois não possuem isômeros IPR (BRINKMAN et al., 2013). Logo, o comentário em (DOŠLIĆ e VUKČEVIĆ, 2007) não se aplica a eles.

Os valores do número de estruturas de Kekulé, do número de Fries, do número de Taylor (e do número mínimo de arestas do emparelhamento ao longo de faces pentagonais) calculados neste trabalho estão todos de acordo com a Tabela 4 de (AUSTIN et al., 1994). Além disso, o valor mínimo (8), o número de isômeros a atingir o mínimo, (2), o valor máximo (20), o número de isômeros a atingir o máximo (apenas um, o $C_{60} : 1812$), e a média (11, 53) para $n = 60$ estão todos de acordo com o exposto na Tabela 5 desse mesmo artigo. Um outro artigo (TORRENS, 2002) (Tabela 1) também apresenta uma seleção de números de estruturas de Kekulé, para n entre 20 e 60, que mais uma vez estão de acordo com os calculados aqui. Por fim, todos os valores do índice de Wiener apresentados em (FOWLER, 2002) para os isômeros de C_{40} conferem com os deste trabalho.

5.2.2 Isômeros Estáveis

Em (FAJTOLOWICZ e LARSON, 2003), a lista de isômeros estáveis é enumerada tomando como base as espirais canônicas apenas dos fulerenos IPR. Fazendo a conversão para a notação mais geral usada aqui (que considera todos os isômeros), a lista é composta pelos isômeros $C_{60} : 1812$, $C_{70} : 8149$, $C_{76} : 19150$, $C_{78} : 24105$, $C_{78} : 24106$, $C_{78} : 24107$, $C_{84} : 51590$, $C_{84} : 51591$. As Tabelas 5.1 a 5.8, que podem ser vistas abaixo, documentam, para cada invariante, a situação de cada um desses isômeros especiais.

Cada tabela contém cinco colunas: *Isômero*, *Máx.?*, *Mín.?*, *% acima*, *% abaixo*. Elas servem para informar, respectivamente, o código do isômero, se ele atinge o valor máximo (em caso positivo, há um número entre parênteses indicando quantos outros isômeros atingem o máximo), se atinge o mínimo (novamente com um número em caso positivo), a porcentagem de isômeros (para um mesmo n) com valor estritamente superior ao dele, e a porcentagem de isômeros com valor estritamente inferior. Perceba que, nos casos em que houver vários isômeros com o mesmo valor que os estáveis, a soma das duas porcentagens pode não chegar nem perto de 100 (já que os valores repetidos não contam nem como estritamente acima nem como estritamente abaixo).

Ao interpretar o conteúdo das tabelas, um sinal de bom desempenho é a capacidade da invariante isolar cada um desses isômeros especiais da maneira mais precisa possível. Nesse sentido, há dois cenários “ideais”: todos os isômeros da tabela estão entre os poucos a atingir o valor máximo, ou todos estão entre os poucos a atingir o valor mínimo. No primeiro caso, a segunda coluna seria toda preenchida com “Sim” (e o valor entre parênteses seria o menor possível), a terceira toda com “Não”, a quarta coluna seria o mais próximo possível de zero, e a quinta coluna o mais próximo possível de 100. No segundo caso, esses valores seriam invertidos.

Embora a expectativa (com base na literatura) fosse de que o critério de Fowler-Manolopoulos se destacasse como a invariante mais precisa, e embora seu desempenho seja razoável, o que se observa é que as Tabelas 5.4 (frustração bipartida de arestas) e 5.5 (número de independência) ilustram claramente os dois cenários ideais. Ao menos nessa faixa limitada, as duas invariantes parecem ser capazes de filtrar mais de 99% dos isômeros candidatos.

Um outro detalhe interessante nesses dois casos é que, conforme o valor de n cresce, percebe-se um ligeiro aumento no número de isômeros a atingir o valor máximo (ou mínimo). Para evitar que isso se torne um problema em valores muito elevados de n , uma boa estratégia parece ser a adotada por (DOŠLIĆ e VUKČEVIĆ, 2007): aplicar a invariante apenas aos isômeros IPR, que por si só já são uma fração do número total de isômeros.

No outro extremo do espectro, o diâmetro parece ser uma das invariantes menos úteis para

prever a estabilidade de fulerenos. Como os isômeros são muito similares, os diâmetros dos grafos correspondentes acabam se repetindo muitas e muitas vezes, o que significa que, mesmo quando um isômero estável atinge o mínimo, o número de outros isômeros que também atinge o mínimo é tão alto que desfaz qualquer vantagem de ter usado a invariante em primeiro lugar. Já o índice de Wiener, que começa de forma bastante promissora (filtrando precisamente vários dos isômeros estáveis), tem sua eficácia diminuída drasticamente conforme n passa de 78. Se esse comportamento se confirmar para valores ainda maiores, a invariante vai acabar sendo tão pouco útil quanto o diâmetro.

Quanto ao número de estruturas de Kekulé, embora de certa forma se saia melhor que o diâmetro (filtrando uma parcela significativa do total de isômeros), a invariante parece ser facilmente substituível pelo número de Fries ou pelo número de Taylor, que apresentam desempenho claramente superior (e quase igual, com uma pequena vantagem para o número de Taylor). Na verdade, e talvez surpreendentemente, o número de Fries e o número de Taylor se saem tão bem ou melhor do que o critério de Fowler-Manolopoulos. Resta descobrir se essa vantagem se mantém para valores maiores de n .

5.2.3 Análise Estatística

Depois de apresentar evidências da validade dos valores computados neste trabalho, e de discutir a eficácia de cada invariante isoladamente, resta mostrar se existe alguma relação entre as invariantes, isto é, se elas concordam entre si. Uma das maneiras de quantificar o relacionamento entre duas variáveis na escala ordinal é através do coeficiente de correlação de Spearman, um método especialmente adequado nos casos em que a relação entre as variáveis não é linear.

Por convenção, o coeficiente de Spearman varia de -1 (maior correlação negativa) até 1 (maior correlação positiva), com zero representando a correlação nula, ou a ausência de correlação entre as variáveis. Neste trabalho, uma correlação positiva indica que as duas invariantes “ordenam” os isômeros aproximadamente da mesma forma, enquanto uma correlação negativa indica que os isômeros são ordenados de forma mais ou menos invertida nas duas invariantes.

Tabela 5.1: Critério de Fowler-Manolopoulos – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Não	Sim(1)	99,94	0,00
C_{70} : 8149	Não	Não	99,21	0,64
C_{76} : 19150	Não	Não	95,97	4,03
C_{78} : 24105	Não	Não	88,63	11,36
C_{78} : 24106	Não	Não	94,48	5,31
C_{78} : 24107	Não	Não	99,87	0,07
C_{84} : 51590	Não	Sim(3)	99,99	0,00
C_{84} : 51591	Não	Sim(3)	99,99	0,00

Tabela 5.2: Diâmetro – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Não	Sim(1261)	30,41	0,00
C_{70} : 8149	Não	Não	26,20	0,02
C_{76} : 19150	Não	Sim(4845)	74,70	0,00
C_{78} : 24105	Não	Sim(2619)	89,14	0,00
C_{78} : 24106	Não	Não	8,51	10,86
C_{78} : 24107	Não	Não	8,51	10,86
C_{84} : 51590	Não	Não	33,59	0,02
C_{84} : 51591	Não	Não	33,59	0,02

Tabela 5.3: Índice de Wiener – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Não	Sim(1)	99,94	0,00
C_{70} : 8149	Não	Sim(1)	99,99	0,00
C_{76} : 19150	Não	Sim(3)	99,98	0,00
C_{78} : 24105	Não	Sim(1)	99,99	0,00
C_{78} : 24106	Não	Não	99,78	0,18
C_{78} : 24107	Não	Não	87,45	11,17
C_{84} : 51590	Não	Não	37,88	60,55
C_{84} : 51591	Não	Não	41,16	57,11

Tabela 5.4: Frustração Bipartida de Arestas – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Sim(1)	Não	0,00	99,94
C_{70} : 8149	Sim(1)	Não	0,00	99,99
C_{76} : 19150	Sim(2)	Não	0,00	99,99
C_{78} : 24105	Sim(5)	Não	0,00	99,98
C_{78} : 24106	Sim(5)	Não	0,00	99,98
C_{78} : 24107	Sim(5)	Não	0,00	99,98
C_{84} : 51590	Sim(33)	Não	0,00	99,94
C_{84} : 51591	Sim(33)	Não	0,00	99,94

Tabela 5.5: Número de Independência – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Não	Sim(1)	99,94	0,00
C_{70} : 8149	Não	Sim(1)	99,99	0,00
C_{76} : 19150	Não	Sim(1)	99,99	0,00
C_{78} : 24105	Não	Sim(3)	99,99	0,00
C_{78} : 24106	Não	Sim(3)	99,99	0,00
C_{78} : 24107	Não	Não	95,30	0,01
C_{84} : 51590	Não	Sim(17)	99,97	0,00
C_{84} : 51591	Não	Sim(17)	99,97	0,00

Tabela 5.6: Número de Estruturas de Kekulé – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Não	Não	1,10	98,84
C_{70} : 8149	Não	Não	3,77	96,22
C_{76} : 19150	Não	Não	9,69	90,31
C_{78} : 24105	Não	Não	13,80	86,20
C_{78} : 24106	Não	Não	4,46	95,54
C_{78} : 24107	Não	Não	30,39	69,60
C_{84} : 51590	Não	Não	28,08	71,92
C_{84} : 51591	Não	Não	29,99	70,01

Tabela 5.7: Número de Fries – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Sim(1)	Não	0,00	99,94
C_{70} : 8149	Sim(8)	Não	0,00	99,90
C_{76} : 19150	Não	Não	0,08	97,52
C_{78} : 24105	Não	Não	0,53	96,50
C_{78} : 24106	Não	Não	0,00	99,97
C_{78} : 24107	Não	Não	0,03	99,47
C_{84} : 51590	Não	Não	0,53	97,02
C_{84} : 51591	Não	Não	0,03	99,47

Tabela 5.8: Número de Taylor – isômeros estáveis.

Isômero	Máx.?	Mín.?	% acima	% abaixo
C_{60} : 1812	Sim(1)	Não	0,00	99,94
C_{70} : 8149	Sim(8)	Não	0,00	99,90
C_{76} : 19150	Não	Não	0,07	99,12
C_{78} : 24105	Não	Não	0,22	99,11
C_{78} : 24106	Não	Não	0,00	99,98
C_{78} : 24107	Não	Não	0,02	99,81
C_{84} : 51590	Não	Não	1,34	94,98
C_{84} : 51591	Não	Não	0,02	99,85

Quanto mais próximo de 1 (ou de -1), mais as duas invariantes *concordam* entre si. Quanto mais próximo de 0, mais as invariantes *discordam* entre si. Note que correlações negativas são esperadas, visto que em algumas invariantes o objetivo é minimizar valores, enquanto em outras o objetivo é maximizar. O que importa, mais do que o sinal, é que a correlação entre as duas invariantes seja distante de zero.

Nas páginas a seguir estão dispostos 28 gráficos, nas Figuras 5.1 a 5.28, um para cada par de invariantes. Em cada gráfico, o eixo X corresponde ao valor de n e está limitado ao intervalo $[20, 160]$, enquanto o eixo Y corresponde ao coeficiente de Spearman. Os pontos, obviamente, correspondem aos coeficientes entre as duas invariantes, calculados para cada n par no intervalo com exceção dos casos especiais mencionados na Seção 5.1 (número de Independência, número de Fries, e número de Taylor, cujas figuras abaixo estão sinalizadas com um asterisco *), dos valores 20, 24, e 26 (nos quais há apenas um isômero) e do 22 (já que não há grafos de fulereno com 22 vértices).

Em primeiro lugar, é interessante observar a relação do critério de Fowler-Manolopoulos, apontado pela literatura como a invariante mais tradicional e promissora, com as demais invariantes. Em cada um dos grafos nas Figuras 5.1 até 5.7 há um padrão semelhante, caracterizado por uma correlação mediana até valores de n entre 80 e 100, seguido por alguns picos e vales em alternância. Essa correlação é particularmente próxima de zero nos casos dos números de Fries e de Taylor, enquanto nos casos do diâmetro, do índice de Wiener, e do número de independência o valor fica bem próximo de 0,5, indicando um bom grau de concordância.

Também é interessante notar que, em pelo menos cinco desses grafos (Figuras 5.1, 5.3, 5.4, 5.6, 5.7), os pontos mais baixos e mais altos parecem se repetir nos mesmos valores de n (100, 120, 140, 160). Embora as observações não sejam suficientes para constatar que existe uma alternância nos múltiplos de 20 acima de 100 (especialmente nos números de independência, de Fries, e de Taylor), há fortes indícios de uma tendência. Se isso se confirmasse, poderia ser um indicativo de características particularmente favoráveis à estabilidade nos isômeros múltiplos de 20 – algo capaz de ser captado igualmente por todas as invariantes.

Outro caso que merece atenção é o da frustração bipartida de arestas e do número de independência, que, ao menos de acordo com o potencial para filtrar os isômeros conhecidos estáveis, parecem ser as invariantes mais promissoras. Na Figura 5.19 é possível ver que a correlação entre elas não é particularmente forte para valores pequenos de n . Porém, conforme n se aproxima e passa de 100, a situação muda e a correlação dá um salto para 1. Essa concordância quase total parece se manter no mínimo até $n = 130$, quando os pontos voltam a descer. Talvez com mais observações nessa faixa fosse possível ver de maneira mais clara alguns picos e vales como no critério de Fowler-Manolopoulos.

O diâmetro, por outro lado, mesmo sendo a invariante menos relevante na prática, mostra uma correlação positiva com a maioria das outras invariantes. Em pelo menos três casos (todos na faixa acima de 100) essa correlação é significativa: nas Figuras 5.9 e 5.10 (frustração bipartida de arestas e número de independência), e na Figura 5.11 (número de estruturas de Kekulé, embora a correlação seja negativa). Na verdade, com exceção do índice de Wiener (que, no geral, apresenta correlações medianas ou fracas) e, parcialmente, do critério de Fowler-Manolopoulos, todas as invariantes apresentam forte concordância nessa faixa que vai aproximadamente de $n = 100$ até $n = 130$, com coeficientes muito próximos de 1 ou -1 .

Por fim, a Figura 5.28 ilustra o caso dos número de Fries e de Taylor, no qual a correlação é consistentemente positiva. Algo que, obviamente, era esperado, visto que o número de Taylor nada mais é do que uma modificação do número de Fries. Essa semelhança é reforçada pelo fato de todas as outras invariantes se relacionarem quase exatamente da mesma forma com elas duas, como pode ser visto nas Figuras 5.6, 5.7, 5.12, 5.13, 5.17, 5.18, 5.21, 5.22, 5.24, 5.25, 5.26, e 5.27.

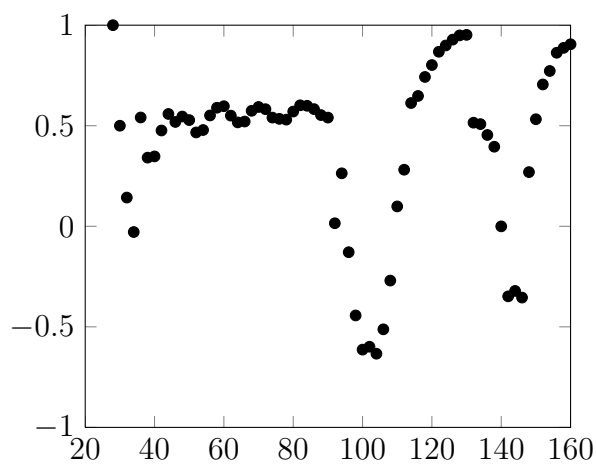


Figura 5.1: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Diâmetro.

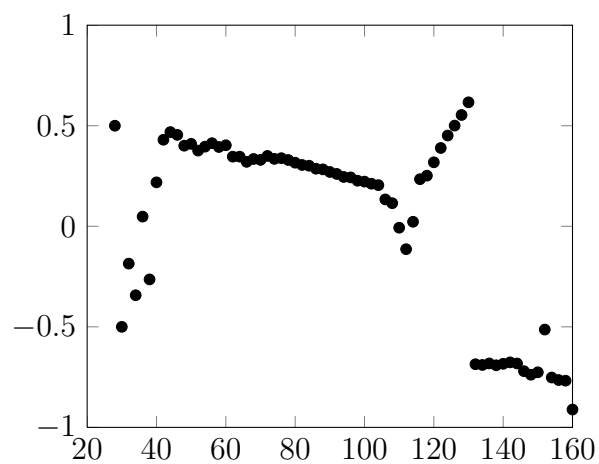


Figura 5.2: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Índice de Wiener.

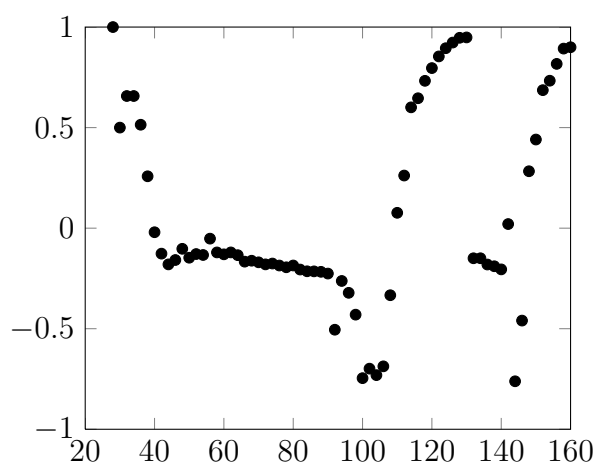


Figura 5.3: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e a Frustração Bipartida de Arestas.

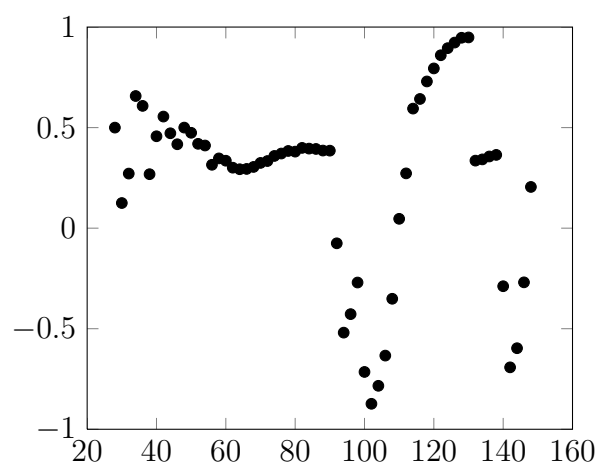


Figura 5.4: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Independência*.

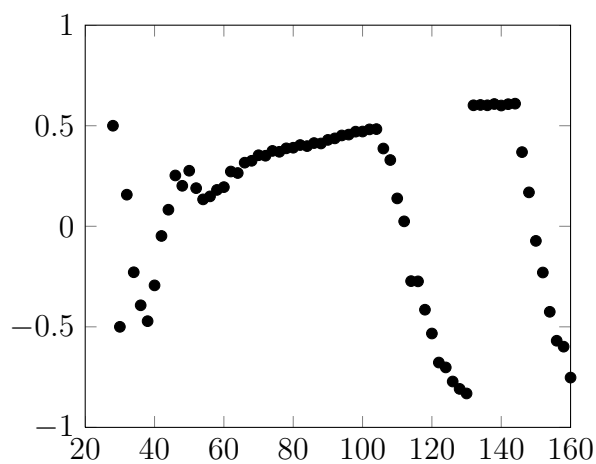


Figura 5.5: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Estruturas de Kekulé.

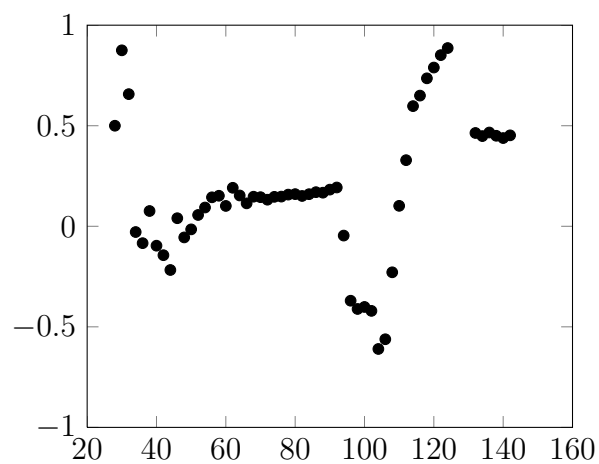


Figura 5.6: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Fries*.

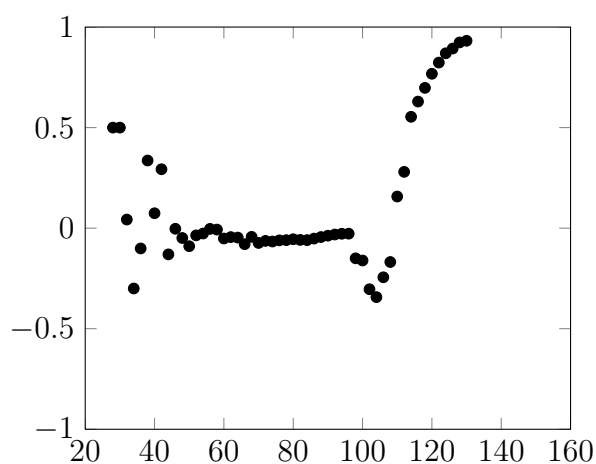


Figura 5.7: Coeficiente de Spearman entre o Critério de Fowler-Manolopoulos e o Número de Taylor*.

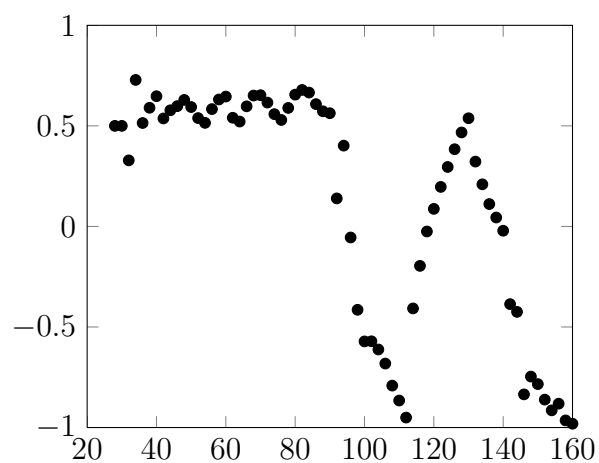


Figura 5.8: Coeficiente de Spearman entre o Diâmetro e o Índice de Wiener.

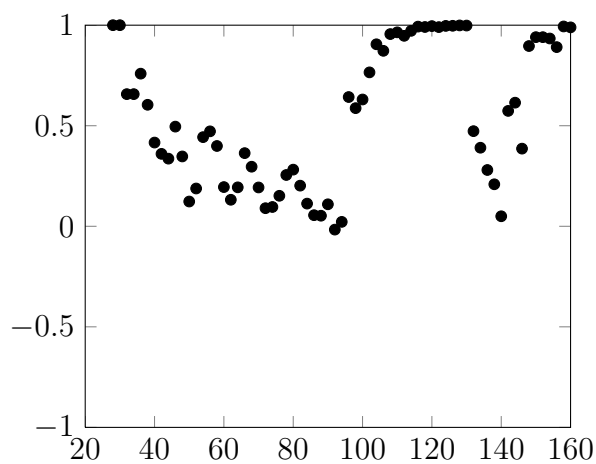


Figura 5.9: Coeficiente de Spearman entre o Diâmetro e a Frustração Bipartida de Arestas.

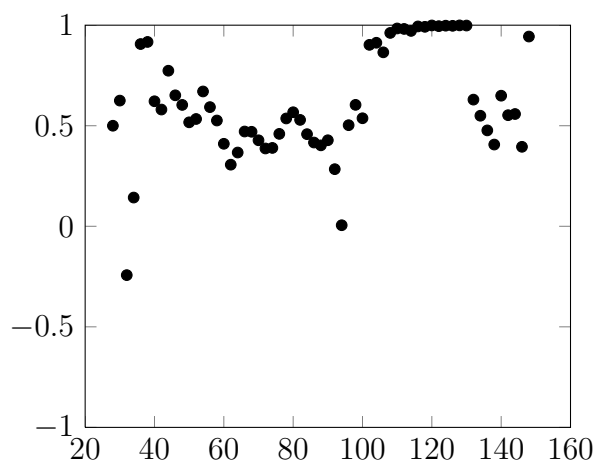


Figura 5.10: Coeficiente de Spearman entre o Diâmetro e o Número de Independência*.

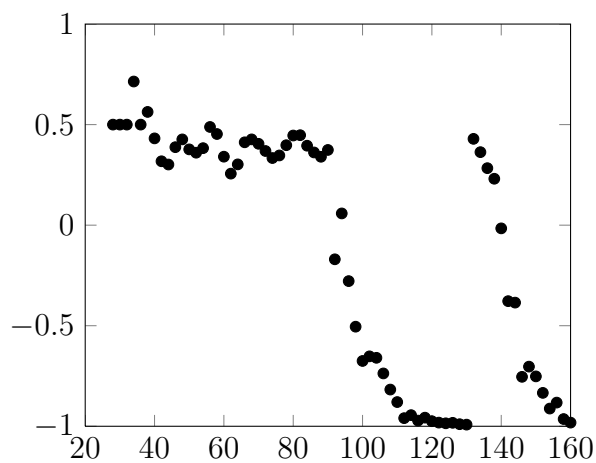


Figura 5.11: Coeficiente de Spearman entre o Diâmetro e o Número de Estruturas de Kekulé.

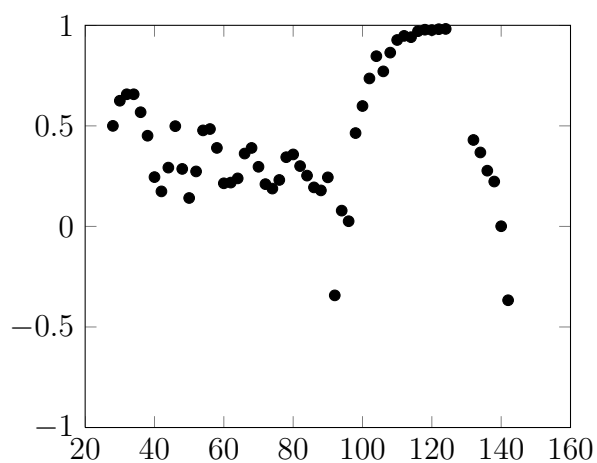


Figura 5.12: Coeficiente de Spearman entre o Diâmetro e o Número de Fries*.

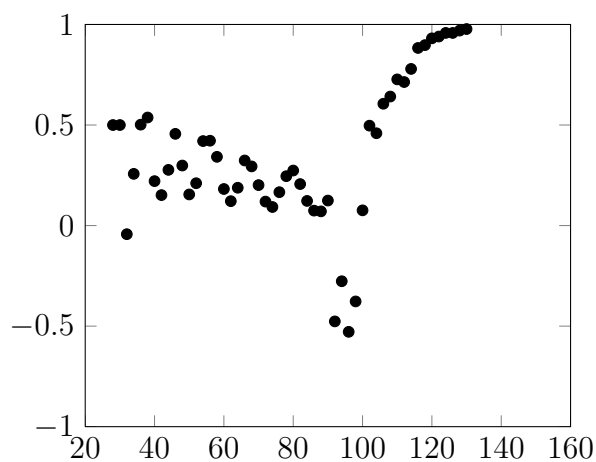


Figura 5.13: Coeficiente de Spearman entre o Diâmetro e o Número de Taylor.*

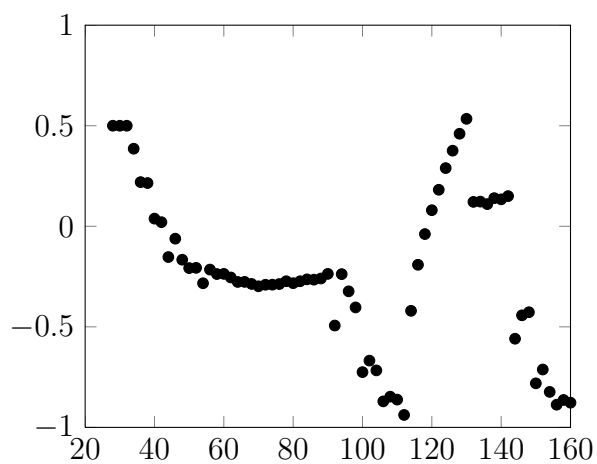


Figura 5.14: Coeficiente de Spearman entre o Índice de Wiener e a Frustração Bipartida de Arestas.

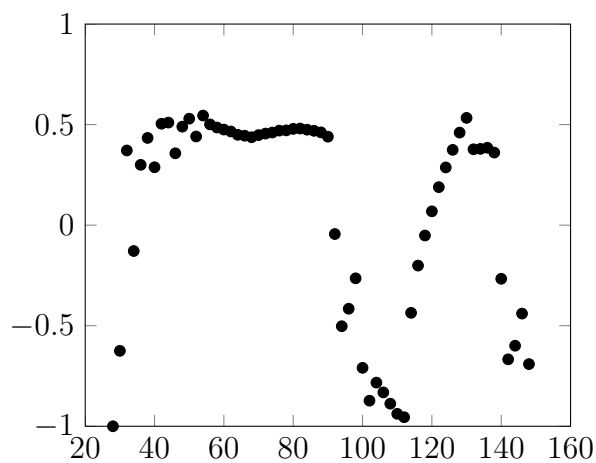


Figura 5.15: Coeficiente de Spearman entre o Índice de Wiener e o Número de Independência*.

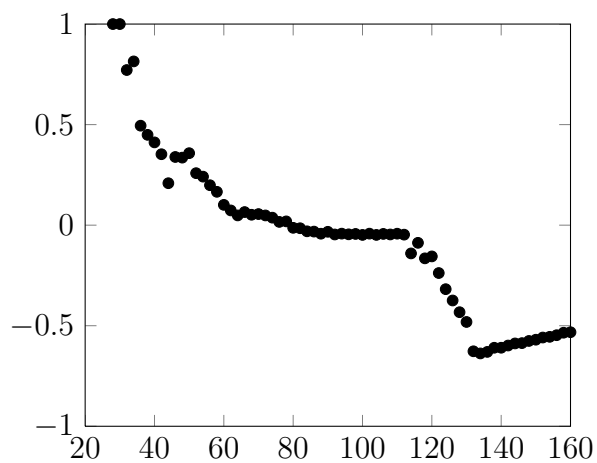


Figura 5.16: Coeficiente de Spearman entre o Índice de Wiener e o Número de Estruturas de Kekulé.

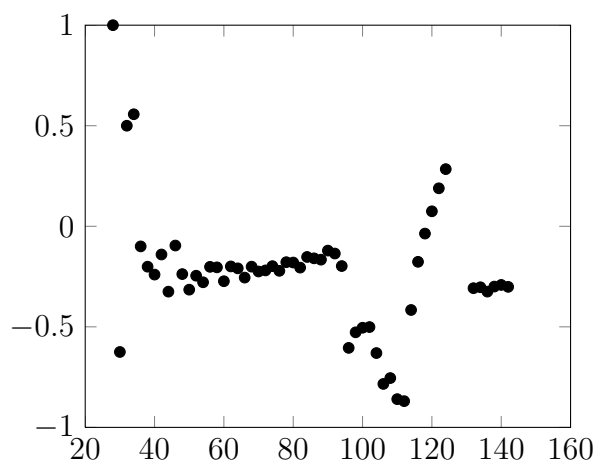


Figura 5.17: Coeficiente de Spearman entre o Índice de Wiener e o Número de Fries*.

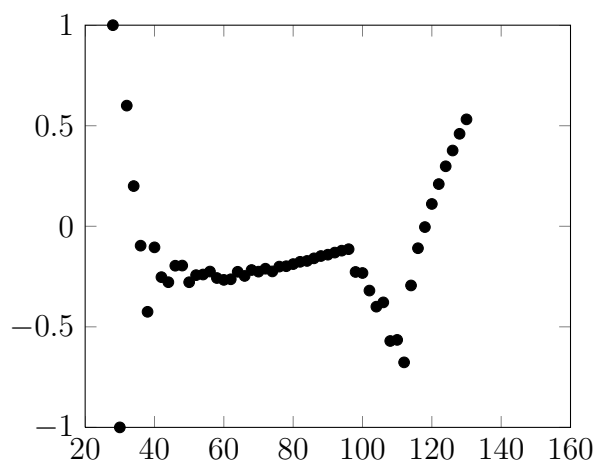


Figura 5.18: Coeficiente de Spearman entre o Índice de Wiener e o Número de Taylor*.

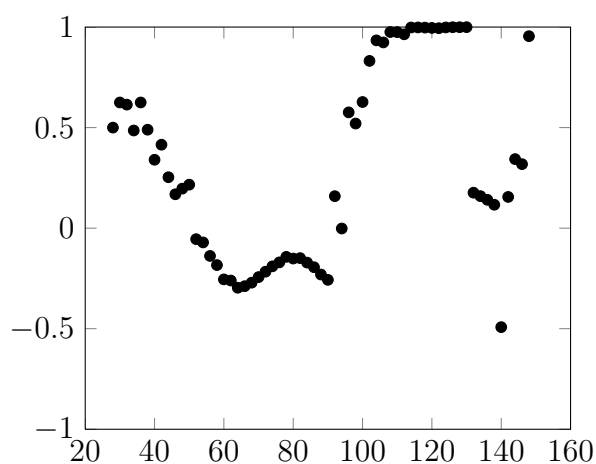


Figura 5.19: Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Independência*.

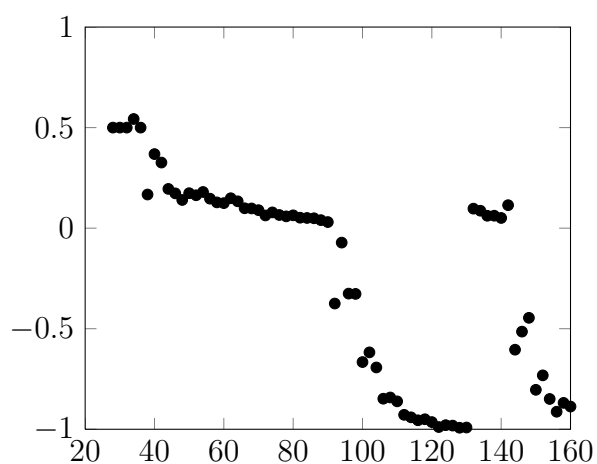


Figura 5.20: Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Estruturas de Kekulé.

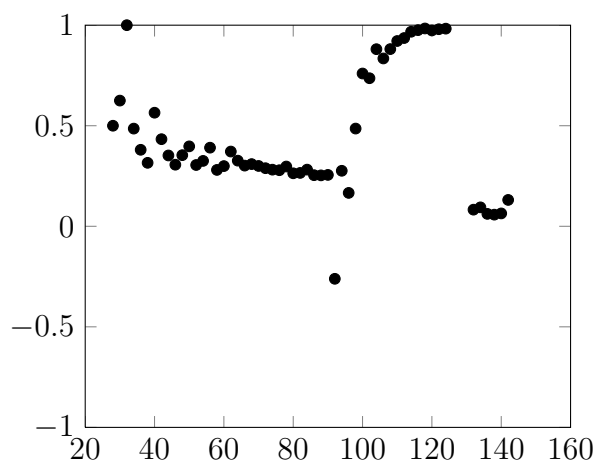


Figura 5.21: Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Fries*.

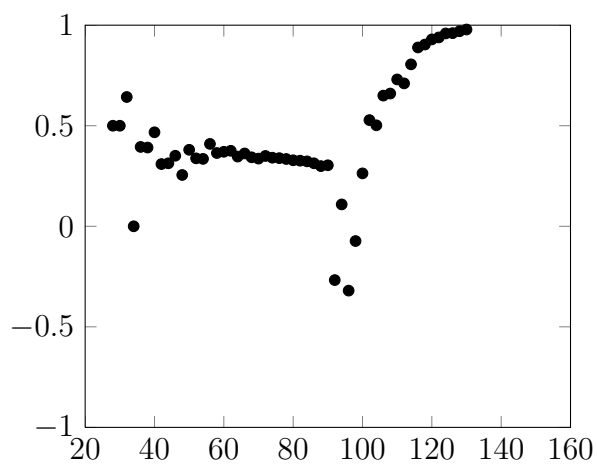


Figura 5.22: Coeficiente de Spearman entre a Frustração Bipartida de Arestas e o Número de Taylor*.

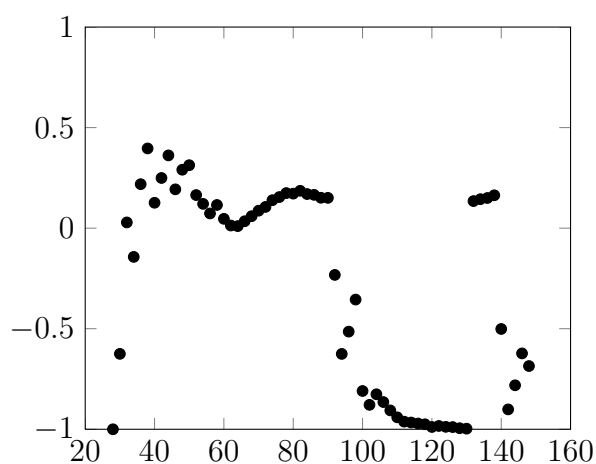


Figura 5.23: Coeficiente de Spearman entre o Número de Independência* e o Número de Estruturas de Kekulé.

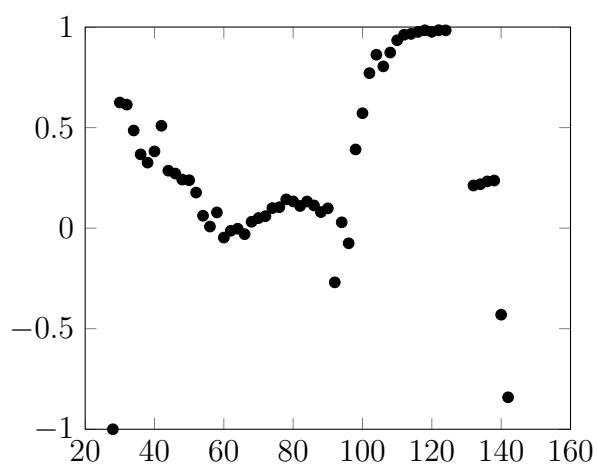


Figura 5.24: Coeficiente de Spearman entre o Número de Independência e o Número de Fries*.

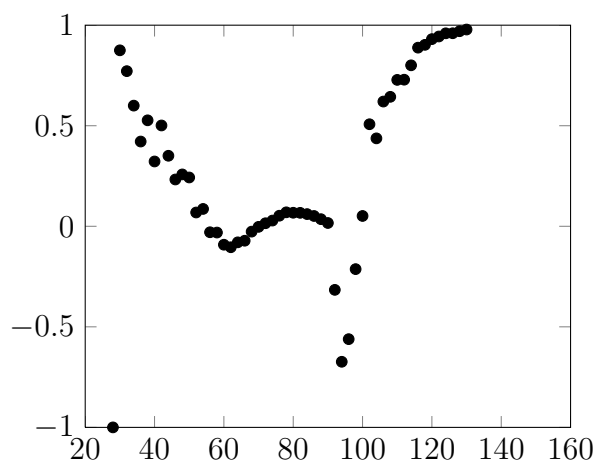


Figura 5.25: Coeficiente de Spearman entre o Número de Independência e o Número de Taylor*.

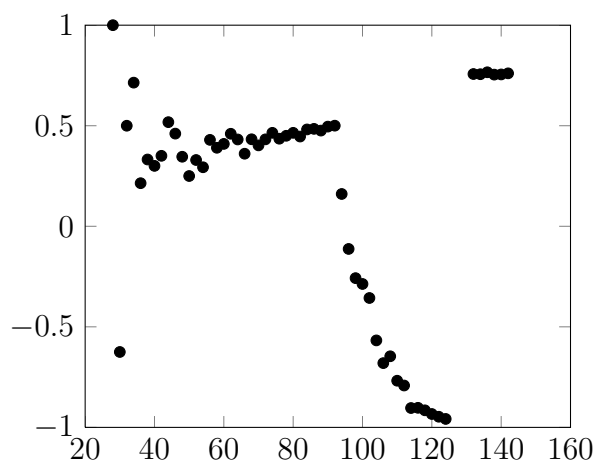


Figura 5.26: Coeficiente de Spearman entre o Número de Estruturas de Kekulé e o Número de Fries*.

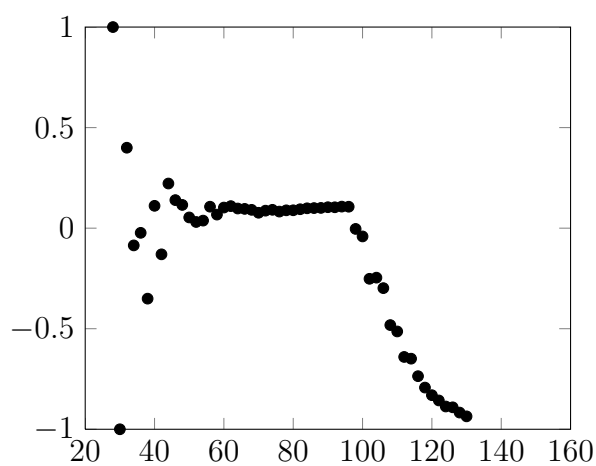


Figura 5.27: Coeficiente de Spearman entre o Número de Estruturas de Kekulé e o Número de Taylor*.

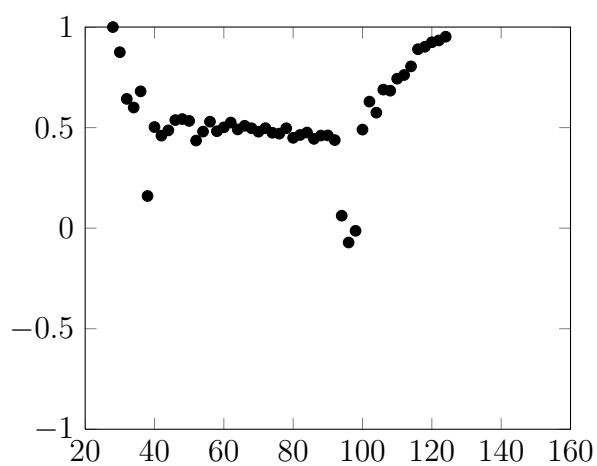


Figura 5.28: Coeficiente de Spearman entre o Número de Fries e o Número de Taylor*.

6 CONCLUSÃO

Este último capítulo encerra a presente dissertação de mestrado. A Seção 6.1 resume brevemente os pontos mais importantes do trabalho. A Seção 6.2 apresenta algumas dificuldades encontradas durante o desenrolar do projeto. Por fim, a Seção 6.3 discute possíveis continuacões deste trabalho.

6.1 Sobre o trabalho desenvolvido

Este texto apresentou uma dissertação de mestrado cujo objetivo foi a comparação de diversas invariantes da teoria dos grafos na previsão da estabilidade de moléculas de fulereno, que podem ser representadas através dos grafos de fulereno definidos na Seção 2.1. As invariantes em questão foram selecionadas com base em sua relevância na literatura sobre o tema, seu suposto potencial de previsão esperado, e sua dificuldade de implementação.

Em particular, foram analisados, nesta ordem: o critério de Fowler-Manolopoulos, o diâmetro, o índice de Wiener, a frustração bipartida de arestas, o número de independência, o número de emparelhamentos perfeitos, o número de Fries, e o número de Taylor. Para cada uma dessas invariantes, este autor estudou o material disponível na literatura, implementou algoritmos para calculá-las, e obteve o valor das invariantes para cada isômero de fulereno com até 130 vértices e para cada isômero IPR com 132 até 160 vértices (com exceção do número de independência, o número de Fries, e o número de Taylor, para os quais apenas parte dessa faixa foi cumprida).

Uma vez de posse dos dados sobre o valor das invariantes em cada isômero, o autor pode fazer uma análise dos resultados, comparando o desempenho de cada invariante com o que era realmente esperado, além de ter comparado a eficácia relativa das invariantes entre si (comparando-as duas a duas utilizando uma métrica estatística conhecida como coeficiente de Spearman). Embora a maioria das invariantes mencionadas no texto já tenha sido estudada no contexto de fulerenos, o principal mérito deste trabalho é o fato de ainda não existir um estudo comparativo que leve todas elas em conta.

Foram aplicadas três abordagens para avaliar e extrair conclusões úteis dos dados obtidos. A primeira foi a comparação com valores calculados em outros trabalhos, que serviu como uma verificação independente da correteza dos programas escritos, dando um grau de confiança maior aos dados. A segunda foi a verificação do desempenho das invariantes na previsão da estabilidade de isômeros que já se sabe terem sido previamente observados em laboratório. Nesse quesito duas invariantes se destacaram claramente acima das outras: a frustração bipartida de arestas e o número de independência. Em particular, esta última merece atenção especial, pois mesmo existindo indicações na literatura de que o número de independência não possui uma relação muito forte com a estabilidade, ele foi capaz de isolar os isômeros conhecidamente estáveis com alta precisão – um indício de que o número de independência closed-shell pode ser ainda melhor.

Finalmente, a terceira abordagem consistiu em uma análise estatística dos valores de cada invariante, quantificando o quanto elas concordam ou discordam entre si através do coeficiente de correlação de postos de Spearman. Aqui é importante destacar que esse método foi escolhido por parecer ao autor o mais apropriado durante a execução do projeto, dada a natureza das variáveis (no sentido estatístico). Entretanto, na prática (isto é, após os cálculos), ele pode não ter sido o melhor método, visto que ficou difícil extrair conclusões muito claras a partir dos valores obtidos.

O que se esperava aqui era que as invariantes “concordassem” na *separação* entre isômeros “bons” (estáveis) e “ruins” (instáveis). Como essa separação envolve pouquíssimos isômeros estáveis e um número (potencialmente) exorbitante de isômeros instáveis, fica difícil averiguar o impacto dos isômeros instáveis no valor do coeficiente de Spearman para a invariante e , conseqüentemente, nos gráficos mostrados anteriormente. Talvez diminuindo o “ruído” causado pela presença de tantos isômeros instáveis (removendo alguma parcela deles da análise) fosse possível enxergar tendências mais claras nos gráficos.

Por exemplo, duas das observações feitas anteriormente foram: i) quase todas as invariantes apresentaram forte concordância para os valores de $n = 100$ até $n = 130$, e ii) o critério de Fowler-Manolopoulos, que apresenta inicialmente uma correlação mediana com ou-

tras invariantes na faixa de n entre 80 e 100, tem um comportamento singular para valores de n maiores que 100, com uma sequência de altos e baixos que parece se repetir sempre nos mesmos múltiplos de 20. Será que isso é mesmo pelas invariantes concordarem na faixa (100, 130) e existir uma sequência de picos e vales, ou será devido à proporção entre isômeros estáveis e instáveis ser extremamente desigual (para o lado dos instáveis) conforme o valor de n cresce? Algumas maneiras de remediar o problema desse possível “ruído” no coeficiente de Spearman são mencionadas na Seção 6.3 abaixo e, embora não sejam incluídas neste trabalho, estão sendo consideradas num artigo escrito dando sequência ao que foi documentado aqui.

6.2 Dificuldades Encontradas

Além de alguns atrasos no cronograma inicialmente planejado, uma das principais dificuldades encontradas por este autor durante a elaboração do trabalho foram os artigos com explicações extremamente curtas sobre as invariantes, por vezes omitindo detalhes importantes para a implementação do que está sendo descrito. Dois casos notáveis foram o número de Fries e o número de Taylor, que teriam sido excluídos da comparação se não fossem os esclarecimentos encontrados em (HANSEN e ZHENG, 1994). Mesmo assim, essas duas invariantes acabaram sendo as mais difíceis de implementar, algo provavelmente agravado pela falta de experiência do autor com programação inteira e com a biblioteca GLPK.

Outra dificuldade está relacionada com a demora dos experimentos para algumas invariantes. Em particular, foi possível completar apenas parcialmente a faixa planejada para o número de independência, o número de Fries, e o número de Taylor. As lacunas permaneceram mesmo após uma alteração no código que permitia execuções paralelas para valores diferentes de n , pois algumas instâncias do programa (no caso do número de independência) chegaram a demorar semanas antes de falhar por algum motivo ainda não determinado.

6.3 Oportunidades Futuras

Finalmente, resta apenas mencionar algumas extensões deste trabalho, que podem acabar motivando trabalhos futuros. Em primeiro lugar, a continuação mais direta seria completar os “espaços em branco” deixados no número de independência, no número de Fries, e no número de Taylor (que poderiam ser calculados até os valores planejados inicialmente, de 20 até 160). Outra ideia é expandir mais ainda os limites da análise – por exemplo, considerando todos os isômeros IPR com n até 200.

Uma terceira alternativa seria a inclusão de outras invariantes, que aqui não puderam ser incluídas nem mesmo na revisão bibliográfica por falta de tempo. Esse é o caso da *distância de resistência* (semelhante ao índice de Wiener, mas se baseia numa analogia elétrica, considerando cada aresta do grafo de fulereno como um fio de resistência unitária (FOWLER, 2002)), do número de árvores geradoras (conceito também conhecido como *complexidade*, que se relaciona de forma inversamente proporcional ao número de adjacências entre pentágonos e, conseqüentemente, de forma proporcional com a estabilidade relativa (FOWLER, 2003)), do descritor topológico Ψ de Réti e László (que se baseia no tradicional índice de pentágonos, acrescentando o conceito novo de *pentagon arm indices* (RÉTI e LÁSZLÓ, 2000)), além de algumas outras invariantes (ALCAMI et al., 2007; GAN et al., 2009; SCHEIN e SANDS-KIDNER, 2008; SLANINA et al., 2001).

Em quarto lugar, mas não menos importante, há a possibilidade de realizar análises estatísticas mais sofisticadas, que permitam extrair conclusões mais claras a respeito da relação entre as invariantes. Três caminhos a considerar são: i) escolher os $X\%$ melhores de acordo com cada invariante (com X , por exemplo, igual a 10 ou 20), fazer a *união* desses conjuntos (para garantir consistência, caso contrário algum isômero poderia não aparecer em alguma invariante e não seria possível calcular o coeficiente de Spearman), e recalcular o coeficiente, ii) da mesma forma que o anterior, escolher os $X\%$ melhores, mas realizar a *interseção* dos conjuntos ao invés da união (que se espera não resulte em um conjunto vazio), e calcular mais uma vez o coeficiente, e iii) ao invés do coeficiente de Spearman, aplicar o *teste de*

Friedman (MILTON, 1937), uma medida especialmente aplicável ao caso em que há várias observações para cada objeto (i.e., diversas invariantes calculadas em cada isômero).

REFERÊNCIAS

- ACHTERBERG, T. Scip: Solving constraint integer programs. **Mathematical Programming Computation**, v. 1, n. 1, p. 1–41, July 2009. <http://mpc.zib.de/index.php/MPC/article/view/4>.
- ALCAMI, M. et al. Structural patterns in fullerenes showing adjacent pentagons: C_{20} to C_{72} . **J. Nanosci. Nanotechnol.**, v. 7, p. 1329–1338, 2007.
- ALON, N. Eigenvalues and expanders. **Combinatorica**, v. 6, n. 2, p. 83–96, 1986.
- ALON, N.; MILMAN, V. D. λ_1 , isoperimetric inequalities for graphs and superconcentrators. **Journal of combinatorial theory, Series B**, v. 38, p. 73–88, 1985.
- ANDOVA, V. et al. On the diameter and some related invariants of Fullerene graphs. **MATCH Communications in Mathematical and in Computer Chemistry**, v. 68, p. 109–130, 2012. ISSN 0340-6253.
- AUSTIN, S. J. et al. Fullerene isomers of C_{60} . Kekulé counts versus stability. **Chemical Physics Letters**, v. 228, p. 478–484, 1994. ISSN 0009-2614.
- BERKELAAR, M.; EIKLAND, K.; NOTEBAERT, P. *Ip_solve 5.5, Open source (Mixed-Integer) Linear Programming system*. 2004. Software. Disponvel em: <http://lpsolve.sourceforge.net/5.5/>. Acesso em: 23/01/2015. Disponível em: <<http://lpsolve.sourceforge.net/5.5/>>.
- BOCHVAR, D. A.; GAL'PERN, E. G. **Proceedings of the USSR Academy of Sciences**, v. 209, p. 239, 1973.
- BONCHEV, D.; TRINAJSTIĆ, N. Information theory, distance matrix, and molecular branching. **Journal of Chemical Physics**, v. 67, p. 4517–4533, 1977.
- BONDY, J. A.; MURTY, U. S. R. **Graph Theory with Applications**. USA: Elsevier Science Ltd. North-Holland, 1976.
- BRESSOUD, D.; PROPP, J. How the alternating sign matrix conjecture was solved. **Not. Amer. Math. Soc.**, v. 46, p. 637–646, 1999.
- BRINKMAN, G. et al. CaGe - a virtual environment for studying some special classes of plane graphs - an update. **Match - Communications in Mathematical and in Computer Chemistry**, v. 63, n. 3, p. 533–552, 2010. ISSN 0340-6253.
- BRINKMAN, G.; GOEDGEBEUR, J.; MCKAY, B. D. Generation of fullerenes. **Journal of Chemical Information Modeling**, v. 52, n. 11, p. 2910–2918, 2012.
- BRINKMAN, G. et al. House of graphs: a database of interesting graphs. **Discrete Applied Mathematics**, v. 161, n. 1-2, p. 311–314, 2013. Disponvel em: <http://hog.grinvin.org>. Acesso em: 13/03/2015.

BRINKMAN, G. et al. Generation of simples quadrangulations of the sphere. **Discrete Mathematics**, v. 305, p. 33–54, 2005.

BRINKMAN, G.; MCKAY, B. D. Construction of planar triangulations with minimum degree 5. **Discrete Mathematics**, v. 301, p. 147–163, 2005.

BRINKMAN, G.; MCKAY, B. D. Fast generation of planar graphs. **Match - Communications in Mathematical and in Computer Chemistry**, v. 58, n. 2, p. 323–357, 2007. ISSN 0340-6253.

CAMI, J. et al. Detection of C_{60} and C_{70} in a young planetary nebula. **Science**, v. 329, n. 5996, p. 1180–1182, 2010.

CHEEGER, J. A lower bound for the smallest eigenvalue of the laplacian. In: GUNNING, R. C. (Ed.). **Problems in Analysis**. New Jersey: Princeton University Press, 1970. p. 195–199.

CORMEN, T. H. et al. **Introduction to Algorithms, Third Edition**. 3rd. ed. [S.l.]: The MIT Press, 2009. ISBN 0262033844, 9780262033848.

DAUGHERTY, S.; MYRVOLD, W.; FOWLER, P. W. Backtracking to compute the closed-shell independence number of a Fullerene. **MATCH Communications in Mathematical and in Computer Chemistry**, v. 58, p. 385–401, 2007. ISSN 0340-6253.

DAUGHERTY, S. M. **Independent Sets and Closed-Shell Independent Sets of Fullerenes**. Tese (Doutorado) — University of Victoria, 2009.

DAVIDOFF, G.; SARNAK, P.; VALETTE, A. **Elementary number theory, group theory, and Ramanujan graphs**. [S.l.]: Cambridge University Press, 2003.

DAVIDSON, R. A. **Theoretica Chimica Acta**, v. 58, p. 193, 1981.

DIESTEL, R. **Graph Theory**. Second. New York, NY, USA: Springer, 2000.

DILLON, A. C. et al. Storage of hydrogen in single-walled carbon nanotubes. **Nature**, v. 386, n. 6623, p. 377–379, 1997.

DOŠLIĆ, T. On some structural properties of fullerene graphs. **J. Math. Chem.**, n. 31, p. 187–195, 2002.

DOŠLIĆ, T. Bipartivity of fullerene graphs and fullerene stability. **Chem. Phys. Lett.**, v. 412, p. 336–340, 2005.

DOŠLIĆ, T. Saturation number of fullerene graphs. **J. Math. Chem.**, n. 43, p. 647–657, 2008.

DOŠLIĆ, T.; VUKČEVIĆ, D. Computing the bipartite edge frustration of fullerene graphs. **Discrete Applied Mathematics**, v. 155, p. 1294–1301, 2007.

DRESSELHAUS, M. S.; AVOURIS, P. Introduction to carbon materials research. In: DRESSELHAUS, M. S.; DRESSELHAUS, G.; AVOURIS, P. (Ed.). **Carbon Nanotubes: Synthesis, Structure, Properties and Applications**. Berlin, Germany: Springer Berlin Heidelberg, 2001, (Topics in Applied Physics, v. 80).

ESTRADA, E.; RODRÍGUEZ-VELÁZQUEZ, J. A. Spectral measures of bipartivity in complex networks. **Physical Review E**, v. 72, n. 4, p. 046105, 2005.

EULER, L. Demonstratio nonnullarum insignium proprietatum, quibus solida hedris planis inclusa sunt praedita. **Novi Commentarii Academiae Scientiarum Petropolitanae**, v. 4, p. 140–160, 1758.

FAJTLOWICZ, S. On conjectures of graffiti. **Discrete Mathematics**, v. 72, p. 113–118, 1988.

FAJTLOWICZ, S. On conjecture of graffiti V. In: _____. **Graph Theory, Combinatorics, and Algorithms: Proceedings of the Seventh Quadrennial Conference on the Theory and Application of Graphs**. [S.l.]: Western Michigan University, 1995. v. 1, p. 367–376.

FAJTOLOWICZ, S.; LARSON, C. E. Graph-theoretic independence as a predictor of fullerene stability. **Chemical Physics Letters**, v. 377, p. 485–490, 2003. ISSN 0009-2614.

FARIA, L.; KLEIN, S.; STEHLÍK, M. Odd cycle transversals and independent seta in fullerene graphs. **SIAM J. Discrete Math.**, v. 26, n. 3, p. 1458–1469, 2012.

FOWLER, P. W. Fullerene graphs with more negative than positive eigenvalues: The exceptions that prove the rule of electron deficiency? **Journal of the Chemical Society, Faraday Transactions**, v. 93, p. 1–3, 1997.

FOWLER, P. W. Resistance distances in fullerene graphs. **Croatica Chemica Acta**, v. 75, n. 2, p. 401–408, 2002. ISSN 0011-1643.

FOWLER, P. W. Complexity, spanning trees and relative energies in fullerene isomers. **Match - Communications in Mathematical and in Computer Chemistry**, v. 48, p. 87–96, 2003. ISSN 0340-6253.

FOWLER, P. W. A note on the smallest eigenvalue of fullerenes. **MATCH Communications in Mathematical and in Computer Chemistry**, v. 48, p. 37–48, 2003. ISSN 0340-6253.

FOWLER, P. W.; DAUGHERTY, S.; MYRVOLD, W. Independence number and fullerene stability. **Chemical Physics Letters**, v. 448, p. 75–82, 2007.

FOWLER, P. W.; JOOYANDEH, M.; BRINKMANN, G. Face-spiral codes in cubic polyhedral graphs with face sizes no larger than 6. **Journal of Mathematical Chemistry**, v. 50, n. 8, p. 2272–2280, 2012. ISSN 0259-9791.

FOWLER, P. W.; MANOLOPOULOS, D. E. **An Atlas of Fullerenes**. Mineola, NY, USA: Dover Publications, Inc., 2006.

FOWLER, P. W.; MYRVOLD, W. Some chemical applications of independent sets. In: HANSEN, P. et al. (Ed.). **Proceedings of the 18th Mini Euro Conference on Variable Neighborhood Search**. Tenerife, Spain: [s.n.], 2005. p. 16–37.

FOWLER, P. W. et al. Facts and conjectures about fullerene graphs: Leapfrog, cylinder and Ramanujan fullerenes. In: BETTEN, A. et al. (Ed.). **Algebraic Combinatorics and Applications**. Berlin: Springer, 2000. p. 134–146.

FOWLER, P. W. et al. Independent sets and the prediction of additional patterns for higher fullerenes. **Journal of the Chemical Society, Perkin Transactions 2**, p. 2023–2027, 1999.

FRIES, K. ' Über bicyclische verbindungen und ihren vergleich mit dem naphtalin. **Justus Liebigs Annalen der Chemie**, v. 454, n. 1, p. 121–324, 1927.

FRIES, K.; WALTER, R.; SCHILLING, K. ' Über tricyclische verbindungen, in denen naphtalin mit einem heterocyclus anelliert ist. **Justus Liebigs Annalen der Chemie**, v. 516, n. 1, p. 248–285, 1935.

GAN, L. H. et al. General geometric rule for stability of carbon polyhedra. **Chemical Physics Letters**, v. 472, p. 224–227, 2009.

GAREY, M. R.; JOHSON, D. S. **Computers and Intractability; A Guide to the Theory of NP-Completeness**. New York, NY, USA: W. H. Freeman & Co., 1990.

GIBBONS, F. P. et al. Fullerene resist materials for the 32nm node and beyond. **Advanced Functional Materials**, v. 18, p. 1977–1982, 2008.

GODSIL, C.; ROYLE, G. **Algebraic Graph Theory**. New York: Springer, 2001.

GOHO, A. Allergy nanomedicine: Buckyballs dampen response of cells that trigger allergic reactions. **Science News**, v. 172, n. 5, p. 5, 2007.

GOLDBERG, M. The isoperimetric problem for polyhedra. **Tohoku Math. J.**, v. 40, p. 226–236, 1935.

GOLDBERG, M. A class of multisymmetric polyhedra. **Tohoku Math. J.**, v. 43, p. 104–108, 1937.

GRÜNBAUM, B.; MOTZKIN, T. S. The number of hexagons and the simplicity of geodesics on certain polyhedra. **Can. J. Math.**, v. 15, p. 744–751, 1963.

GUEDES, A. L. P. **Representao de Variedades Combinatrias de Dimenso n**. Dissertaço (Mestrado) — COPPE/Engenharia de Sistemas e Computao - UFRJ, Rio de janeiro, RJ, 1995.

GUENNEBAUD, G.; AL., B. J. et. *Eigen v3*. 2010. Disponvel em: <http://eigen.tuxfamily.org>. Acesso em: 13/05/2015.

GUIBAS, L.; STOLFI, J. Primitives for the manipulation of general subdivisions and the computation of voronoi diagrams. **ACM Transactions on Graphics**, v. 4, n. 2, p. 74–123, 1985.

GUTMAN, I. et al. Some recent results in the theory of the Wiener number. **Indian Journal of Chemistry**, v. 32A, p. 651–661, 1993.

HANSEN, P.; ZHENG, M. Numerical bounds for the perfect matching vectors of a polyhex. **Journal of Chemical Information and Computer Sciences**, v. 34, n. 2, p. 305–308, 1994.

HECKMAN, C. C.; THOMAS, R. Independent sets in triangle-free cubic planar graphs. **Journal of Combinatorial Theory, Series B**, v. 96, p. 253–275, 2006.

HERNDON, W. C. Resonance energies of aromatic hydrocarbons. quantitative test of resonance theory. **Journal of the American Chemical Society**, v. 95, n. 7, p. 2404–2406, 1973.

HERNDON, W. C.; ELLZEY, M. L. Resonance theory. v. resonance energies of benzonoid and nonbenzonoid .pi. systems. **Journal of the American Chemical Society**, v. 96, n. 21, p. 6631–6642, 1974.

HOGG, R. V.; MCKEAN, J.; CRAIG, A. T. **Introduction to Mathematical Statistics**. 7th. ed. [S.l.]: Pearson, 2012. ISBN 0321795431, 978-0321795434.

HOLME, P. et al. Network bipartivity. **Physical review. E, Statistical, nonlinear, and soft matter physics**, v. 68, n. 5, p. 056107, 2003.

HOSOYA, H. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. **Bull. Chem. Soc. Japan**, v. 44, p. 2332–2339, 1971.

JONES, D. E. H. **New Scientist**, v. 35, p. 245, 1966.

JUSTUS, C. **Boundaries of triangle-patches and the expander constant of fullerenes**. Tese (Doutorado) — Universität Bielefeld, October 2007.

KASTELEYN, P. The statistics of dimers on a lattice: I. the number of dimer arrangements on a quadratic lattice. **Physica**, v. 27, n. 12, p. 1209 – 1225, 1961. ISSN 0031-8914. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0031891461900635>>.

KASTELEYN, P. W. Dimer statistics and phase transitions. **Journal of Mathematical Physics**, v. 4, n. 2, p. 287–293, 1963.

KASTELEYN, P. W. Graph theory and crystal physics. In: HARARY, F. (Ed.). **Graph Theory and Theoretical Physics**. New York: Academic Press, 1967. p. 43–110.

KOLMOGOROV, V. Blossom v: a new implementation of a minimum cost perfect matching algorithm. **Mathematical Programming Computation**, Springer-Verlag, v. 1, n. 1, p. 43–67, 2009. ISSN 1867-2949. Disponível em: <<http://dx.doi.org/10.1007/s12532-009-0002-8>>.

KONC, J.; JANEŽIČ, D. An improved branch and bound algorithm for the maximum clique problem. **Match - Communications in Mathematical and in Computer Chemistry**, v. 58, p. 569–590, 2007. ISSN 0340-6253.

KOSHLAND JR., D. E. Molecule of the year. **Science**, v. 254, p. 1705, 1991.

KRÄTSCHMER, W. et al. Solid C_{60} : a new form of carbon. **Nature**, v. 347, p. 354–358, 1990.

KROTO, H. W. The stability of the fullerenes C_n , with $n = 24, 28, 32, 36, 50, 60$ and 70. **Nature**, v. 329, p. 529–531, 1987.

KROTO, H. W. et al. C_{60} : Buckminsterfullerene. **Nature**, v. 318, p. 162–163, 1985.

LI, Q. et al. Antimicrobial nanomaterials for water disinfection and microbial control: Potential applications and implications. **Water Research**, v. 42, n. 18, p. 4591–4602, 2008. ISSN 0043-1354. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0043135408003333>>.

LIERS, F.; PARDELLA, G. **A simple MAX-CUT algorithm for planar graphs**. [S.l.], 2008. Disponível em: <Disponvel em: <http://e-archive.informatik.uni-koeln.de/579/>. Acesso em: 13/05/2015.>.

LOVÁSZ, L.; PLUMMER, M. D. **Matching Theory**. In: . Amsterdam: North-Holland, 1986, (Annals of Discrete Mathematics, v. 29).

LUBOTZKY, A.; PAK, I. The product replacement algorithm and kazhdan's property (t). **Journal of the American Mathematical Society**, v. 14, n. 2, p. 347–363, 2000.

LUBOTZKY, A.; PHILLIPS, R.; SARNAK, P. Ramanujan graphs. **Combinatorica**, v. 8, n. 3, p. 261–277, 1988.

MILTON, F. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. **Journal of the American Statistical Association**, v. 32, n. 200, p. 675–701, 1937.

MOHAR, B. Isoperimetric numbers of graphs. **Journal of combinatorial theory, Series B**, v. 47, p. 274–291, 1989.

MOSSMAN, D. et al. Testing for fullerenes in geologic materials: Oklo carbonaceous substances, karelian shungites, sudbury black tuff. **Geology**, v. 31, n. 3, p. 255–258, 2003.

NAOR, J.; NAOR, M. Small-bias probability spaces: Efficient constructions and applications. **SIAM Journal on Computing**, v. 22, p. 838–856, 1993.

NEWMAN, A. Max cut. In: KAO, M.-Y. (Ed.). **Encyclopedia of Algorithms**. Springer US, 2008. p. 489–492. ISBN 978-0-387-30770-1. Disponível em: <http://dx.doi.org/10.1007/978-0-387-30162-4_219>.

NONENMACHER, R. A. *Graph of a 60-fullerene (a truncated icosahedral graph) using one of the pentagons as a base (perimeter)*. 2008. Disponvel em: https://commons.wikimedia.org/wiki/File:Graph_of_60-fullerene_w-nodes.svg. Acesso em: 13/03/2015. Wikimedia Commons.

OSAWA, E. **Kagaku (Kyoto)**, v. 25, p. 854, 1970.

PAPADIMITRIOU, C. H.; STEIGLITZ, K. **Combinatorial Optimization**. In: . Englewood Clifss, NJ, USA: Prentice-Hall, 1982, (Algorithms and Complexity).

PIPPENGER, N. Sorting and selecting in rounds. **SIAM Journal on Computing**, v. 16, p. 1032–1038, 1987.

RAGHAVACHARI, K. Ground state of C_{84} : Two almost isoenergetic isomers. **Chem. Phys. Lett.**, v. 190, p. 397–400, 1992.

RANDIĆ, M. Comment on the difference between the bond orders calculated by SCF MO and simple MO method. **The Journal of Chemical Physics**, v. 34, n. 2, p. 693–694, 1961.

RANDIĆ, M. Conjugated circuits and resonance energies of benzonoid hydrocarbons. **Chemical Physics Letters**, v. 38, n. 1, p. 68–70, 1976. ISSN 0009-2614.

RÉTI, T.; LÁSZLÓ, I. On the combinatorial characterization of fullerene graphs. **Acta Polytechnica Hungarica**, v. 6, n. 5, p. 85–93, 2000. ISSN 1785-8860.

ROUVRAY, D. H. The modeling of chemical phenomena using topological indices. **Journal of Computational Chemistry**, v. 8, n. 4, p. 470–480, 1987.

ROUVRAY, D. H. The rich legacy of half a century of the Wiener index. In: ROUVRAY, D. H.; KING, R. B. (Ed.). **Topology in Chemistry: Discrete Mathematics of Molecules**. [S.l.]: Horwood Publishing, 2002. p. 16–37.

SAH, C.-H. A generalized leapfrog for fullerene structures. **Fullerene Science and Technology**, v. 2, n. 4, p. 445–458, 1994.

SARNAK, P. What is an expander? **Notices of the AMS**, v. 51, p. 762–763, 2004.

SCHEIN, S.; SANDS-KIDNER, M. A geometric principle may guide self assembly of fullerene cages from Clathrin Triskella and from carbon atoms. **Biophysical Journal**, v. 94, p. 958–976, 2008.

SCHMALZ, T. G. et al. C_{60} carbon cages. **Chem. Phys. Letters**, v. 130, n. 3, p. 203–207, 1986.

SCHMALZ, T. G. et al. Elemental carbon cages. **J. Am. Chem. Soc.**, v. 110, p. 1113–1127, 1988.

SIPSER, M. Expanders, randomness, or time versus space. **Journal of Computer System Sciences**, v. 36, p. 379–383, 1988.

SIPSER, M.; SPIELMAN, D. A. Expander codes. **IEEE Transactions on Information Theory**, v. 42, p. 1710–1722, 1996.

SLANINA, Z. et al. Geometrical and thermodynamic approaches to the relative stabilities of fullerene isomers. **Communications in mathematical and in computer chemistry**, v. 44, p. 335–348, 2001.

STRÖCK, M. *A 3D model of a C_{60} molecule, also called a "Buckyball"*. 2006. Disponível em: <http://commons.wikimedia.org/wiki/File:C60a.png>. Acesso em: 13/03/2015. Wikimedia Commons.

STRÖCK, M. *This illustration depicts eight of the allotropes (different molecular configurations) that pure carbon can take*. 2006. Disponível em: http://commons.wikimedia.org/wiki/File:Eight_Allotropes_of_Carbon.png. Acesso em: 13/03/2015. Wikimedia Commons.

STREITWIESER JR., A. **Molecular Orbital Theory for Organic Chemists**. New York: John Wiley & Sons, 1961.

TANS, S. J. et al. Individual single-wall carbon nanotubes as quantum wires. **Nature**, v. 386, p. 474–477, 1997.

TAYLOR, R. Rationalisations of the most stable isomer of a fullerene *C*. **Journal of the Chemical Society, Perkin Transactions 2**, n. 1, p. 3–4, 1992.

TAYLOR, R. (Ed.). **The Chemistry of Fullerenes**. 5 Toh Tuck Link, Singapore: World Scientific Pub. Co. Pte. Ltd., 1995. v. 4.

TEGOS, G. P. et al. Cationic Fullerenes are effective and selective antimicrobial photosensitizers. **Chemistry & biology**, v. 12, n. 10, p. 1127–1135, 2005.

TEMPERLEY, H. N. V.; FISHER, M. E. Dimer problem in statistical mechanics-an exact result. **Philosophical Magazine**, v. 6, n. 68, p. 1061–1063, 1961. Disponível em: <<http://dx.doi.org/10.1080/14786436108243366>>.

THOMPSON, B. C.; FRÉCHET, J. M. J. Polymer-Fullerene composite solar cells. **Angewandte Chemie International Edition**, v. 47, p. 58–77, 2008.

TORRENS, F. Computing the permanent of the adjacency matrix for fullerenes. **Internet Electron. J. Mol. Des.**, v. 1, n. 7, p. 351–359, 2002. <http://www.biochempress.com>.

VAZIRANI, V. V. {NC} algorithms for computing the number of perfect matchings in $k_{3,3}$ -free graphs and related problems. **Information and Computation**, v. 80, n. 2, p. 152 – 164, 1989. ISSN 0890-5401. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0890540189900175>>.

WEST, D. B. **Introduction to Graph Theory**. Patparganj, Delhi, India: Pearson Education, 2002.

WIENER, H. Structural determination of paraffin boiling points. **Journal of the American Chemical Society**, v. 69, n. 1, p. 17–20, 1947.

ZEEMAN, E. C. **Uma introdução informal à topologia das superfícies**. Rio de Janeiro: Publicações do IMPA, 1975.

ZHAO, Y. et al. Hydrogen storage in novel organometallic Buckyballs. **Physical Review Letters**, American Physical Society, v. 94, n. 15, p. 155504, 2005. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.94.155504>>.