مجله‌ی ریاضی شریف

سال سوم شماره‌ی هفتم

مجله‌ی ریاضی شریف
سال سوم شماره‌ی هفتم

صاحب امتیاز: انجمن علمی و فوق‌برنامه‌ی دانشکده‌ی علوم ریاضی؛ مدیر مسوول: دکتر امیر جعفری؛ سردبیر: خشایار فیلم؛ همکاران این شماره: دکتر امیر جعفری، عباس محرابیان، روزبه فرهودی، عرفان صلواتی، آرش حدادان، حمید ملک، محی‌الدین متوصل، امید رفیعیان، سامان حبیبی اصفهانی، یاسمن بقایی، ابوالفضل طاهری؛ هیات تحریریه: دکتر امیر جعفری، علی کمالی‌نژاد، خشایار فیلم، علی قصاب، روزبه فرهودی، عرفان صلواتی، نوید هاشمی، اوژن غنی‌زاده‌ی خوب، احمدرضا حاج‌سعیدی؛ طراحی: اوژن غنی‌زاده‌ی خوب؛ طراحی سایت: محسن منصوریار؛ ویراستاری: خشایار فیلم، ابوالفضل طاهری، یاسمن بقایی، حمید ملک؛ با تشکر از دکتر مصطفی اصفهانی‌زاده، دکتر سیدرضا مقدسی، محمدامین فضلی، حمید پرنیان، محسن یاورتنو، حمید ملک، محمد امین شعبانی، مهتاب کریمی، علی خزلی، محمد قیاسی، ساینا غفرانی

فهرست مطالب

عمرخیام، رنه دکارت و حل معادلات جبری ۱
توپولوژی کاربردی و دانته: مصاحبه‌ای با رابرت کریس ۱۰
گذری بر فلسفه و آموزش ریاضی ۱۴
تاریخچه‌ی تقابل مربعی ۱۸
چند حقیقت درباره‌ی گروه‌ها ۳۰
کمی دورتر از فضای اقلیدسی ۳۳
بعد در ریاضیات ۵۳
نگاهی هندسی به مساله‌ای حدی در احتمال ۶۱
الگوریتم جستجوی گراور ۶۶
بهینه‌سازی ترکیبیاتی روی گراف‌های با عرض درختی محدود ۷۱
مساله‌ها ۷۷
پاسخ مساله‌ها ۷۸
فعالیت‌های درون دانشکده ۸۴

بسمه تعالی

آن‌چه که در دست دارید هفت‌امین شماره از دوره‌ی جدید مجله‌ی ریاضی شریف است. آن‌چه که در این هفت شماره –با احتساب دو شماره‌ی اول که تنها روی وب قرار گرفت– از زمستان ۹۰ تاکنون انجام گردید؛ تلاشی بود در جهت احیای مجله‌ای که زمانی بسیاری از اساتید کنونی دانشکده در آن قلم می‌زده‌اند و با این وجود متأسفانه مدت‌ها از زمان انتشار آخرین شماره‌ی آن می‌گذشت. سخن گفتن درباره‌ی این تجربه برای من دو جنبه دارد. اول شرحی مختصر از آن‌چه که در این مدت نه چندان کوتاه فعالیت مجله گذشت و دوم این‌که این تجربه برای خود من به چه معنی بوده است. با مورد اول شروع می‌کنم.

بی‌هیچ تردیدی کسی که نخستین‌بار و البته بیش از هر شخص دیگری در جهت شروع دوره‌ی جدیدی از مجله‌ی ریاضی همت گمارد، ابوالفضل طاهری بود. من و ابوالفضل هردو ورودی ۸۵ بودیم و اگر حافظه یاری کند پاییز سال ۹۰ بود که دریافتم ابوالفضل در پی احیای مجدد مجله‌ی ریاضی شریف است. شروع کار با انتشار دو شماره‌ی اینترنتی بود که ابوالفضل و دوستان دیگری همچون اوژن غنی‌زاده زحمت آن را تماماً تقبل کردند و من همکاری مختصری در حد نوشتن دو مقاله‌ی کوتاه داشتم. قدم بعدی گرفتن مجوز برای انتشار نسخه‌ی چاپی مجله بود. دکتر جعفری لطف کردند و پذیرفتند مدیرمسؤول دوره‌ی جدید مجله باشند و ابوالفضل که به دلایلی کاملاً غیرعلمی و غیرمنطقی نمی‌توانست سردبیر مجله باشد این مسؤولیت را به من پیشنهاد کرد و قرعه‌ی این کار به من افتاد! هرچند کماکان ابوالفضل عمده‌ی بار زحمت را به دوش می‌کشید و اوژن همواره گزینه‌ی اول ما برای طراحی جلد بود. در این پنج شماره‌ای که از زمستان ۹۱ تا این شماره‌ای که اکنون در دست شماست چاپ گردید؛ هم از کمک دانشجویان کارشناسی دانشکده استفاده کردیم و هم از دوستانی که در مقطع دکترا مشغول به تحصیل بودند. چرا که هدف ما آن بود که مجله تا آن‌جا که مقدور است از همکاری دانشجویان بهره گیرد و فعالیت‌های انجام شده در دانشکده را منعکس کند به این امید که با دخیل کردن افرادی بیشتر شاهد وقفه‌ای مجدد در انتشار مجله نباشیم. تصور می‌کنم حداقل در دو شماره‌ی اخیر تا حدی به این هدف نائل آمده‌ایم. شماره‌ی پیشین که زمستان گذشته منتشر گردید گزارش مفصلی درباره‌ی همایش مرزهای دانش ریاضی – که بی‌تردید

برگزاری آن یکی از مهم‌ترین رخدادهای علمی دانشکده‌مان در سال گذشته بود- دربرداشت و آماده‌سازی این شماره نیز در یک بازه‌ی زمانی به نسبت کوتاه بدون همکاری همه‌جانبه‌ی دوستانمان در انجمن علمی و فوق‌برنامه که انجام بسیاری از کارها را تقبل کردند به هیچ‌وجه میسر نمی‌شد. ضروری است این‌جا از همه‌ی این دوستان تشکر کنم و امید است این روحیه‌ی کار جمعی انجمن را قادر سازد روند انتشار مجله را ادامه دهد.

برای خود من که تا پیش از این کار جمعی به این صورت را نیازموده بودم؛ همکاری با مجله‌ی ریاضی شریف تجربه‌ای بی‌بدیل بود که در جریان آن بسیاری از مطالب و مهارت‌ها را فراگرفتم و مهمتر آن‌که از جهت دخیل بودن در یک فعالیت داوطلبانه‌ی علمی و تلاش در جهت انجام آن به نحو احسن برایم بسیار آموزنده بود. شاید اگر به عقب بازگردیم باید از کسی که مانع شد ابوالفضل سردبیر شود تشکر کنم که عدو شود سبب خیر اگر خدا خواهد!

و حرف آخر این‌که تصور می‌کنم گروهی که در این هفت شماره مجله را منتشر کرده، آن‌چه که داشته عرضه کرده و وقت آن رسیده است که مسؤولیت سردبیری و انتشار مجله به گروهی دیگر از دوستان منتقل شود. همکاری صمیمانه‌ای که بچه‌های انجمن علمی و فوق‌برنامه در جریان آماده‌سازی این شماره داشتند مرا دلگرم می‌سازد که انجمن به عنوان صاحب‌امتیاز مجله‌ی ریاضی شریف این سنت را بی‌آن‌که وقفه‌ی طولانی دیگری ایجاد شود ادامه خواهد داد. در انتها باید تشکر کنم از همه‌ی اساتید و دانشجویانی که در این مدت ما را یاری کردند. فهرست اسامی‌شان قطعاً طولانی‌تر از آن است که این‌جا بگنجد.

یک چند به خیره عمر بگذشت   من بعد بر آن سرم که چندی
بنشینم و صبر پیش گیرم
دنباله‌ی کار خویش گیرم[1]

خدانگهدار
خشایار فیلُم

---
[1] سعدی

عمر خیام، رنه دکارت و حل معادلات جبری

ترجمه: یاسمن بقایی

نویسنده‌ی این مقاله پروفسور ماردیا [۱] در جشن هزاره‌ی خیام در نوامبر سال ۱۹۹۹ از طرف باشگاه خیام در لندن برای سخنرانی دعوت شد. قسمتی از این مقاله ازمتن سخنرانی او گرفته شده است. باشگاه خیام در سیزدهم اکتبر سال ۱۸۹۲ در لندن تِأسیس شد و همچنان به فعالیت خودش پیرامون آثار خیام ادامه می‌دهد.

۱ چکیده

خیام تنها فردی است که علاوه بر ریاضی‌دانی بزرگ از او به عنوان شاعری توانمند نیز یاد می‌شود. در این مقاله می‌خواهیم تلاش‌های خیام در قرن ۱۲ میلادی پیرامون حل معادلات جبری را توضیح دهیم. تلاش‌های ارزشمندی که رنه دکارت [۲] را نیز در قرن ۱۷ تحت تاثیر خود قرار می‌دهد. مشخصاً در این مقاله به حل معادلات درجه‌ی سه می‌پردازیم؛ معادلاتی که قرن ۹ تا ۱۶ دنیای ریاضیات را مجذوب خود کرده بود .

خیام سهم عمده‌ای در پیدا کردن ریشه‌های مثبت معادله با استفاده از استدلال هندسی دارد. امری که باعث می‌شود او سرانجام بتواند هندسه‌ی تحلیلی دکارت را تا حدی پیش‌بینی کند. ابتدا تاریخچه و طبقه‌بندی معادلات درجه سوم را توسط خیام توضیح می‌دهیم و سپس روش او برای حل آن را شرح خواهیم داد.

۲ مقدمه

برای ریاضی‌دانی برجسته بسیار نادر است که اشعار او همانند کارهای ریاضی‌اش به یاد آورده شود اما خیام تنها فردی است که همزمان کارهای ریاضی‌اش به خوبی اشعارش به خاطر آورده می‌شود. یک ریاضی‌دان بزرگ در کنار شاعری برجسته. افرادی همچون جورج بول[۳]و ماکسول[۴] هم شعر می‌نوشتند اما برای آن به شهرت نرسیدند یا لانگ فلو[۵] شاعر آمریکایی که تنها یک ریاضی‌دان آماتور بود.

در این‌جا به توضیح فعالیت خیام در قرن ۱۲ پیرامون حل معادلات جبری و به طور مشخص حل معادلات درجه سوم می‌پردازیم. همچون دکارت که متهم به انکار کردن خدا شد، به نظر می‌رسد خیام هم زندگی آسوده‌ای نداشته و به خاطر اعتقاداتش تحت فشار بوده است. (Kasir,1931,p.3)

۳ پیشینه

در این‌جا فقط به حل معادلات درجه‌ی سه توسط خیام می‌پردازیم. ریاضی‌دانان در قرون ۹ تا ۱۶ مجذوب معادلات درجه‌ی سه بودند. در نهایت دو ریاضی‌دان ایتالیایی کاردانو[۶] و تارتالیا[۷]موفق به حل کامل این معادلات در قرن ۱۶ شدند. از طرف دیگر خیام سهم چشم‌گیری در پیدا کردن ریشه(های) مثبت معادله‌ی درجه سوم با استفاده از روش هندسی را دارد در حالی که رابطه‌ی میان جبرو هندسه در قرن ۱۷ توسط دکارت بیان شد. دکارت از ابتدا به حل معادله علاقمند بود. روش‌هایی که در حل معادله به کار برد بیشتر شبیه کار خیام بود ولی او دریافت که بعضی از نقاط تقاطع، ریشه‌ی موهومی را نشان می‌دهد (Katz,1998,p.448). دکارت راه‌حل کاملی برای ریشه‌های منفی، مثبت و موهومی را ارائه داد تا جایی که حتی حل معادلات درجه بالاتر از سه هم در کارهایش قرارگرفت در شرایطی که به گواه تاریخ کارهای خیام در این زمینه بسیار محدود بوده است. (Kasir,1931,p.34)

۴ پیش‌نیازها

ابتدا به خصوص برای افراد غیرریاضی‌دان، به توضیح الفاظ رایجی می‌پردازیم که به شکل روزمره در ریاضی به کار می‌روند:

لغت‌نامه‌ی آکسفورد بیان می‌کند :

Algebra (جبر) = پیدا کردن خواص اعداد و کمیت با استفاده از نمادهای عمومی

و همچنین درباره‌ی ریشه‌ی عربی این واژه می گوید :

Aljebr = پیوند مجدد قسمت‌های شکسته (دوباره به هم پیوستن = jabara)

"جبر " روح ریاضیات مدرن محسوب می‌شود . نام جبر از "الجبر و المقابله " که نام یکی از کارهای خوارزمی است گرفته شده است (حدود ۸۲۵ بعد از میلاد).

[۱] Kantilal Mardia
[۲] René Descartes
[۳] George Boole
[۴] Maxwell
[۵] Longfellow
[۶] Cardano
[۷] Tartaglia

کلمه‌ی بعدی معادله است. دوباره از لغت‌نامه‌ی آکسفورد :
معادله(equation) = فرمولی که بر هم‌ارزی دو عبارت ریاضی، که با علامت "=" به هم ربط داده شده‌اند صحه می‌گذارد.
در ادامه تلاش‌های عمر خیام را براساس دست‌نوشته‌های موجود از او که قبلاً توسط وُپکه[۸] در ۱۸۵۱ و کثیر[۹] در ۱۹۳۱ استفاده شده، توضیح خواهیم داد .

کثیر (1931,p.9)می‌گوید دست‌نوشته‌ای که او روی آن کار کرده است همان دست‌نوشته‌ی موجود در کتابخانه‌ی شخصی پروفسور اسمیت از دانشگاه کلمبیا است که بسیار شبیه به دست‌نوشته‌ی موجود در کتابخانه لیدن[۱۰] است. وُپکه از دست‌نوشته‌ی کتابخانه لیدن برای ترجمه استفاده کرده بود. به گفته‌ی کثیر (p.9) این دست‌نوشته بخش اساسی متن عربی استفاده شده توسط وُپکه است. در ادامه دو صفحه‌ی اول متن عربی مقاله از آن را مشاهده می کنید .

ساده‌ترین نوع معادله، معادله‌ی خطی با یک متغیر که آن را x می‌نامیم است، به عنوان مثال:

$$2x - 3 = 4x + 2$$

که تنها یک جواب $x = -\frac{5}{2}$ دارد. این معادله خطی است زیرا بیشترین درجه‌ی مجهول یک است. معادله‌ی بعدی یک معادله درجه‌ی دو است:

$$ax^2 + bx + c = 0$$

که در آن a ، b و c ثابت هستند. در این‌جا بالاترین درجه‌ی مجهول دو است. این معادله سالها قبل حل شده بود و دو جواب دارد. در واقع جواب‌هایی که امروزه برای این معادله می‌دانیم توسط دو ریاضی‌دان بزرگ هندی ارائه شد. براهماگوپته[۱۱] (قرن۶) که یکی از جواب‌های معادله و بهاسکره[۱۲] (قرن ۱۲) هر دو جواب معادله را بدست آوردند. (Katz, 1998 , pp. 226 – 227)
دو جواب معادله می‌توانند هر دو موهومی یا هر دو حقیقی باشند. به عنوان مثال معادله‌ی

$$x^2 = -1$$

دو ریشه‌ی موهومی دارد (مجذور ریشه‌ها عدد منفی می‌شود!) با گسترش بیشتر این ایده‌ها، معادله‌ی درجه‌ی سه به صورت طبیعی دارای شکل زیر است:

$$ax^3 + bx^2 + cx + d = 0$$

که در آن a ، b ، c و d ثابت هستند. با یک تغییر متغیر خطی ساده $x = y - \frac{b}{3a}$ معادله تبدیل به $y^3 + py + q = 0$ (فرم متعارف) می‌شود که در آن p و q ثابت هستند. این معادله سه جواب دارد که حداقل یکی از آن‌ها حقیقی است .

به پیروی از یونانیان باستان، خیام برای بیان معادله از توان‌های متوالی مجهول که آن‌ها را ریشه، مربع، مکعب می‌نامید استفاده می‌کرد.

به عنوان مثال جمله‌ی: «یک مکعب و مربع برابر است با ریشه‌ها و عددها»، در زبان ریاضی و نشانه‌گذاری امروز به صورت زیر خواهد بود:

$$x^3 + ax^2 = bx + c$$

(مکعب) + a(مربع) = b(ریشه) + عدد

منظور خیام از «حل عددی» اعداد صحیحی بود که درمعادله صدق کنند. از طرف دیگر منظور او از «حل هندسی»، تعیین مجهول به عنوان کمیتی قابل‌اندازه‌گیری به عنوان طول یک پاره‌خط بود. (Kasir,1931,p.23)

۵ رده‌بندی خیام

خیام اولین فردی بود که معادلات را به طور مفصل رده‌بندی کرد هرچند در زبان امروزی کاری که او انجام می‌داده در واقع بررسی درجه‌ی معادله بوده است. اولین دسته از معادلات که شامل دو قسمت هستند - یا به عبارت دیگر معادلات دوجمله‌ای- توسط خود او «معادلات ساده» نام‌گذاری شده‌اند. این معادلات به شرح زیرند:
۱) $a = x$ ۲) $a = x^2$ ۳) $a = x^3$ ۴) $bx = x^2$ ۵) $cx^2 = x^3$ ۶) $bx = x^3$

دسته‌ی دوم که آن‌ها را« معادلات ترکیبی» می‌نامد به دو نوع تقسیم می‌شوند: نوع اول معادلاتی که شامل سه جمله یا سه عبارت هستند و نوع دوم معادلاتی که شامل چهار عبارت هستند.

الف) معادلات سه‌جمله‌ای درجه‌ی دو

۷) $x^2 + bx = a$ ۸) $x^2 + a = bx$ ۹) $bx + a = x^2$

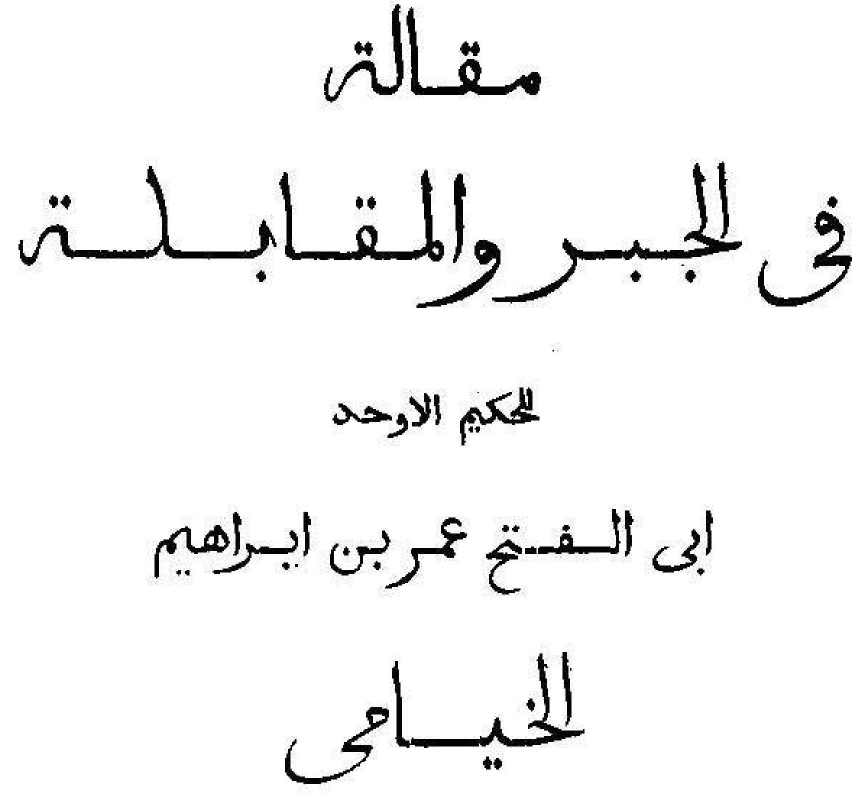
ب) معادلات سه‌جمله‌ای درجه‌ی سه که قابل تقلیل به معادلات درجه‌ی دوم هستند

۱۰) $x^3 + cx^2 = bx$ ۱۱) $x^3 + bx = cx^2$ ۱۲) $cx^2 + bx = x^3$

ج) معادلات سه‌جمله‌ای درجه‌ی سه

۱۳) $x^3 + bx = a$ ۱۴) $x^3 + a = bx$ ۱۵) $bx + a = x^3$

---
[۸] Woepcke
[۹] Kasir
[۱۰] Leyden
[۱۱] Brahmagupta
[۱۲] Bhaskara


مقالة
فی الجبر والمقابلة
للحکیم الاوحد
ابی الفتح عمر بن ابراهیم
الخیامی

شکل ۱:
عنوان کتاب جبر و مقابله‌ی حکیم عمر خیام (دست‌نوشته‌ی کتابخانه‌ی لیدن)

رسالة الحكيم الفاضل

غياث الدين ابى الفتح عمر بن ابراهيم الخيّامى النيشابورى (۱)

قدّس الله روحه العزيز

فى البراهين على مسائل الجبر والمقابلة

بسم الله الرحمن الرحيم

الحمد لله رب العالمين والعاقبة للمتقين ولا عدوان الا على الظالمين والصلوة على الانبيآء وخصوصا على محمد واله الطاهرين اجمعين (۲) ان احد المعانى التعليميّة المحتاج اليها فى جزء الحكمة المعروف بالرياضى هو صناعة الجبر والمقابلة الموضوعة لاستخراج المجهولات العددية

شکل ۲:
صفحه‌ی آغازین کتاب جبر و مقابله‌ی حکیم عمر خیام (دست‌نوشته‌ی کتابخانه‌ی لیدن)

۱۶) $x^3 + cx^2 = a$ ۱۷) $x^3 + a = cx^2$ ۱۸) $cx^2 + a = x^3$

د) معادلات چهارجمله‌ای که در آن‌ها حاصل‌جمع سه جمله برابر با جمله‌ی چهارمی است

۱۹) $x^3 + cx^2 + a = bx$ ۲۰) $x^3 + cx^2 + bx = a$

۲۱) $x^3 + bx + a = cx^2$ ۲۲) $cx^2 + bx + a = x^3$

ح) معادلات چهارجمله‌ای که در آن‌ها جمع دوجمله مساوی با جمع دو جمله‌ی دیگر است

۲۳) $x^3 + cx^2 = bx + a$ ۲۴) $x^3 + bx = cx^2 + a$

۲۵) $x^3 + a = cx^2 + bx$

خیام پیشینه‌ی معادلات را در مقاله‌اش توضیح می‌دهد و این‌که کدام‌یک از معادلات تا آن زمان به طور کامل حل شده‌اند. در ادامه روش خود را برای حل با توجه به این دسته‌بندی ۲۵ گانه شرح می‌دهد. بعضی از این دسته‌ها از نظر جبری یکسانند اما از نظر هندسی با هم تفاوت دارند.

۶ راه‌حل‌های ارائه شده توسط خیام

روش‌های خیام با نمادگذاری مدرن به صورت زیر بیان می‌شود. فرض کنید:

$$x^3 + px = q \quad p, q > 0 \tag{1}$$

فرم متعارف معادله‌ی کلی باشد. است حال با تغییر متغیر، y را به صورت زیر تعریف می‌کنیم :

$$y = p^{-\frac{1}{2}} x^2 \tag{2}$$

این معادله‌ی یک سهمی است .

در معادله‌ی (۱)، یک x در طرفین ضرب می‌کنیم. حاصل برابر است با:

$$x^4 + px^2 = qx$$

با توجه به رابطه‌ی (۲) داریم:

$$py^2 + px^2 = qx$$

با انجام عملیات جبری این معادله تبدیل می‌شود به:

$$\left(x - \frac{q}{2p}\right)^2 + y^2 = \left(\frac{q}{2p}\right)^2 \tag{3}$$

که یک دایره به مرکز $\left(\frac{q}{2p}, 0\right)$ به شعاع $\frac{q}{2p}$ و به قطر $\frac{q}{p}$ است. بنابراین ریشه‌ی مثبت معادله درجه‌ی سومی که در (۱) آمده برابر با مختصه‌ی x نقطه‌ی تقاطع دایره‌ی (۳) و سهمی (۲) است .

به عنوان مثال می‌توانیم معادله‌ی درجه‌ی سه با $p = 4$ و $q = 8$ را بررسی کنیم. در این‌جا سهمی مورد بحث $y = \frac{x^2}{2}$ و دایره به مرکز $(1, 0)$ و شعاع ۱ خواهد بود. معادله‌ی مورد نظر به صورت روبه‌رو است: $x^3 + 4x - 8 = 0$ که ریشه‌ی مثبت آن $x = 1/365$ است.

روش هندسی‌ای که امروزه به کار می‌رود، رسم نمودار تابع $y = x^3 + 4x - 8$ (شکل ۳) است و محل تقاطع منحنی با محورxها ریشه‌ی (ریشه‌های) حقیقی معادله را به‌دست می‌دهد. این مثال تنها یک ریشه‌ی حقیقی دارد (op در شکل ۳) و دو ریشه دیگر موهومی هستند زیرا منحنی در نقاط دیگری محور xها را قطع نکرده است.

به وضوح روش خیام یک روش هندسی بود. به عبارت دیگر معادله‌ی درجه‌ی سه، معادله‌ای بین احجام در نظر گرفته می‌شود. لذا x نشان‌دهنده‌ی یک یال از مکعب و p بیان‌کننده‌ی یک مساحت ($p > 0$) در حالی که q دلالت بر حجم مکعب می‌کند.

حال به ترسیم شکل هندسی با استفاده از معادلات (۲) و (۳) می‌پردازیم: ابتدا مربعی به مساحت p که ضلع آن $\sqrt{p}$ است رسم می‌کنیم. ضلع مربع را OA می‌نامیم (شکل ۴). خطی عمود بر OA در نقطه O رسم می‌کنیم. به اندازه‌ی $\frac{p}{q}$ روی این خط جدا کرده و آن را OE می‌نامیم. روی این پاره‌خط دایره‌ای به مرکز D و قطر $\frac{p}{q}$ رسم می‌کنیم. با در نظر گرفتن نقطه O به عنوان رأس، سهمی‌ای را با پارامتر OA رسم می‌کنیم.(محور سهمی خط OR که ادامه OA است) (سهمی قبل از خیام در کارهای آپولونیوس[۱۳] (۲۱۰ قبل از میلاد) تعریف و رسم شده بود). از نقطه P محل تقاطع دایره و سهمی، عمودی بر پاره‌خط OE رسم می‌کنیم. نقطه‌ی پای عمود را Q می‌نامیم. در نهایت طول OQ پاسخ معادله است. در این‌جا محورها تنها برای خوانندگانی که مأنوس با امروزی نگارش ریاضیات هستند مشخص شده و کاری به رسم هندسی ما ندارد.

لازم به ذکر است که فرمول کاردانو [۱۴]در سال ۱۵۴۵ که در رساله‌اش Ars Magna ظاهر می‌شود؛ به طور دقیق ریشه‌ی مثبت برای این معادله را مشخص می‌کند:

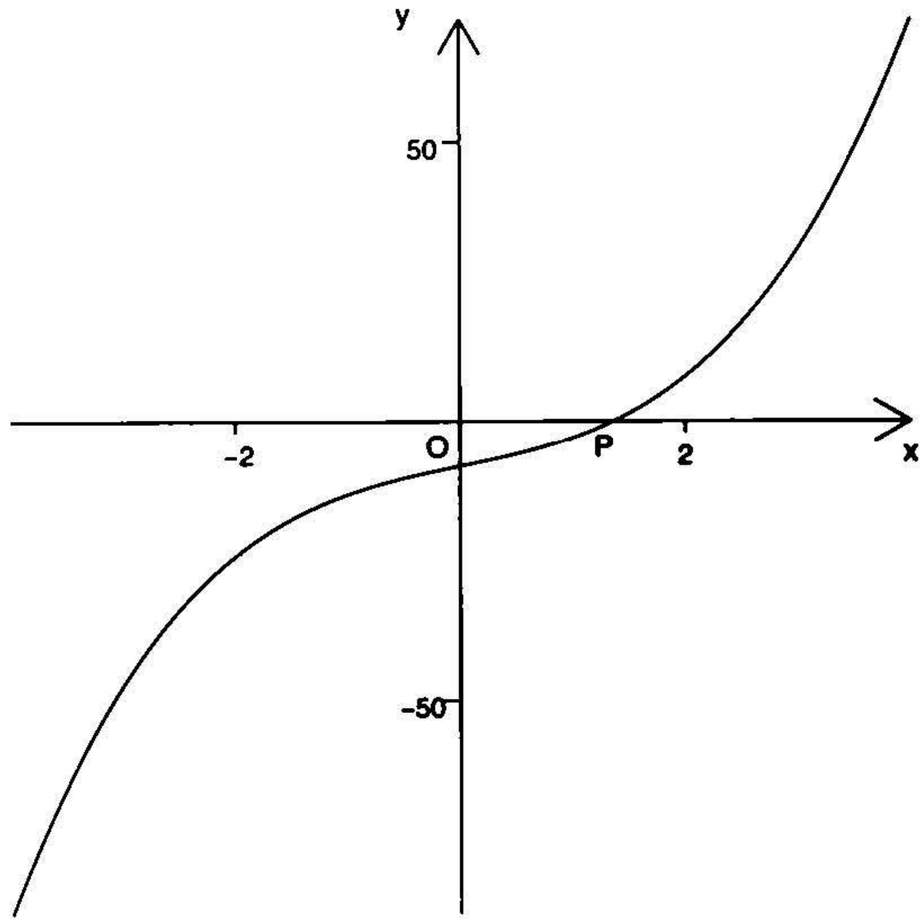
$$\left[\frac{q}{2} + \left(\frac{q^2}{4} + \frac{p^3}{27}\right)^{\frac{1}{2}}\right]^{\frac{1}{3}} + \left[\frac{q}{2} - \left(\frac{q^2}{4} + \frac{p^3}{27}\right)^{\frac{1}{2}}\right]^{\frac{1}{3}}$$

به علاوه توجه کنید که معادلات درجه‌ی سه می‌توانند سه ریشه‌ی مثبت داشته باشند. یکی از معادلاتی که خیام در نظر گرفت $x^3 + \frac{13}{2}x + 5 = 10x^2$ بود (Kasir,1931,pp. 92-93) که دو ریشه‌ی مثبت و یک ریشه‌ی منفی دارد: $x = 2$ و $x = 4 \pm \frac{1}{2}\sqrt{74}$.

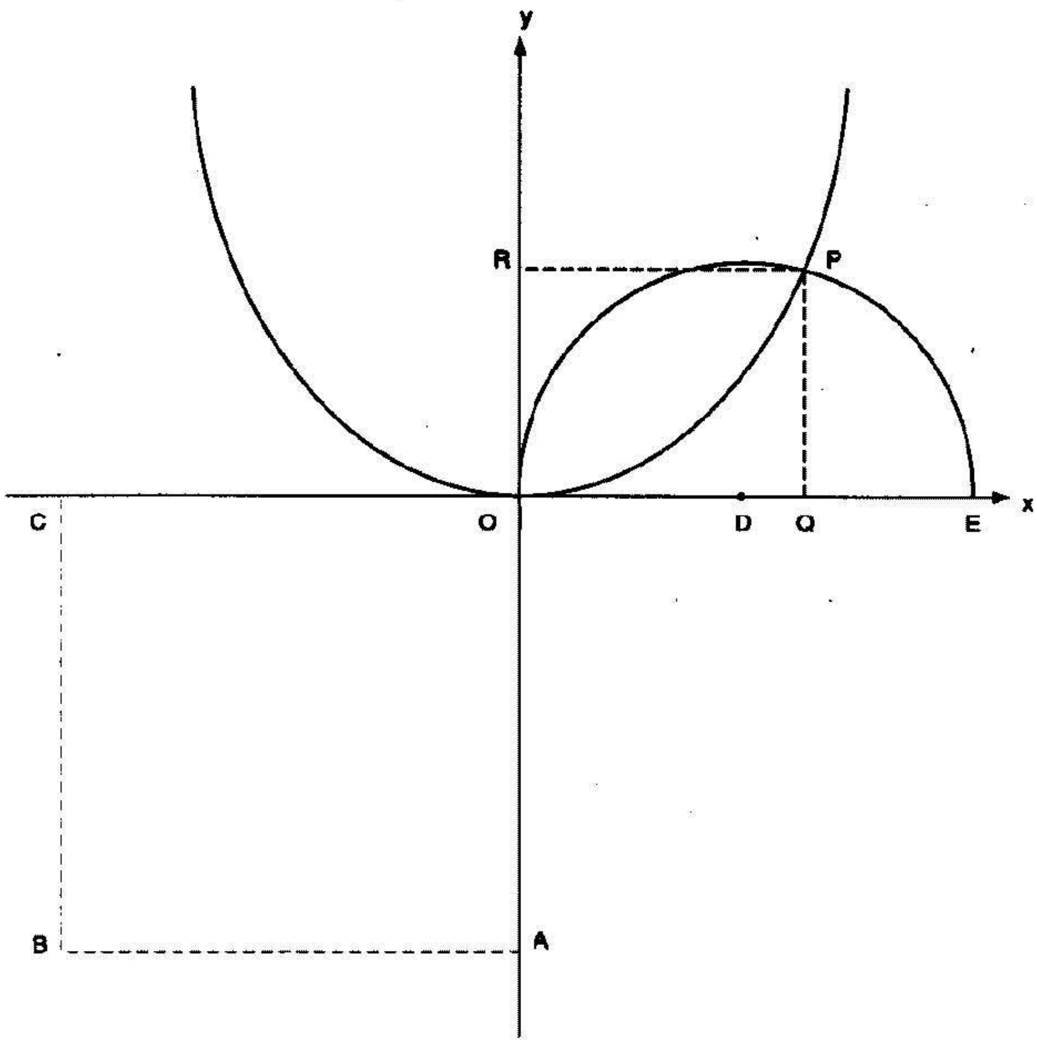
شکل ۵ نشان می‌دهد که منحنی در سه نقطه محور xها را قطع می‌کند P, Q, R.

روش خیام تنها یک یا دو ریشه‌ی مثبت را از طریق تقاطع دو مقطع مخروطی پیدا می‌کند. تمامی ۲۵ دسته‌ای که در بخش ۵ به آن‌ها

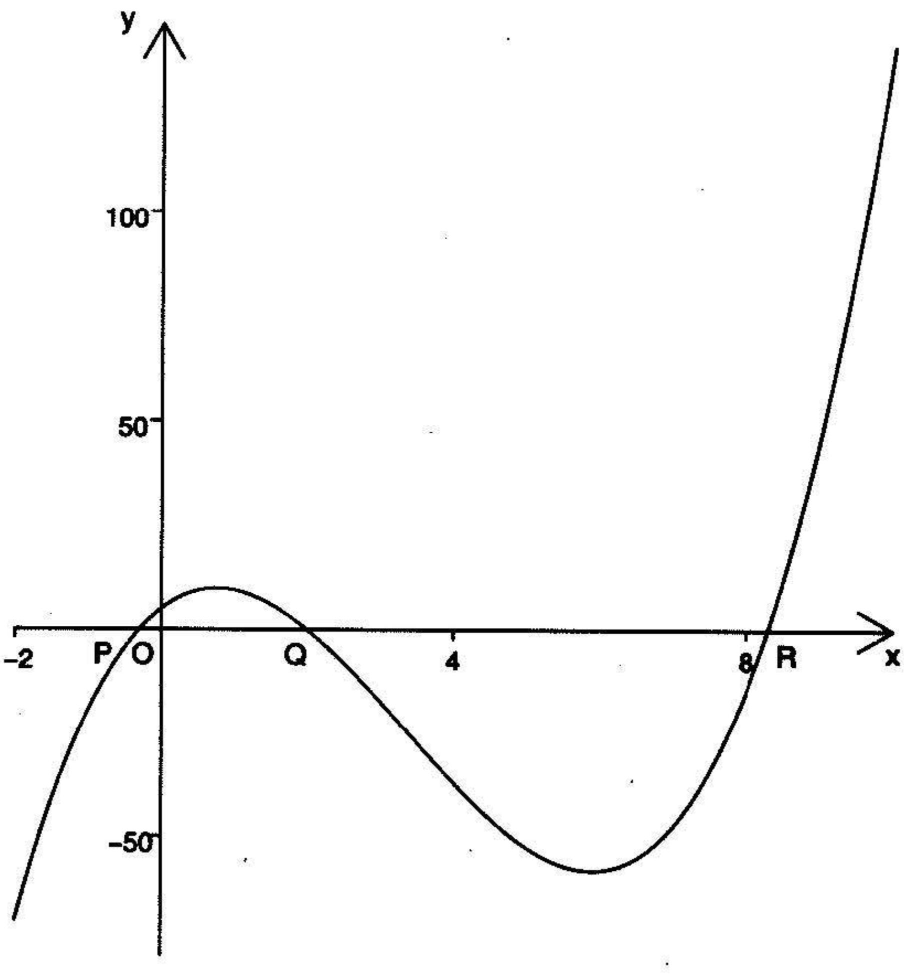
[۱۳] Apollonius
[۱۴] Cardano



شکل ۳:
منحنی $y = x^۳ + ۴x - ۸$، طول OP ریشه‌ی حقیقی معادله‌ی $x^۳ + ۴x - ۸ = ۰$ است.



شکل ۴:
روش خیام برای حل معادله‌ی درجه سوم $x^۳ + ۴x - ۸ = ۰$، پاره‌خط OQ ریشه‌ی مثبت را می‌دهد.
روش رسم مربع $ABCO$، نیم‌دایره و سهمی را از متن ببینید. طول OQ ریشه‌ی معادله است.



شکل ۵:
منحنی $y = x^{۳} - ۱۰x^{۲} + \frac{۱۳}{۲}x + ۵$ و سه ریشه‌ی حقیقی OQ، OR و OP معادله‌ی $x^{۳} - ۱۰x^{۲} + \frac{۱۳}{۲}x + ۵$

اشاره شد، توسط او به طور دقیق و منظم مورد مطالعه قرار گرفت.
وُپکه برای هریک از دسته‌های معرفی شده در بخش ۵، زوج مناسبی از مقاطع مخروطی که آن معادله را به‌دست می‌دهد معرفی کرده است.

مراجع

[۱] Halbach,H. (1975) Romance of the Rubaiyat. Helen Halbach Santa Barbara

[۲] Kasir,D.S. (1931) the Algebra of Omar Khayyam Teachers College, Columbia University, NY

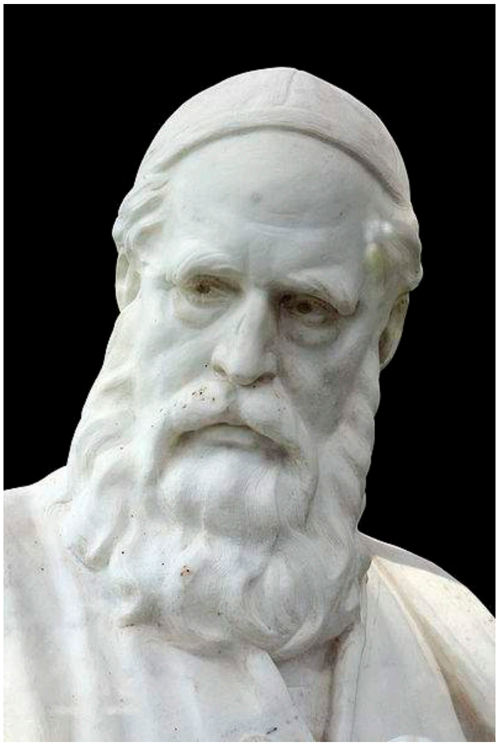
[۳] Katz,V.J. (1998) A History of Mathematics 2nd edition. Addison Wesley Longman Inc., Harlow, England.

[۴] Potter,A.G. (1929) A Bibliography of the Rubaiyat of Omar Khayyam Igpen and Grant, London

[۵] Story,W.E. (1918) Omar Khayyam as a Mathematician. Rosemary Press, Boston.

[۶] The Book of the Omar Khayyam Club (1910),1892-1910.London.Printed for the members for private circulation.

[۷] Woepcke, F. (1851)L’Algebred’Omar Al Khayyam i.Paris



هرگز دل من ز علم محروم نشد
کم ماند ز اسرار که معلوم نشد
هفتاد و دو سال فکر کردم شب و روز
معلومم شد که هیچ معلوم نشد

توپولوژی کاربردی و دانته[۱]: مصاحبه‌ای با رابرت کریس[۲]

اخیراً کاربردهایی از توپولوژی در مسائل دنیای واقع صورت گرفته است که بسیاری را شگفت‌زده کرده است. چرا که ساختار انتزاعی توپولوژی تصور کوچک‌ترین کاربردها را نیز سخت می‌کند. رابرت کریس یکی از افرادی است که در این زمینه تحقیقاتی انجام داده است. جان کوک[۳] که خود از کسانی است که در زمینه‌ی ریاضیات کاربردی فعالیت دارد و هم‌چون بسیاری از این موضوع به شگفت آمده است؛ مصاحبه‌ای با رابرت انجام داده که ترجمه‌ی نه چندان دقیق آن را در این‌جا می‌بینیم.[۴]

شکل ۱: جان کوک

[۱] Dante Alighieri شاعر ایتالیایی قرن‌های ۱۳ و ۱۴ که کمدی الهی از آثار معروف اوست.
[۲] Robert Ghrist
[۳] John D. Cook
[۴] http://www.johndcook.com/blog/2010/09/13/applied-topology-and-dante-an-interview-with-robert-ghrist/

چند هفته پیش بود که رابرت کریس را از طریق سایتش [۵] کشف کردم. رابرت پروفسور ریاضیات و مهندسی برق است. او تحقیقاتش را با عنوان توپولوژی کاربردی[۶] توصیف می‌کند؛ چیزی که من هیچ‌گاه نشنیده بودم. (توپولوژی در بخش‌های بی‌شماری از ریاضیات کاربرد دارد، اما تاکنون کاربردی از آن را در مسائل واقعی و دنیای بیرون ندیده‌ام.) علاوه بر کارهای او در توپولوژی کاربردی، علاقه‌ی رابرت به کتاب‌های قدیمی نیز مرا شیفته‌ی خودش کرد.

آنچه در ادامه می‌آید، صحبت تلفنی من با رابرت در ۹ سپتامبر ۲۰۱۰ و با اندکی ویرایش است.

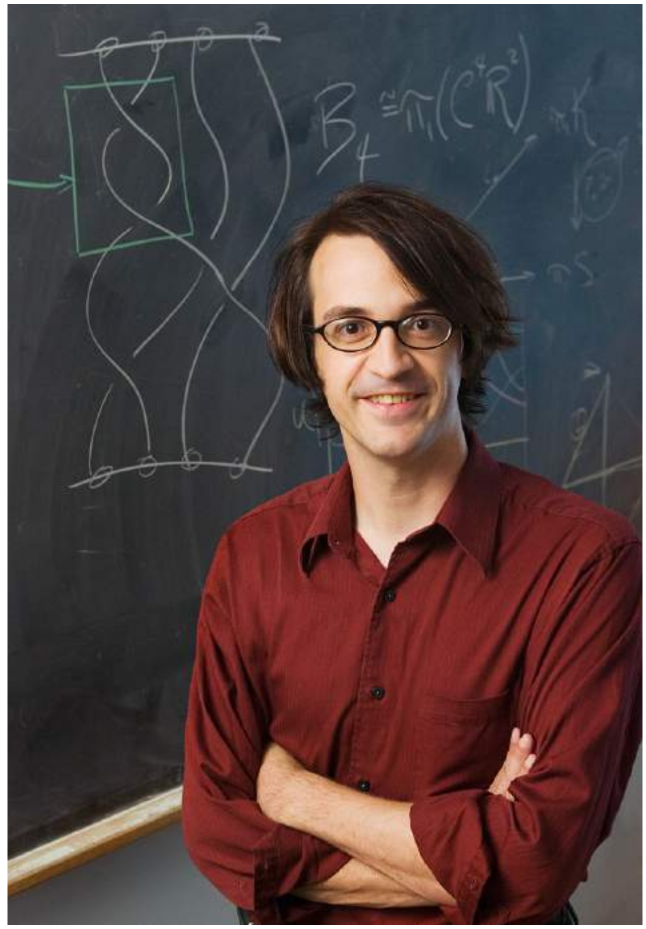
زمانی که وب سایت شخصی شما را دیدم؛ نکته‌ای که توجه مرا به خود جلب کرد تحقیقات شما در زمینه‌ی توپولوژی کاربردی بود. من در زمینه‌ی ریاضیات کاربردی و توپولوژی مطالعاتی داشته‌ام؛ اما این دو در ذهن من بسیار جدا از هم محسوب می‌شوند. بسیار هیجان‌زده‌ام که بدانم چگونه آن‌ها را با یکدیگر ترکیب کرده‌اید.

بسیاری از مردم این دو را جدا از هم تصور می‌کنند؛ اما همیشه این‌گونه نخواهد بود. و در حال حاضر می‌بینیم ابزارهایی که برای مسائل بسیار مجرد و غامض طراحی شده بودند، در مسائل نوین در آنالیز داده‌ها و سیستم‌ها کاربردهایی واقعاً ملموس دارند.

می‌توانید مثال‌هایی از این دست مطرح کنید؟

البته. یکی از اولین گروه‌هایی که در مقیاس وسیع در توپولوژی جبری کاربردی[۷] کار کردند؛ گروه گانر کارلسون[۸] در دانشگاه استنفورد بود که در زمینه‌ی آنالیز داده‌ها کار می‌کنند. فرض اولیه این است که مجموعه‌ی داده‌های در دست نقاطی در فضا هستند که به آن ابرِ نقاط[۹] می‌گوییم؛ یعنی، نمایشی گسسته از یک ساختار جالب. به طور مثال، می‌خواهید بدانید این مجموعه‌ی داده، چند مؤلفه‌ی همبندی دارد؟ این می‌تواند متناظر ویژگی‌های گوناگونی باشد. به طور مثال ممکن است این داده‌ها نظرسنجی‌هایی باشد که از مشتریان یک شرکت به دست آمده است و این مؤلفه‌ها سعی در دسته‌بندی مشتریان دارد. یا داده‌های پزشکی باشد که سعی در تشخیص انواع مختلف سرطان را دارد. سپس می‌توانید ببینید که این مؤلفه‌های همبندی، همان چیزی است که محققین آمار دسته‌بندی [۱۰] می‌نامند؛ تقسیم کردن مجموعه‌ی داده‌ها به مؤلفه‌های همبندی‌شان.

[۵] http://www.math.upenn.edu/ ghrist
[۶] Applied Topology
[۷] Applied Algebraic Topology
[۸] Gunner Carlsson
[۹] Point Cloud
[۱۰] Clustering



شکل ۲: رابرت کریس

خب؛ توپولوژی‌دانان می‌دانند که این، تنها اولین گام در برنامه‌ی بزرگتر یافتن ساختار سرتاسری فضا است. در کنار خاصیت همبندی، فضاها در حالات بسیاری حفره‌هایی دارند و روش‌های صوری جبری برای پیدا کردن و دسته‌بندی آن‌ها وجود دارد. این روش‌ها، همولوژی[۱۱]، کوهمولوژی[۱۲] و نظریه‌ی هموتوپی[۱۳] هستند. اعمال این روش‌ها روی داده‌ها تکنیک‌های واقعاً انقلابی به دست می‌دهد که نیازی به تصویر کردن داده‌ها به یک فضای دوبعدی و تلاش برای مصورسازی آن تا ببینم چه اتفاقی رخ می‌دهد نیست. این اتفاقی است که به طور خودکار می‌افتد!

کاربردهای مشابهی با همین مضمون در کارهایی که من در زمینه‌ی سیستم‌های مهندسی انجام می‌دهم وجود دارد؛ جایی که داده‌ها از شبکه‌ای از حسگرها یا یک ارتباط شبکه‌ای میان کامپیوترها می‌آیند. شما همه‌ی این داده‌های موضعی را که مثلاً از این حسگرها حاصل

[۱۱] Homology
[۱۲] Cohomology
[۱۳] Homotopy Theory

می‌شوند گردآوری می‌کنید و تلاش می‌کنید با سرهم کردن آنها به فهمی سرتاسری از محیط برسید. . این نوع تغییر از حالت موضعی به سرتاسری [۱۴] چیزی است که تکنیک‌های توپولوژی به آن منظور ابداع شده‌اند و این تکنیک‌ها در مسائلی بسیار کاربردی به طرز شگفت‌انگیزی موثر واقع شده‌اند.

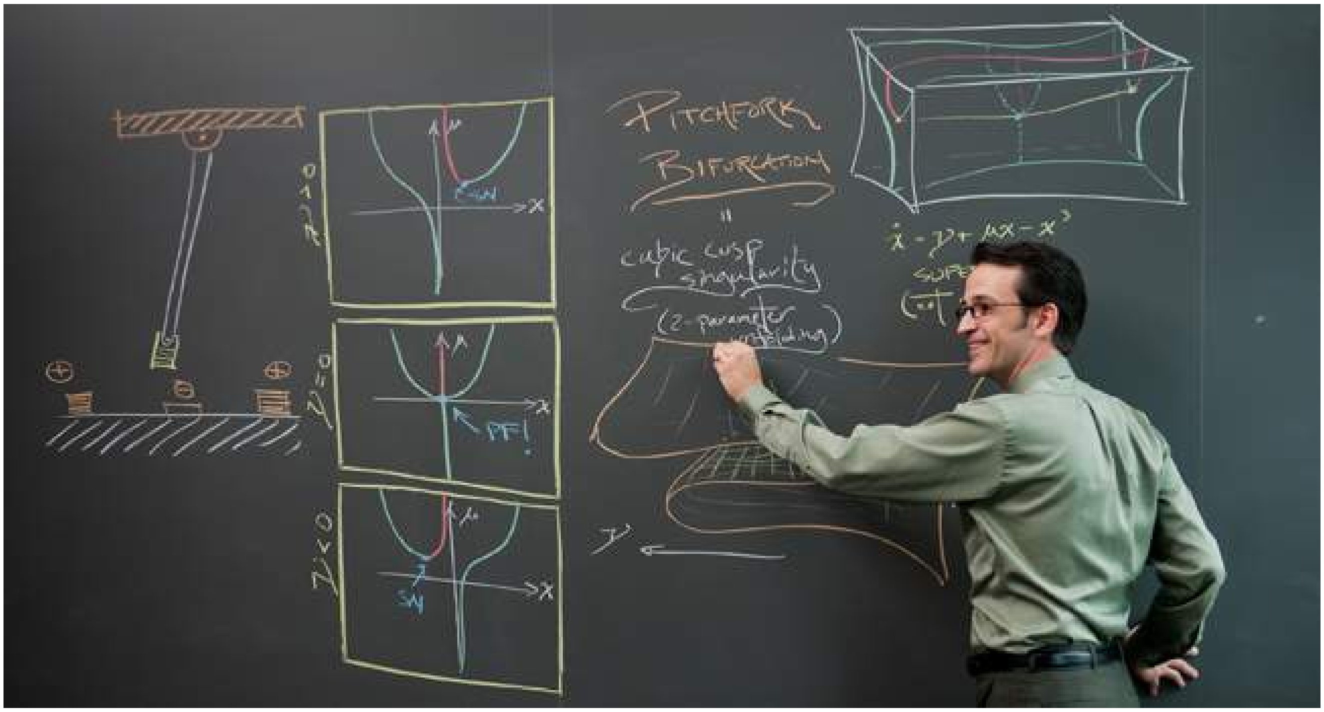
من با قضیه‌ی ون کمپن[۱۵] در هموتوپی آشنا هستم. آیا این قضیه از آن نوعی است که مدنظر شماست؟

بله. همولوژی نسبت به هوموتوپی قابلیت محاسباتی بیشتری دارد. همولوژی گزینه‌ای بسیار طبیعی‌تر است. متناظر قضیه‌ی ون کمپن، در همولوژی دنباله‌ی میر–ویتروس[۱۶] را داریم که به راستی طبیعت آن ادغام کردن است؛ چگونه داده‌های کوچک را به یکدیگر پیوند دهیم و با یک‌پارچه کردن آن‌ها در مورد کلیت شبکه اطلاعاتی به دست آوریم. این ایده‌ی بسیار عمیق و مهمی در رسیدن از داده‌های موضعی به خواص سرتاسری فضا است.

یکی از مسائلی که در این مورد ذهن مرا درگیر می‌کند، آن است که توپولوژی برای من به نوعی شکننده به نظر می‌رسد؛ به این معنی که اتصالی تصادفی می‌تواند چیزی را عوض می‌کند. دو نقطه را به هم وصل کنید و ناگهان فضای ناهمبند، همبند می‌شود. بنابراین به نظر می‌رسد که این چینش توپولوژیک مشکلی دارد: از بین رفتن مقدار جزئی داده یا خطا در داده‌ها نتایج حاصل را باطل می‌کند. اما به یاد دارم در وب سایت شما، نظرتان کاملاً خلاف این تصور بود و متدهای توپولوژیک را پایدارتر می‌دانستید. من در تصمیم گرفتن میان این دو دیدگاه مشکل دارم.

انواع مختلفی از پایداری [۱۷] وجود دارد که در موارد کاربردی گوناگونی اهمیت پیدا می‌کنند. چون ساختارهای توپولوژی جبری تحت هموتوپی و دگردیسی [۱۸] ناوردا هستند، تحت این تغییرات –به عنوان نمونه تغییر مختصات– بسیار پایدار خواهند ماند. این زمانی که شما با داده‌های در دنیای واقع سروکار دارید، بسیار مفید است. به عنوان مثال فرض کنید داده‌ها توسط تلفن‌های همراهمان جمع‌آوری شده باشند. یک جفت حسگر روی یک تلفن همراه نصب کنید و از این طریق مقدار زیادی جریان داده‌ی جذاب به دست می‌آید که هر یک از داده‌ها به یک مکان فیزیکی پیوند خورده است.

[۱۴] local-to-global transition
[۱۵] Van Kampen's Theorem
[۱۶] Mayer-Vietoris
[۱۷] Robustness
[۱۸] Deformation



اما به طور قطعی نمی‌دانید کجا؟ به طور مثال، زمانی که داخل یک ساختمان هستید GPS به خوبی کار نمی‌کند. در مواردی مشابه این، شما گونه‌ای از پایداری را می‌خواهید که مستقل از دستگاه مختصات باشد.

پایداری نسبت به نویز، و به طور خاص نسبت به خطا، مسأله‌ای به مراتب سخت‌تر از آن است که در حالت کلی قابل حل باشد. اما حتی در این مورد، ابزارهایی توپولوژیک وجود دارند که در مثال‌های خاص قابل استفاده‌اند و به موارد فنی‌تر هم‌چون همولوژی پایدار[۱۹] و خواص توپولوژیکی که روی طیفی از نمونه‌ها پایدار می‌مانند، مربوط می‌شوند.

می‌توانید مثالی بیاورید که چطور دانستن همولوژی یک مجموعه از داده‌ها ممکن است اطلاعاتی در مورد پدیده‌ی فیزیکی مربوط به آن به دست دهد؟

به عنوان نمونه یکی از مسائلی که من زیاد روی آن کار کرده‌ام، مسأله‌ی منطقه‌ی پوشش داده شده توسط شبکه‌ی حسگرهاست. برای نمونه در مورد پوشش شبکه‌ی تلفن‌های همراه صحبت می‌کنیم که همگی با آن آشنا هستند. بسیار ناخوشایند خواهد بود اگر شما وارد حفره‌ای در شبکه‌ی تلفن همراه شده و آنتن‌دهی خود را از دست دهید. پس شما می‌خواهید بدانید آیا در یک منطقه پوشش کامل دارید یا خیر؛ آیا آنجا حفره وجود دارد؟

این مسأله زمانی تنظیمات امنیتی یک منطقه مدنظر است بسیار حیاتی‌تر جلوه می‌کند. دوربین‌ها و ماهواره‌ها یک منطقه را پوشش داده‌اند و می‌خواهید بدانید آیا همه چیز پوشش داده می‌شود یا آنکه حفره‌هایی وجود دارند که شما آنجا اطلاعات را از دست می‌دهید؟ یکی از مواردی که از نظریه‌ی همولوژی استفاده می‌کنم یافتن محکی است که پوشش کامل را ضمانت کند؛ تضمین کند که شبکه‌ی حسگرهای شما هیچ حفره‌ای ندارد. پس هرچند ممکن است ندانیم که حسگرها در کجا قرار دارند و تنها از حسگرهای نزدیک به آن اطلاع داشته باشیم. همچنان می‌توان به کمک محک همولوژی برقراری پوشش را تحقیق کرد.

بسیار جالب. گمان می‌کنم اگر به داده‌های در ابعاد بالاتر نگاه کنید نمی‌توانید فقط تعدادی دایره روی نقشه بکشید و ببینید همپوشانی دارند یا خیر. به چیزی محاسبه‌پذیرتر نیاز خواهید داشت.

دقیقاً! به طور خاص اگر دایره‌ها حرکت کنند و بخواهید بدانید با

[۱۹] Persistent Homology

گذر زمان زمان چه اتفاقاتی رخ می‌دهد. به خصوص زمانی که شما مختصات ندارید و نمی‌دانید دایره‌ها را چطور رسم کنید و هم‌پوشانی آن‌ها را ببینید.

وقتی با دیگران صحبت می‌کنم؛ تمایل دارم این مطلب را مطرح کنم که یک کتاب‌خانه‌ی ریاضی را همان‌طوری می‌بینم که یک باستان‌شناس به یک سایت حفاری می‌نگرد. آنجا پر است از گنجینه‌های شگفت‌انگیزی که به خاک سپرده شده‌اند و از دید جهانیان پنهان مانده‌اند. برای مثال یک کتاب نوعی در مورد نظریه‌ی بافه‌ها[۲۰] غیرقابل خواندن است! اما پر از مطالبی است که برای حل مسائل واقعاً سخت، بسیار مهم جلوه می‌کنند و من با این دید، از طریق کندوکاو این متون مبهم و یافتن این گنجینه‌های باارزش، آن‌ها را وارد مهندسی و دیگر حوزه‌ها می‌کنم؛ جایی که این ابزارها کارایی خواهند داشت.

در حال حاضر، بسیاری از این ریاضیات غامض از این حصار گذشته‌اند. دیگر کسی ادعا نخواهد کرد که نظریه‌ی اعداد بی‌فایده است. ولی توپولوژی جایی است که شما سودمندترین و کمتر به کار گرفته شده‌ترین ابزار را دارید! این دیدگاه من بود در مورد آنچه که می‌خواهم در ریاضیات رخ دهد و آنچه که خودم می‌خواهم موفق به انجامش شوم.

کمی موضوع بحث را عوض کنیم، مطلب دیگری که در سایت شما توجه مرا به خود جلب کرد، نقل قول‌های شما از کتاب‌های قدیمی بود.

خواندن هر چیزی کمتر از ۵۰ سال مانند نوشیدن شراب تازه است: می‌توانید یک یا دو بار در سال استفاده کنید و معمولاً با پشیمانی و سردرد همراه است.

من یاد نصیحت لوئیس[۲۱] افتادم که می‌گوید در کنار خواندن کتاب‌های جدید، کتاب‌های قدیمی را هم وارد کنید زیرا هر دوره‌ای نقاط کور خود را دارد. نویسندگان قدیمی نقاط کور خود را دارند اما ممکن است مکمل دیدگاه ما باشند.

دقیقاً! من قطعاً از این حکم پیروی کرده‌ام، شاید حتی کمی زیاده از حد! من به ندرت کتابی مدرن می‌خوانم. اما در مورد موسیقی، فیلم و وبلاگ‌ها از این قانون پیروی نمی‌کنم. اما در مورد کتاب، چیزهای جالب بسیاری وجود دارند که از آزمون زمان سربلند بیرون آمده‌اند و با این حساب چیزهای جالبی که باید بخوانم تمام نمی‌شوند! من در دوران کالج، تجربه‌ی فوق‌العاده‌ی گرفتن درس‌هایی کتاب‌محور [۲۲] که مقدار زیادی بحث و مطالعه را دربرمی‌گرفت داشته‌ام. در واقع این دوره دیدگاه مرا عوض کرد و باعث شد که عاشق کلاسیک‌ها باشم و با بسیاری از این کتاب‌ها زندگی کنم.

زمانی که چنین درس‌هایی را می‌گرفتید، مشغول خواندن ریاضی بودید؟

نه، من مهندسی خواندم. مدرک دوره‌ی کارشناسی من، مهندسی است. در زندگی کمی دیرتر وارد ریاضیات شدم!

شما گفتید که عاشق دانته هستید.

بله، صحیح است. من مدت زیادی را با کتاب او، کمدی الهی[۲۳]، زندگی کرده‌ام و هنوز هم هر بار که آن را باز می‌کنم ایده‌های جدیدی در آن می‌بینم. اکثر مردم چندان از دوزخ دور نمی‌شوند. دوزخ [۲۴] هیجان‌انگیز و قسمت اکشن داستان است. اما قسمت‌های بعدی داستان -برزخ و بهشت- مکان‌هایی واقعاً زیبا برای زندگی هستند.

به منظور آنکه دانته را بخوانید مجبور شدید مطالعات تاریخی زیادی انجام دهید؟

اگر یک ترجمه‌ی خوب و مجموعه‌ای از کامنت‌های مناسب داشته باشید، خواندن را بسیار ساده‌تر می‌کند. من یک ترجمه از دروتی سایرز[۲۵] پیدا کردم که کامنت‌های عالی‌ای داشت. او در اواخر عمر با دانته آشنا شد و داستانش بسیار جالب است. در ابتدای ۱۹۴۰، زمانی که بمباران شروع شد، او در راه رفتن به پناهگاه تصمیم گرفت یک کتاب از قفسه بردارد و یک نسخه از دانته را دید. خودش می‌گوید: "می‌دانید؟ من در واقع هیچ‌گاه دانته نخوانده‌ام." آن را از قفسه بیرون کشید. طی دو روز بعد، نه غذا خورد و نه خوابید. او مجذوب داستان و استادانه بودن آن شد و تمام باقی زندگی‌اش را برای تسلط به زبان ایتالیایی و ترجمه‌ی دانته گذاشت.

[۲۰] Sheaf Theory
[۲۱] C. S. Lewis نویسنده‌ی کتاب سرگذشت نارنیا
[۲۲] book-type courses
[۲۳] The Divine Comedy
[۲۴] کمدی الهی شامل سه بخش دوزخ، برزخ و بهشت است.
[۲۵] Dorothy Sayers

گذری بر فلسفه و آموزش ریاضی

سامان حبیبی اصفهانی

یک سؤال همیشگی دانش‌آموزان این است که: " چرا ما این‌قدر ریاضی می‌خوانیم؟!! "

تقریباً در تمام طول تاریخ آموزش، از زمان فیثاغورسیان تا به امروز و در تمام کشورها ریاضیات پررنگ‌ترین نقش را در برنامه‌های آموزشی مدارس داشته است. ساعت‌های بسیاری از وقت کودکان و نوجوانان همواره سر کلاس‌های ریاضی گذشته. ولی آیا این اتفاق واقعاً ضروری است یا صرفاً سخت‌گیری در آموزش ریاضیات تبدیل به یک عادت شده‌است؟ کافی است بیاندیشید که چه جایگزین‌های هیجان‌انگیزی برای پر کردن وقت مدارس وجود دارد.

به نظر می‌رسد گاهی نه دانش‌آموزان از هدف ریاضی خواندن‌شان مطلع‌اند و نه حتی دبیران‌شان. شاید یافتن پاسخی مناسب به سود همه باشد و ما را در تشخیص این‌که تحت عنوان ریاضی "چه درس بدهیم" و "چگونه درس بدهیم" راهنمایی کند.

اما برای پاسخ دادن به این پرسش باید ابتدا روی این‌که ریاضی چیست به یک توافق نظر برسیم.

اگر در مسیر تاریخ عقب برویم تا به افلاطونیان برسیم خواهیم دید آن‌ها این طور می‌اندیشیدند که جهانی به نام عالم مُثُل، خارج از زمان و مکان، خارج از ماده و خارج از اندیشه وجود دارد که شامل حقیقت هر چیزی است. اشیاء ریاضی نیز از جمله عدد، نقطه، خط، دایره، تابع و ... متعلق به این جهان مثلی بوده و مستقل از هرگونه آگاهی و مستقل از انسان وجود دارند. ازلی و ابدی بوده و دنیای پیرامون‌مان را تنها سایه‌ای از آن جهان می‌دانستند. از آن‌جایی که این جهان اثری از جهان مثل است، مفاهیم و گزاره‌های ریاضی بر خلاف سایر علوم همواره صادق و درست بوده و پیشینی‌اند. افلاطونیان گزاره‌های ریاضی را ضروری می‌پنداشتند. گزاره‌هایی که شکی در صحت آن‌ها نیست. آن‌ها معتقد بودند ما اشیاء و حقایق ریاضی را کشف می‌کنیم نه این‌که آن‌ها را بسازیم. در نتیجه یک ریاضی‌دان یک ابداع‌کننده نیست بلکه یک کاشف است درست همانطور که یک زمین شناس دست به اکتشاف می‌زند. راسل در جایی می‌نویسد که ریاضی‌دانان حساب را کشف کردند همان‌طور که کریستف کلمب هند غربی را کشف کرد. ما اعداد را خود نساختیم همان‌طور که کلمب سرخ‌پوستان را نیافرید. در این فلسفه دلیلی محکم برای تدریس ریاضی خواهیم داشت چرا که کشف حقایق ریاضی همان کشف حقیقت است.

به مرور زمان دو اتفاق مهم برای ریاضی رخ داد. یکی این‌که تا حدودی جایگاه خود را به عنوان علمی قطعی از دست داد و به مجموعه‌ای از اصول که ساخته ذهن انسان‌هاست تغییر هویت داد و دیگری این‌که از نظر عموم کاربرد ریاضیات از خود آن مهم‌تر شد. ریاضی تبدیل شد به ابزاری قوی برای مدل‌سازی تا حقیقتی که خود ارزش مطالعه داشته باشد.

نگاه اخیر، ریاضی را بیشتر برای سایر علوم به خصوص رشته‌های مهندسی می‌خواهد تا برای خود ریاضی. به نظر می‌رسد این همان نگاهی است که معمولاً در تدریس ریاضی در مدارس حاکم است؛ یعنی، ریاضیات دبیرستانی پیش‌نیازی برای مهندس شدن است تا برای ریاضی‌دان شدن. نتیجه‌ی این نگاه به وجود آمدن حفره‌های فراوان در ریاضیات مقدماتی است که در مدارس آموزش داده می‌شود و این در تضاد با این فرهنگ ریاضی است که میل دارد همه چیز را با دقت اثبات کند و به پیش رود. برای نمونه در اثبات این‌که $\sqrt{۲}$ عددی گنگ است؛ فرض را بر گویا بودن آن می‌گیریم و سپس به تناقض می‌رسیم. در این اثبات حقیقی بودن $\sqrt{۲}$ بدیهی فرض شده که برای اثبات آن به خاصیت تمامیت اعداد حقیقی احتیاج داریم و چندان هم بدیهی نیست که یک عددی اعشاری وجود دارد که به توان ۲ می‌شود ۲. در واقع یک دانش‌آموز که پیش‌زمینه‌ای از ریاضی ندارد هیچ دلیلی برای بدیهی دانستن آن هم نخواهد داشت. این بدیهی‌انگاری در بسیاری از بخش‌های ریاضی مقدماتی وجود دارد و معلول روش آموزش است تا این‌که واقعاً بدیهی باشد.

به عنوان مثالی دیگر، معمولاً اعمال حساب برای اعداد گویا تعریف می‌شوند و دانش‌آموزان مدت‌ها با آن‌ها کار می‌کنند. در واقع اعداد گویا تمام جهان ریاضی آن‌ها را می‌سازند. سپس در دبیرستان اعداد به مجموعه‌ی اعداد حقیقی گسترش می‌یابند و بدون هیچ توضیحی همه‌ی دانسته‌های قبلی در مورد اعداد حقیقی نیز به کار می‌روند. اگر قبلاً در مورد درستی اتحادی کنکاش کرده‌اند (در صورتی که اعداد موجود در رابطه گویا بوده‌اند.) حال هیچ شکی در مورد صحت آن در مورد همه‌ی اعداد حقیقی پیش نمی‌آید و این قطعاً چیزی است که از طرف دبیرانشان به آن‌ها آموخته می‌شود یا دانش‌آموز که قبلاً با تساوی

$$a^{m/n}.a^{p/q}=a^{m/n+p/q}$$

روبه‌رو شده که m/n و p/q اعدادی گویا هستند، حال همین رابطه

را به صورت

$$a^s.a^t = a^{s+t}$$

می‌پذیرد که s و t حقیقی‌اند اما این‌که واقعاً a^s -هنگامی که توان عددی گنگ است- یعنی چه؟ برای دانش‌آموز اصلاً واضح نیست اما بدون هیچ ابایی با این روابط به صورت فرمال کار می‌شود. یادآوری می‌کنم که اگر ریاضیات را به عنوان پیش‌نیازی برای رشته‌های مهندسی بخواهیم؛ وجود چنین حفره‌هایی چندان اهمیتی نخواهد داشت و صرف توانایی کار کردن با روابط کافی به نظر می‌رسد.

ولی آیا باید این حفره‌ها پر شوند و آیا اصلاً چنین کاری ممکن است؟ به این سؤال باز خواهیم گشت.

حال به این مشکل معلم‌ها را نیز اضافه کنید. بسیاری از آن‌ها مشکل‌ترین درس ریاضی که گذرانده‌اند احتمالاً ریاضی عمومی بوده است. برای نمونه یک امتحان ساده از دبیران ریاضی نشان می‌دهد بسیاری از آن‌ها حتی به‌درستی تمایز تعریف، اصل و قضیه را نمی‌دانند و این‌که برای مثال آیا ض ز ض برای همنهشتی دو مثلث تعریف است، یا اصل یا قضیه!؟

آیا می‌توان ادعا کرد که اگر مدرسین ریاضی همگی تحصیل کرده‌ی همین رشته بودند بسیاری از مشکلات آموزشی حل می‌شد؟!

اچ. وو در [۱] تحت عنوان گم‌شده‌های آموزش معلمان ریاضی چنین می‌نویسد :

> ....یک سناریو مطرح است. اگر بخواهیم معلم‌های زبان فرانسه‌ی خوبی داشته باشیم آیا بایستی آن‌ها را ملزم کنیم که به جای فرانسوی در کالج مشغول به تحصیل لاتین شوند؟ هر چه باشد لاتین زبان مادر فرانسوی است و به لحاظ ساختار زبانی بسیار پیچیده‌تر است. با مسلط شدن بر زبان لاتین معلم‌ها باید قادر باشند تا درک‌شان از زبان فرانسه را که در مدرسه آموخته‌اند بالا ببرند.
>
> تصور این‌که برای تدریس ریاضیات مقدماتی معلمین باید مباحث پیشرفته ریاضی دانشگاهی را بلد باشند همان قدر مضحک است که سناریوی بالا در مورد تدریس زبان فرانسه. تصور کنید برای آموزش جمع و ضرب ابتدایی دبیر ملزم باشد نظریه‌ی گروه‌ها بداند ...

در هر صورت این یک سؤال چالش برانگیز است که انتظار چه سطحی از سواد از دبیران ریاضی می‌رود. در این زمینه تا حد زیادی می‌توان همان سندروم فرانسه-لاتین را مشاهده کرد. جامعه‌ی ریاضی‌دانان از دبیرهای ریاضی مدارس (مخصوصاً در سطح دبیرستان) انتظار داشتن سواد عمیقی در مباحث پیشرفته ریاضی را دارند. باید به نوعی بتوان رابطه سطح دانش دبیران (در دو بخش ریاضی مقدماتی و ریاضی پیشرفته) و دانش‌آموزان‌شان را برای پاسخ دادن به این پرسش سنجید. یکی از افرادی که سعی کرده به این مسأله پاسخ دهد بلج (مدیر SMSG) است. وی در سال ۱۹۷۲ از میان ۳۰۸ معلم جبر سال اول دبیرستان برای هر دو گروه معلمین و دانش‌آموزان آزمونی چند گزینه‌ای برگزار کرد تا رابطه‌ی دانش دبیران با دانش‌آموزان‌شان را به‌دست آورد. با مسامحه می‌توان گفت که نتیجه‌ی آزمایش وی نشانی از این ادعا که " در حالت کلی هر چه بیشتر درباره‌ی موضوعی بدانی معلم بهتری نیز خواهی بود " نداشت . ([۱])

امتحان وی برای معلمین شامل دو بخش بود. یکی جبر اعداد حقیقی و دیگری جبر مجرد (شامل نظریه گروه‌ها، حلقه‌ها و میدان‌ها). آنالیز نتایج حاصل تاکید می‌کرد که

> ... تقریباً هیچ رابطه‌ی معنی‌داری میان نمرات دبیران در جبر مجرد با نمره دانش‌آموزان‌شان نیست اما از طرفی نمرات دبیران در جبر مربوط به اعداد حقیقی همبستگی مثبت کوچکی با نمرات دانش‌آموزان‌شان در جبر سال اول دبیرستان دارد...

نتیجه‌ی وی از مشاهداتش این بود که داشتن دانش عمیق از مباحث مجرد ریاضی برای تدریس در دبیرستان نه لازم است و نه کمکی می‌کند اما هر چه عمیق‌تر بودن درک دبیر از همان مباحث دبیرستانی تاثیر مثبتی بر سطح آموزش وی دارد. ولی قطعاً این همه‌ی ماجرا نیست.

اساساً مشکلاتی که در تدریس ریاضی برای دبیران پیش می‌آید غالباً مشکلات وابسته به سواد ایشان نیست. برای مثال یکی از بزرگ‌ترین سختی‌های آموزش ریاضی، آموزش تقسیم است.

پیش از این‌که به بررسی این مشکل بپردازیم؛ نگاهی به این داشته باشیم که همین مفهوم در ریاضیات پیشرفته چطور تدریس می‌شود. اگر

$$S = \{(x,y) | x \in \mathbb{Z}, y \in \mathbb{Z} - \{۰\}\}$$

رابطه‌ی $\sim$ را روی S چنین تعریف می‌کنیم:

$$(x,y) \sim (z,w) \iff xw = yz$$

کلاس هم‌ارزی $[(x,y)]$ را روی S با $\frac{x}{y}$ نمایش می‌دهیم و مجموعه‌ی همه‌ی کلاس‌های هم‌ارزی را $\mathbb{Q}$ می‌نامیم. حال با تعریف زیر $\mathbb{Q}$ را به یک حلقه تبدیل می‌کنیم:

$$\frac{x}{y} + \frac{z}{w} = \frac{xw+zy}{yw}, \frac{x}{y}.\frac{z}{w} = \frac{xz}{yw}$$

بیایید بررسی کنیم که یک مدرس به چه روشی می‌تواند این مفاهیم را به یک دانش‌آموز ابتدایی منتقل کند؟ احتمالاً به هیچ روشی!

اجازه دهید مراحل کار را با دقت بیشتری بررسی کنیم. مرحله‌ی اول نیاز به فهمیدن افراز S به کلاس‌های هم‌ارزی دارد به طوری که به هر کلاس به چشم یک عضو نگاه کنیم. سپس یکی بودن $\mathbb{Z}$ و $\{\frac{x}{1} : x \in \mathbb{Z}\}$. مشکل دیگر تعریف جمع و ضرب اعداد گویا است. تصور کنید که عمل جمع قبلاً تعریف شده‌است و حال می‌خواهیم ضرب را تعریف کنیم. تعریف ضرب به صورت $\frac{x}{y}.\frac{z}{w} = \frac{xz}{yw}$ برای ما معنی‌دار است چرا که ما می‌خواهیم روی $\mathbb{Q}$ ساختار حلقه بگذاریم و این تعریف بدیهی‌ترین روش ممکن برای انجام این کار است. ولی آیا می‌توانیم به یک دانش‌آموز توضیح دهیم که ساختار حلقه برای ما اهمیت دارد و این تعریف ضرب روش مناسبی برای ساختن آن است؟ چرا به جای تعریف رایج جمع به سراغ $\frac{x}{y} + \frac{z}{w} = \frac{x+z}{y+w}$ نمی‌رویم؟ (که احتمالاً برای دانش‌آموزان بسیار خوشحال کننده خواهد بود!)

در واقع دانش‌آموز هیچ شهودی در مورد چرایی تعریف جمع دو عدد گویا کسب نمی‌کند. (این که مخرج مشترک گرفتن به معنی یکی کردن واحد شمارش برای دو نسبت مختلف است ...)

الگوریتم جمع دو کسر، محاسبه ک م م مخرج‌ها و سپس جمع صورت‌ها کاملاً صوری باقی می‌ماند و در بهترین حالت به کمک مثال معروف "پیتزا" توصیفی برای آن ارائه می‌شود. در واقع تدریس تقسیم، نمونه‌ی مناسبی برای این ادعای دانش‌آموزان است که وظیفه‌ی ما در درس ریاضی چرایی نیست بلکه ضرب و تقسیم است .

به طوری ضمنی به دانش‌آموزان القا می‌شود که مفاهیم تقسیم، نسبت و کسر هم‌معنی هستند و می‌توان آن‌ها را به جای هم به کار برد. در واقع این سه مفهوم با یک الگوریتم تقسیم در ذهن دانش‌آموز ضمیمه می‌شوند حال آن‌که نه تنها هیچ شهودی در مورد جمع کسرها ندارد بلکه الگوریتم تقسیم هم برای وی یک بازی با اعداد است و نه چیزی بیشتر. این مشکلات باز هم در تعریف اخیر از ریاضی که آن را برای سایر علوم می‌خواهد اهمیت چندانی ندارد. یادگیری جمع دو عدد کسری و الگوریتم تقسیم بدون فهم مفاهیم پشت آن هم کافی به نظر می‌آید. در واقع همان‌طور که بارها در زندگی‌مان با این گزاره مواجه شده‌ایم؛ هدف تعیین می‌کند چه چیزی درست است و چه چیزی نادرست.

اما اگر یکی از اهداف تدریس ریاضی را کمک به درست استدلال کردن و منطقی فکر کردن دانش‌آموزان بدانیم (چیزی که واقعاً ارزش وقت زیادی را که از دانش‌آموزان می‌گیرد، دارد.) و این‌که از اهداف ریاضی این بدانیم که برای هر چیزی باید دلیلی داشت و بسیاری از چیزهایی که بدیهی به نظر می‌آیند واقعاً بدیهی نیستند؛ این روش چندان خوشایند نیست.

یکی از پیشنهادهایی که برای حل مشکلات ذکر شده مطرح می‌شود؛ این است که مجدداً به این فرهنگ بازگردیم که ریاضی را برای ریاضی تدریس کنیم. بسیاری از حفره‌های پیش آمده به راحتی با استدلال‌های شهودی و نسبتاً ساده قابل پوشش‌اند. حتی اگر کاملاً به طور ریاضی دقیق نباشند، این فرهنگ ریاضی آموزش داده می‌شود که برای هر چیزی باید استدلالی ارائه کرد.

در مورد مثال تقسیم، این مشکل تاریخی که نسبت و عدد برای سال‌ها دو مفهوم متمایز دانسته می‌شدند برای دانش‌آموز ابتدایی هم رخ خواهد داد. یک پیشرفت در تدریس مفهوم کسر استفاده از نمایش هندسی است تا ارائه مفهوم جبری. به این‌گونه که اگر اعداد حقیقی را به عنوان خط متصل در نظر بگیریم و کسر ۴/۳ را روی خط کمی جلوتر از ۱، به عنوان یک نقطه معرفی کنیم تا یک موجود ریاضی جدید با تعریف جمع و ضرب خاص خود. به همین شکل می‌توان برای جمع کسرها -به عنوان جمع طول دو پاره خط- یک استدلال هندسی ساده آورد. به نظر می‌آید همین نکته راه‌حلی در مورد بسیاری از حفره‌های دیگر و روشی برای پوشش آن‌ها به ما می‌دهد. این‌که گاهی استفاده از ایده‌های هندسی و بیان اثبات‌های تصویری می‌توانند جایگزین اثبات‌های جبری شوند.

در مورد $\sqrt{۲}$ که پیش‌تر بحث شد؛ اگر دانش‌آموز متوجه شود که به مجموعه‌ی اعداد حقیقی می‌توان به دو صورت نگاه کرد -یکی جبری (مجموعه‌ی اعداد اعشاری) و یکی هندسی (خط متصل)- می‌توان به جای اثبات این‌که معادله $x^۲ = ۲$ در مجموعه اعداد حقیقی ریشه دارد، $\sqrt{۲}$ را روی خط اعداد حقیقی نشان داد. در واقع یکی از مهم‌ترین خاصیت‌های آموختن ریاضی درست و منطقی فکر کردن و توانایی ارائه‌ی استدلال است که با بدیهی‌انگاری‌ها، این بخش از آموزش آن کم‌رنگ شده‌است. به نظر می‌رسد بیش از آن‌که به دنبال دبیرانی دارای سوادی عمیق از ریاضیات دانشگاهی باشیم بایستی دبیرانی تربیت کرده باشیم که شناختی از فرهنگ ریاضی داشته و توانایی آموزش هنر استدلال کردن به دانش‌آموزان خود را به‌دست آورده باشند.

به هر حال باز اشاره به این نکته ضروری است: این‌که آموزش ریاضی باید چگونه باشد با هدف ما از تدریس آن رابطه‌ی مستقیم دارد.

مراجع

[۱] Hung-Hsi Wu. *The Mis-Education of Math Teachers*. Notices of the American Mathematical Society, 2011.

[۲] Platonism in the Philosophy of Mathematics. *http://plato.stanford.edu/index.html*.

[۳] محمد صالح مصلحیان. *فلسفه‌ی ریاضی*.

بررسی تاریخچه‌ی قانون تقابل مربعی

محی‌الدین متوصل، امید رفیعیان، حمید ملک

۱ مقدمه

آن‌چه در وهله‌ی اول، درباره‌ی قانون تقابل مربعی جلب توجه می‌کند، تعداد اثبات‌های بسیار زیاد ارائه شده برای آن است. اگر تعریف صریح و کامل این قانون را متعلق به سال ۱۷۸۳ بدانیم، از آن سال تاکنون ۲۳۳ اثبات مختلف برای این قانون ارائه شده است. رقم خوردن میانگین بیش از یک اثبات در هر سال، برای یک قضیه، پیش از هر چیز، دلالت بر وجود یک نقش تاریخی در این قانون می‌نماید. حجم دانشمندان نامداری که به نوعی با اثبات این قانون درگیر بوده‌اند، به قدری زیاد است که کنت آیرلند و مایکل روزن با طنازی ظریفی، در کتاب A Classical Introduction to Modern Number Theory چنین تمرینی را مطرح می‌کنند:

> تعداد اثبات‌های قانون تقابل مربعی را که تاکنون در این کتاب[۳] ارائه شده بشمارید و اثبات دیگری از آن ارائه دهید.

اما هدف از این نوشتار نه پشت هم نوشتن تمامی اثبات‌هاست و نه ارائه‌ی یک اثبات جدید! خوشبختانه آدرس تمامی این اثبات‌ها را در کتاب‌ها و سایت‌های مختلف می‌توان یافت و خواننده‌ی علاقه‌مند می‌تواند به آن‌ها مراجعه نماید[۲]. تاریخ قانون تقابل مربعی را می‌توان از نقطه‌نظرهای دیگری مورد بررسی قرار داد. در این نوشته، بیشتر بنابر به دست آوردن یک سیر تحول کلی در روند پیشرفت این مساله است؛ این‌که چه شد چنین مساله‌ای پدید آمد، تلاش برای اثبات آن در چه جهاتی انجام گرفت و هر کدام از ریاضی‌دانان بزرگ چگونه با آن مواجه شدند. بر همین اساس بیشتر نگاهی تطبیقی مد نظر است و نگاه تاثیرگذارترین ریاضی‌دانان، و هم‌چنین چگونگی تلاش آن‌ها برای اثبات این قانون، بررسی خواهد شد.

۲ صورت مساله

یکی از جالب‌ترین نکات درباره‌ی قانون تقابل مربعی، نحوه‌ی بیان این قانون است. اگر به مقاله‌های اویلر[۱] ، لژاندر[۲] و گاوس[۳] ، سه نفر از تاثیرگذارترین ریاضی‌دانان در تاریخ این قانون، نگاهی بیندازیم، متوجه خواهیم شد که هر کدام به شیوه‌ی خود این مساله را بیان کرده‌اند. در انتها، با پشتوانه‌ی قوی‌تری به بررسی این تفاوت‌ها خواهیم پرداخت، اما برای ورود به مبحث بد نیست با صورت امروزی قانون تقابل مربعی که در کتب مرجع نظریه‌ی اعداد به کار می‌رود، شروع کنیم:

فرض کنید p و q دو عدد اول فرد و متمایز باشند. در این صورت

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

طبیعی‌ست خواننده‌ای که تاکنون با این قانون آشنایی نداشته، ممکن است در خواندن آن نیز به مشکل برخورد کند ($\left(\frac{p}{q}\right)$ نماد لژاندر است که در ادامه به آن پرداخته می‌شود). بر همین اساس، در ادامه تاریخچه‌ای از بارقه‌های اولیه تا اثباتِ قانون تقابل مربعی را بیان خواهیم کرد.

۳ فرما

به عنوان اولین قدم در بررسی روند تکامل قانون تقابل مربعی، می‌توان از کارهای پی‌یر دو فرما[۴] شروع کرد. پی‌یر دو فرما(۱۶۰۱–۱۶۶۵)، در شهر کوچکی نزدیک تولوز و در جنوب فرانسه متولد شد. تحصیلات دانشگاهی خود را سال ۱۶۲۳ در دانشگاه اورلئان آغاز کرد و سال ۱۶۲۶ توانست در رشته‌ی «حقوق مدنی» اخذ مدرک کند و در سال ۱۶۳۱ توانست به عضویت «پارلمان» دادگاه عالی ایالتی درآید. اما چه شد که فرمای حقوق‌دان، به ریاضیات رو آورد، تا بدانجا که امروزه او را بیشتر به یک ریاضی‌دان می‌شناسیم؟ درباره‌ی این‌که دقیقاً چه زمانی فرما به ریاضیات علاقه‌مند شد، اطلاعاتی در دست نیست. گفته می‌شود که در آخرین سال‌های اقامتش در شهر بوردو، آشنایی او با یک دادرس موجب علاقه‌مندی او به ریاضیات شد. او در واقع اولین فردی بود که استعداد ریاضی فرما را شناخت و او را به فراگیری ریاضیات تشویق کرد و تعدادی از آثار چاپ نشده‌ی ویت[۵] را در اختیار فرما قرار داد.

[۱] Euler
[۲] Legendre
[۳] Gauss
[۴] Pierre de Fermat
[۵] Viete

پس از آن به تدریج علاقه‌مندی فرما به ریاضیات افزایش یافت و کم‌کم شروع به نامه‌نگاری با گروهی از ریاضی‌دانان کرد که یکی از مهم‌ترین آن‌ها، مرسن [۶] بود. نکته‌ی جالب آن‌که، فرما که ریاضی‌دانی حرفه‌ای نبود و بطور مستقیم با ریاضی‌دانان ارتباطی نداشت، بیشتر دلگرمی خود را از این نامه‌نگاری‌ها می‌گرفت. وی با وجود حرفه‌ای نبودن، تاثیرات به‌سزایی در حوزه‌ی نظریه اعداد گذاشت تا آن‌جا که می‌توان او را پدر نظریه اعداد مدرن به حساب آورد. وی معادله‌ی پل[۷] و اعدادی را که امروزه آن‌ها را به اعداد فرمایی می‌شناسیم، مورد مطالعه قرار داد و «قضیه‌ی کوچک» خود را پیشرفت داد. هم‌چنین وی مسائل مهمی درباره‌ی نوشتن اعداد به صورت مجموع ۲ و یا ۴ مربع را مطرح کرد که این مسائل زمینه‌ساز تحقیقات گسترده‌ای در آینده شد. هم‌چنین «قضیه‌ی آخر» وی که یکی از مهم‌ترین عوامل شهرت اوست، سال‌ها مورد بحث و مطالعه‌ی ریاضی‌دانان بود و تا دو دهه قبل به اثبات نرسیده بود.

همانند بسیاری از پیشرفت‌های امروز نظریه‌ی اعداد، قانون تقابل هم نتیجه‌ی تفکر بر روی مسائل مطرح شده توسط فرما است؛ این‌در حالیست که فرما خود از این قانون مطلع نبود. در این زمینه نخستین مسائل راجع به نوشتن اعداد اول به صورت مجموع دو مربع است که در یکی از نامه‌های وی به مرسن در سال ۱۶۴۰ دیده می‌شود. مسائل دیگر مانند نوشتن اعداد اول به صورت x^2+2y^2 و x^2+3y^2 ، اولین‌بار در نامه‌ی وی به پاسکال در سال ۱۶۵۴ مطرح شده است. نکته جالب مربوط به این نامه‌ها این است که فرما با وجود این‌که هیچ گونه اثباتی از نتایج خود ارائه نمی‌دهد، آن‌ها را قضیه می‌نامد.

او در نامه‌ای در سال ۱۶۵۸ به دیگبی[۸] می‌نویسد:

> هر عدد اول که از مضرب چهار، یکی بیشتر باشد، از دو مربع ساخته می‌شود. مثال‌ها چنین‌اند: ۵، ۱۳، ۱۷، ۲۹، ۳۷، ۴۱ و غیره.
>
> هر عدد اول که از مضرب سه، یکی بیشتر باشد از یک مربع و سه برابر مربع دیگری ساخته می‌شود. مثال‌ها چنین‌اند: ۷، ۱۳، ۱۹، ۳۱، ۳۷، ۴۳ و غیره.
>
> هر عدد اول که از مضرب هشت، یکی یا سه تا بیشتر باشد از یک مربع و دو برابر مربع دیگری ساخته می‌شود. مثال‌ها چنین‌اند: ۳، ۱۱، ۱۷، ۱۹، ۴۱، ۴۳ و غیره.

فرما اضافه می‌کند اثبات‌های محکمی برای ادعاهای فوق دارد، اما بعید به نظر می‌رسد این اثبات‌ها، در زبان امروزی ریاضیات چندان پذیرفته باشند. در حقیقت نگاه ریاضی‌دانان به مفهوم اثبات در طول تاریخ دستخوش تغییراتی شده و به عنوان نمونه، آن چیزی که فرما به عنوان اثبات به حساب آورده، با مفهوم دیگر ریاضی‌دانان متفاوت است. در انتهای این بخش، به عنوان حسن‌ختام یکی از این "اثبات‌ها" را خواهیم آورد.

کار فرما اما، به این مرحله محدود نشد. برای مثال، او شرط لازم و کافی برای این‌که یک عدد دلخواه(نه لزوماً اول) را بتوان به صورت جمع دو مربع نوشت، به دست آورد. هم‌چنین تعداد راه‌های نمایش یک عدد به صورت مجموع دو مربع را می‌دانست. علاوه بر این، فرما فرم‌های دیگری به غیر از x^2+y^2، x^2+2y^2 و x^2+3y^2 را نیز بررسی می‌کرد. او در سال ۱۶۵۸ در نامه‌ای به دیگبی، حدس‌هایی درباره‌ی فرم x^2+5y^2 مطرح می‌کند اما اقرار می‌کند که نمی‌تواند آن را اثبات کند. در واقع، تحقیق بر روی این مسائل بود که اویلر را به سمت تقابل مربعی هدایت کرد و زمینه‌ساز تحقیقات مهمی در نظریه‌ی اعداد شد. اکنون به عنوان حسن‌ختام، این بخش را با یکی از اثبات‌های فرما به پایان می‌بریم و در ادامه‌ی آن به سوالاتی که از دل این نوع نگرش خارج می‌شود، اشاره خواهیم کرد. در این اثبات روشی که فرما از آن استفاده می‌کند، همان روشی‌ست که امروزه آن را به نزول نامتناهی می‌شناسیم. اثبات زیر توسط فرما، به منظور اثبات این قضیه که اعداد اول به صورت $4k+1$ را می‌توان به صورت جمع دو مربع نوشت، ارائه شد:

> اگر یک عدد اول دلخواه انتخاب شود که از مضرب چهار یکی بیشتر باشد و جمع دو مربع نباشد، در اینصورت عدد اولی از همان فرم، کوچک‌تر از این عدد هست و باز هم عدد سومی کوچک‌تر از آن موجود است و به همین شکل تعداد زیادی نزول انجام دهید تا به عدد ۵ برسید، که کوچک‌ترین عدد از همه‌ی این نوع اعدادی‌ست که جمع دو مربع نیست. در این مرحله به وسیله‌ی غیر ممکن بودن کاهش، هر کسی باید استنتاج کند که نتیجتاً همه‌ی این اعداد از فرم از دو مربع ساخته می‌شوند.

همان‌طور که می‌بینیم، در این اثبات از روش نزول استفاده می‌شود، اما فرما توضیحی درباره‌ی چگونگی روش نزول در این مساله نمی‌دهد. در واقع، فرما تنها یک اثبات کامل به زبان امروزی ریاضیات دارد که مربوط به یادداشت‌های حاشیه‌ای اوست که اثبات این قضیه است که مساحت مثلثی قائم‌الزاویه با اضلاع صحیح، نمی‌تواند مربع کامل باشد.

سوالی که در این‌جا پیش می‌آید این است که آیا ممکن است فرما ادعاهای بی‌اساسی مطرح کرده باشد؟ آیا امکان دارد وی در عین

[۶] Mersenne
[۷] Pell's equation
[۸] Digby

نداشتن اثبات برای بسیاری از گزاره‌های خود، ادعا کند آن‌ها را اثبات کرده است؟

با توجه به بررسی‌های صورت گرفته، در نظر گرفتن چنین فرضیه‌ای چندان معقول به نظر نمی‌رسد، زیرا بسیاری از قضایای فرما، بعدها توسط اویلر، با استفاده از ایده‌های فرما اثبات شده است. آندره ویل[۹] در کتاب خود[۵] مطالعه‌ی دقیقی بر روی نامه‌ها و نوشته‌های حاشیه‌ای فرما انجام داده و با استفاده از اشاره‌های اویلر، برخی اثبات‌های فرما را بازنویسی کرده است. هم‌چنین فرما در رابطه با برخی از حدس‌های خود بیان کرده که موفق به ارائه‌ی اثباتی برای آن‌ها نشده است. به طور مثال وی درباره‌ی حدس خود مبنی بر اول بودن اعداد به فرم $2^{2^n}+1$ ، گفته است که نتوانسته آن را اثبات کند و در این نمونه، حدس مورد نظر بعدها رد شده است. در حقیقت با بررسی این شواهد، می‌توان چنین استدلال کرد که احتمالاً فرما دلایل قانع‌کننده‌ای برای بسیاری از حدس‌های خود داشته است.

۴ اویلر

اویلر برای اولین بار توسط نامه‌های گلدباخ[۱۰]، با نتیجه‌های فرما آشنا شد. اولین نامه‌ی گلدباخ به اویلر، در سال ۱۷۲۹ نوشته شده که او در آن‌ها، حدس فرما درباره‌ی اول بودن اعداد به فرم $2^{2^n}+1$ را مطرح کرده است. پس از آن اویلر به کارهای فرما علاقه‌مند شد و تعدادی از آن‌ها را مورد مطالعه قرار داد و پس از مدتی تحت تاثیر نتایج فرما قرار گرفت و در پی تکمیل و اثبات آن‌ها برآمد، که پایه‌ی بسیاری از تحقیقات او گردید.

اویلر اولین مقاله‌ی خود در نظریه اعداد را در سن ۲۵ سالگی در سال ۱۷۳۲ نوشت. او در این مقاله حدس فرما را رد کرد(نشان داد $2^{32}+1$ بر ۶۴۱ بخش‌پذیر است). از جمله کارهایی که اویلر از فرما مطالعه کرد، نامه او به دیگبی است که در فصل قبل به آن اشاره شد. اویلر ابتدا سراغ اثبات این قضیه از فرما رفت :

> هر عدد اول که باقی‌مانده‌اش بر ۴، ۱ باشد، جمع دو مربع کامل است.

او موفق به ارائه اثبات کاملی از این قضیه شد. اثبات او را می‌توان به دو بخش تقسیم کرد. در بخش اول او ثابت می‌کند که اگر p چنین شرطی داشته باشد، $x,y \in \mathbb{Z}$ که $(x,y)=1$ موجود است که $p|x^2+y^2$ و در بخش دوم ثابت می‌کند که اگر $p|x^2+y^2$ و $(x,y)=1$ در این صورت p خود جمع دو مربع است. برای بخش دوم او شبیه فرما عمل می‌کند. اگر عدد اول p موجود باشد که $p|x^2+y^2$ و خود جمع دو مربع نباشد، نشان می‌دهد عدد اولی کوچک‌تر موجود است که x^2+y^2 را عاد کند و خود جمع دو مربع نباشد. بخش دیگر اثبات آن بخشی است که به موضوع ما مربوط می‌شود. اویلر باید نشان می‌داد که اگر باقی‌مانده‌ی p بر ۴ برابر ۱ باشد، آن‌گاه p، جمع دو مربع را عاد می‌کند. همین‌طور درباره‌ی حدس دیگر فرما باید نشان می‌داد اگر p، به صورت $8k+1$ یا $8k+3$ باشد، عددی به صورت x^2+2y^2 ($(x,y)=1$) و اگر باقی‌مانده‌ی p بر ۳ برابر ۱ باشد، آن‌گاه عددی به صورت x^2+3y^2 ($(x,y)=1$) را عاد می‌کند. توجه کنید که اگر اویلر این احکام را ثابت می‌کرد، با روش‌های مشابه قضیه‌ی جمع دو مربع، می‌توانست قضایای فرما را اثبات کند. همین مسائل اویلر را به این سمت هدایت کرد که عوامل اول اعداد به صورت x^2+ny^2 ($(x,y)=1$) را برای مقادیر مختلف n مورد مطالعه قرار دهد و در پی شرط لازم و کافی برای این‌که عددی اول، عددی به شکل x^2+ny^2 ($(x,y)=1$) را عاد کند، باشد. به زبان ریاضیات امروزی، اویلر دنبال اعداد اول p بود که $-n$ مانده‌ی مربعی به هنگ p باشد.

اولین نتیجه‌ای از اویلر که به طور مستقیم با قانون تقابل مربعی در ارتباط است، گزاره‌ای است که امروزه به محک اویلر مشهور است و در کتب امروزی این چنین نوشته می‌شود:

> فرض کنید $a \in \mathbb{Z}$ و p عددی اول است به طوری که $p \nmid a$. در این صورت a مانده‌ی مربعی است، اگر و تنها اگر $a^{\frac{p-1}{2}} \overset{p}{\equiv} 1$ ، و a نامانده‌ی مربعی است اگر و تنها اگر $a^{\frac{p-1}{2}} \overset{p}{\equiv} -1$ (a مانده‌ی مربعی است زمانی که معادله $x^2 \overset{p}{\equiv} a$ جواب داشته باشد.)

اما آیا اویلر به صورت فوق محک خود را بیان کرد؟ یقیناً پاسخ منفی است؛ زیرا همان‌طور که می‌دانیم، تعریف نماد همنهشتی یکی از کارهای گاوس است و قاعدتاً اویلر می‌بایست به گونه‌ی دیگری این مساله را بیان کرده باشد. به علاوه، تعریف مانده و نامانده، خود در ادبیات اویلر دستخوش تغییراتی شد که در ادامه بیان خواهیم کرد. اولین بیان او در این راستا، که امروزه به «محک اویلر» شناخته می‌شود در مقاله‌ی $numerorum\ theoremata\ circa\ divisores$ (مقاله‌ی ۱۳۴ اویلر) مطرح شده است. با مطالعه‌ی این مقاله، در نتیجه‌ی ۴ام قضیه ۱۱، چنین آمده است:

> بنابراین با بررسی مقادیری برای a که a^m-1 بر عدد اول $2m+1$ بخش‌پذیر باشد، مانده‌هایی که بعد از تقسیم هر عدد مربع بر $2m+1$ باقی می‌مانند، بدست می‌آید. به این

[۹] André Weil
[۱۰] Goldbach

دلیل که اگر r باقی‌مانده‌ای از این فرم باشد، در این صورت $(۲m+۱)p+r$ یک مقدار مناسب برای a است.

اثبات‌های رایج محک اویلر از دوری بودن ساختار $\mathbb{Z}/p\mathbb{Z}$ استفاده می‌کند؛ گزاره‌ای که لامبرت[۱۱] آن را حدس زده بود و اویلر نتوانست آن را تا سال ۱۷۷۲ اثبات کند. یک نکته‌ی حائز اهمیت دیگر، بررسی روند تحول مفاهیم مانده و نامانده در ادبیات اویلر است. در این مورد، کاکس در کتاب خود [۱۹] چنین آورده است:

> در سال ۱۷۴۴ او می‌نویسد «عوامل اول اعداد به شکل $aa-Nbb$»؛ در ۱۷۴۷ این بیان به «باقیمانده‌های ایجاد شده از تقسیم مربع‌ها بر عدد اول p» تبدیل می‌گردد؛ و در سال ۱۷۵۱ این تحول کامل می‌گردد. از این‌جا اویلر آزادانه از اصطلاحات مانده ($residua$) و نامانده ($non-residua$)، استفاده کرد.

لازم به ذکر است که اینجا a و b نسبت به هم اول در نظر گرفته شده‌اند. در پی همین تلاش‌ها، اویلر در ۲۸ آگوست سال ۱۷۴۲ در نامه‌ای به گلدباخ تعدادی از نتایج به دست آمده توسط خود را می‌نویسد. نتایجی که او در سال ۱۷۴۴ در مقاله‌ای به نام $(E۱۶۴) Theoremata \ldots forma\ paa \pm qbb\ contentorum$ به معنی قضایایی درباره مقسوم علیه‌های اعداد به شکل $paa \pm qbb$ منتشر کرد.

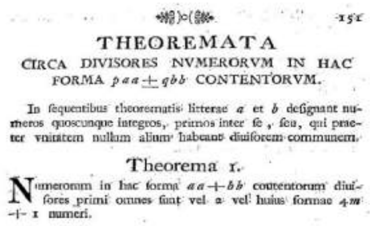
این مقاله ساختار متفاوتی از تمام مقالاتی که اویلر در طول عمر خود منتشر کرد داشت و در آن هیچ اثباتی نمی‌توان یافت. در عوض لیست بلندی از ۵۹ قضیه (بدون اثبات) و ۱۷ یادداشت در آن دیده می‌شود. اویلر مثال‌ها وحدس‌های خود را "قضیه" می‌نامد، اما می‌توان گفت منظور او از قضیه مفهوم مدرن آن نیست، بلکه گزاره‌هایی است که به درستی آن‌ها اطمینان دارد و برای موارد متعدد و بزرگی درستی آن‌ها را امتحان کرده است. در واقع این‌ها نتایجی به دست آمده از یک رویکرد تجربی هستند و به معنای امروزی هنوز "قضیه" نیستند.

برای پیش بردن اهدافمان باید درباره‌ی این مقاله‌ی مهم بیشتر توضیح دهیم. شاید مشاهده‌ی اولین صفحه‌ی این مقاله جالب باشد؛ اولین قضیه‌ی این مقاله در این تصویر قابل مشاهده است و معنای آن چنین است:

قضیه ۱ . *تمام عوامل اول اعداد به صورت $aa+bb$، ۲ یا اعداد به صورت $۴m+۱$ هستند.*

قضیه ۲. *تمام اعداد اول به فرم $۴m+۱$، به شکل $aa+bb$ هستند.*

[۱۱] Lambert


·151

THEOREMATA

CIRCA DIVISORES NVMERORVM IN HAC FORMA $paa \pm qbb$ CONTENTORVM.

In ſequentibus theorematis litterae a et b deſignant numeros quoscunque integros, primos inter ſe, ſeu, qui praeter vnitatem nullum alium habeant diuiſorem communem.

Theorema 1.

Numerorum in hac forma $aa+bb$ contentorum diuiſores primi omnes ſunt vel 2 vel huius formae $4m+1$ numeri.

قضیه ۳. *بنابراین مجموع دو مربع‌ها، که اعداد به صورت $aa+bb$ هستند، قابل تقسیم بر هیچ عددی به صورت $۴m-۱$ نیستند.*

بعد از این قضیه، اویلر با سه قضیه، با ساختار مشابه به سراغ صورت $a^۲+۲b^۲$ می‌رود و همین روند را در این مقاله، برای صورت‌های $a^۲+۳b^۲$ ، $a^۲+۵b^۲$ ، $a^۲+۷b^۲$ و ... (برای بیش‌تر صورت‌ها، سه قضیه به شکل مشابه) دنبال می‌کند. البته، همه‌ی قضایا مثل قضیه‌ی اول صورت کوتاهی ندارند. به طور مثال قضیه ۱۹ به صورت زیر است:

قضیه ۱۹. *همه‌ی عوامل اول اعداد به صورت $aa+۱۳bb$، ۲ یا ۱۳ یا به شکل یکی از ۱۲ فرمول زیر هستند.*

$۵۲m+۱$	$۵۲m+۷$
$۵۲m+۴۹$	$۵۲m+۳۱$
$۵۲m+۹$	$۵۲m+۱۱$
$۵۲m+۲۵$	$۵۲m+۱۹$
$۵۲m+۲۹$	$۵۲m+۴۷$
$۵۲m+۱۷$	$۵۲m+۱۵$

پس از بیان ۴۰ قضیه درباره‌ی این فرم‌ها، اویلر تعدادی یادداشت می‌آورد و پس از آن‌ها فرم‌هایی به شکل $a^۲-nb^۲$ را بررسی می‌کند. مثل فرم‌های $a^۲-b^۲, a^۲-۲b^۲, a^۲-۳b^۲$ و ...

در نهایت پس از ۵۰–امین قضیه، اویلر مقاله را با ۵ یادداشت مهم تمام می‌کند(یادداشت‌های ۱۳ تا ۱۷). بخشی از یادداشت ۱۳ چنین است:

> بنابراین تمام عوامل اول اعداد به فرم $aa-Nbb$، ۲ یا مقسوم‌علیهی از N و یا به فرم $۴Nm \pm \alpha$ هستند. اگر $۴Nm+\alpha$ یک فرم از عوامل باشد، در این صورت $۴Nm-\alpha$ هم یک فرم از عوامل خواهد بود.

هم‌چنین او در بخشی از یادداشت ۱۴ چنین می‌نویسد:

> در واقع از بحث روی عوامل اول، در بین مقادیر برای α، عدد زوج و مقسوم‌علیه‌ای از N وجود ندارد. به علاوه

> واضح است که همه مقادیر α، می‌توانند طوری اختیار شوند که هر کدام از $۲N$ کم‌تر باشند. برای زمانی که $۴Nm+۲N+b$ یک مقسوم‌علیه باشد، آن‌گاه با قرار دادن $m-۱$ به جای m، $۴Nm-(۲N-b)$ یک مقسوم‌علیه خواهد بود. از تمام اعداد فرد اول نسبت به N و کم‌تر از $۲N$، دقیقاً نیمی، مقادیر مناسب برای α خواهد داد.

هم‌چنین ابتدای یادداشت ۱۶ چنین است:

> به علاوه از آن‌جا که واحد همیشه در بین مقادیر α ظاهر می‌شود، پس برای همه اعداد مربعی که نسبت به $۴N$ اولند، یک مقدار مناسب برای α ایجاد می‌شود.

این سه بخش از این یادداشت‌ها، حاوی اطلاعات مهم و عمیقی هستند و می‌توان گفت این سه گزاره در کنار هم، به طور کامل معادل با قانون تقابل مربعی‌اند. با اطلاعات امروزی فهمیدن این موضوع چندان سخت نیست. در واقع با دانستن نظریه مقدماتی اعداد می‌توان نشان داد که قانون تقابل مربعی این ۳ گزاره را نتیجه می‌دهد. برعکس اگر N را عددی اول و فرد بگیریم، این سه گزاره به زبان امروزی نتیجه می‌دهد که N مانده‌ی مربعی به هنگ عدد اول فرد p است $(p \neq N)$ که $p \equiv \pm\alpha(۴N)$ اما تعداد باقی‌مانده‌های $\pm\alpha$ در واقع $\frac{\phi(۴N)}{۲} = N-۱$ است. از طرفی طبق یادداشت ۱۶ اگر p به هنگ $۴N$ با $\pm\beta^۲$ $((\beta,۲N)=۱)$ همنهشت باشد، آن‌گاه N به هنگ p مانده است، اما به سادگی می‌توانید نشان دهید اعداد $\pm۱^۲, \pm۳^۲, \ldots, \pm(N-۲)^۲$ به هنگ $۴N$ متمایزند و می‌دانیم تعدادشان $N-۱ = ۲\times\frac{N-۱}{۲}$ است. پس در واقع تمام مقادیر مجاز α همین اعداد هستند . در نتیجه به گزاره‌ی زیر می‌رسیم که امروزه از معادل‌های معروف قانون تقابل به شمار می‌رود:

> فرض کنید Nو P دو عدد اول فرد و متمایز باشند. در این صورت N به هنگ P مانده‌ی مربعی است، اگر و تنها اگر $\beta \in \mathbb{Z}$ موجود باشد که $P \overset{۴N}{\equiv} \pm\beta^۲$.

این نقطه از تاریخ نظریه‌ی اعداد را می‌توان نقطه‌ی کشف قانون تقابل مربعی به حساب آورد. همان‌طور که کرونکر[12] در سال ۱۸۷۵، این یادداشت‌ها در مقاله‌ی اویلر را اولین حکم معادل قانون تقابل مربعی می‌خواند. البته در این‌جا می‌توان سوالی مطرح کرد که چرا اویلر، خود در این مقاله قانون تقابل مربعی را به صورت صریح بیان نکرد؟ پاسخ دشوار نیست. همان‌طور که آندره ویل اشاره می‌کند هدف اویلر بررسی برای حالتی که N اول باشد نبود، بلکه سعی داشت مقادیر α را برای هر N، مشخص کند.

[12] Kronecker

جوزف لاگرانژ[13] یکی از ریاضی‌دان‌های آن دوره (متولد ۱۷۳۶) است که کارهای برجسته‌ای در حوزه‌ی نظریه‌ی اعداد انجام داده است. اگر چه لاگرانژ در روند پیشرفت قانون تقابل مربعی چندان موثر نبود، اما آشنایی او با این تحقیقات اویلر و کار کردن روی آن‌ها بین سال‌های ۱۷۷۳ تا ۱۷۷۵ الهام‌بخش اویلر شد، به گونه‌ای که باعث شد او کار در این حوزه خاص را ادامه دهد و بالاخره در مقاله‌ای به نام $observations\ circa \ldots primos$ (مقاله‌ی ۵۵۲ اویلر) قانون تقابل مربعی را به صورت صریح بیان کند. البته تاریخ انتشار این مقاله مربوط به سال ۱۷۸۳ است که پس از مرگ اویلر است. ذکر این نکته لازم است که اویلر نتوانست در زمان حیات خود اثباتی برای این قانون پیدا کند، اگرچه در حالات خاصی مثل مسائل درباره‌ی عوامل اول اعداد به فرم $x^۲+۳y^۲$ (به زبان امروزی اعداد اولی که ۳−به هنگ آن‌ها مانده‌ی مربعی است.) موفق به ارائه‌ی اثبات نشد.

۵ لژاندر

لژاندر از چند جهت در تاریخچه‌ی قانون تقابل مربعی نقش مهمی داشته است. نخست به خاطر نمادهایی که استفاده کرده که امروزه کاربرد زیادی دارد و به نماد لژاندر معروف است، و دوم به این علت که یک صورت‌بندی از قانون تقابل ارائه داد که امروزه متداول‌ترین صورت قانون تقابل محسوب می‌شود. همین‌طور می‌توان گفت او اولین کسی است که اثباتی کامل برای تعدادی از حالات (نه مثال خاص؛ توجه کنید که هر حالت شامل بی‌شمار مثال است.) ارائه کرده است.

کار لژاندر روی این قانون از یکی از مشهورترین مقالات او به نام $Reacherches\ d'analyse\ indéterminée$ شروع می‌شود. او در این مقاله ابتدا محک اویلر را ثابت می‌کند. هم‌چنین او تلاشی برای خلاصه نوشتن این محک انجام می‌دهد. صورت این محک بدین شکل بود که عامل اول فرد c، فرمول $x^۲+dy^۲$ $((x,y)=۱)$ را عاد می‌کند اگر و تنها اگر $(-d)^{\frac{c-۱}{۲}}$ در تقسیم بر c، باقی‌مانده ۱ داشته باشد. لژاندر در ادامه‌ی مقاله، این نتیجه را به صورت $(-d)^{\frac{c-۱}{۲}}=۱$ یا $(-d)^{\frac{c-۱}{۲}}=-۱$ می‌نویسد و توضیح می‌دهد که این رابطه «با از قلم انداختن مضارب c» درست است. نکته‌ی جالب این‌جاست که لژاندر برای راحتی نوشتار خود سعی دارد از همان مفهومی که امروزه آن را به عنوان همنهشتی می‌شناسیم استفاده کند و در واقع تساوی‌های او همنهشتی است، ولی در آن زمان هنوز مفهوم همنهشتی تعریف نشده بود. پس از این قرارداد، لژاندر بیان خود را از قانون تقابل ارائه می‌دهد. برای این کار a و A را دو عدد اول متمایز به فرم $۴x+۱$

[13] Lagrange

و b و B را دو عدد اول متمایز از فرم $4x-1$ می‌گیرد و قانون تقابل مربعی را در ۸ قضیه به صورت زیر بیان می‌کند:

قضیه ۱. *اگر* $b^{\frac{a-1}{2}}=1$ *در این صورت* $a^{\frac{b-1}{2}}=1$.

قضیه ۲. *اگر* $a^{\frac{b-1}{2}}=-1$ *در این صورت* $b^{\frac{a-1}{2}}=-1$.

قضیه ۳. *اگر* $a^{\frac{A-1}{2}}=1$ *در این صورت* $A^{\frac{a-1}{2}}=1$.

قضیه ۴. *اگر* $a^{\frac{A-1}{2}}=-1$ *در این صورت* $A^{\frac{a-1}{2}}=-1$.

قضیه ۵. *اگر* $a^{\frac{b-1}{2}}=1$ *در این صورت* $b^{\frac{a-1}{2}}=1$.

قضیه ۶. *اگر* $b^{\frac{a-1}{2}}=-1$ *در این صورت* $a^{\frac{b-1}{2}}=-1$.

قضیه ۷. *اگر* $b^{\frac{B-1}{2}}=1$ *در این صورت* $B^{\frac{b-1}{2}}=-1$.

قضیه ۸. *اگر* $B^{\frac{b-1}{2}}=-1$ *در این صورت* $b^{\frac{B-1}{2}}=1$.

توجه کنید که قضایای ۱ و ۲، ۳ و ۴، ۵ و ۶ معادل هستند. پس در واقع ۵ حالت مختلف وجود دارد که هر کدام باید ثابت شود. قبل از توضیح درباره‌ی این‌که لژاندر چه تلاش‌هایی برای اثبات این قضایا انجام داد، لازم است توضیحاتی درباره‌ی صورت‌بندی لژاندر از قانون تقابل مربعی ارائه دهیم. در واقع درست است که لژاندر ابتدا قانون را به صورت فوق بیان کرد، اما بعدها توانست صورت‌های کوتاه‌تر و بهتری ارائه دهد. در کتاب نوشته شده توسط او در سال ۱۷۹۷ دو تغییر عمده نسبت به کارهای قبلی او دیده می‌شود. تغییر اول مربوط به آن چیزی است که امروزه به «نماد لژاندر» معروف است و اکنون در تمامی کتاب‌های درباره مانده‌ی مربعی دیده می‌شود. او در صفحه ۱۸۶ کتاب چنین می‌آورد:

> از آنجا که مقادیر مشابه با $N^{\frac{c-1}{2}}$ اغلب در تحقیقات ما واقع خواهد شد، ما به اختصار، $\left(\frac{N}{c}\right)$ را برای بیان باقیمانده‌ای که $N^{\frac{c-1}{2}}$ در تقسیم بر c می دهد به کار خواهیم برد، که بر اساس آن چه که مشاهده کرده‌ایم، تنها مقادیر ۱+ یا ۱− به خود می‌گیرد.

تغییر دوم معرفی اصطلاح «تقابل» است. او در جایی از کتاب درباره‌ی این قانون چنین می‌نویسد:

> «قضیه‌ای شامل قاعده‌ای از تقابل که بین دو عدد اول دلخواه وجود دارد.»

ما می‌دانیم که امروزه این قانون به «قانون تقابل مربعی» مشهور است. تغییر سوم نیز مربوط به صورت‌بندی قانون تقابل مربعی است. در صفحه ۲۱۴ کتاب، او می‌نویسد:

> اعداد اول m و n هر چه باشند اگر هر دو از فرم $4x-1$ نباشند، آن‌گاه $\left(\frac{n}{m}\right)=\left(\frac{m}{n}\right)$. و اگر هر دو از فرم $4x-1$ باشند آن‌گاه $\left(\frac{n}{m}\right)=-\left(\frac{m}{n}\right)$. که دو عبارت قبل ترکیب می‌شوند در فرمول $\left(\frac{n}{m}\right)=(-1)^{\frac{n-1}{2}\frac{m-1}{2}}\left(\frac{m}{n}\right)$.

و این همان صورتی از قانون تقابل است که امروزه در بسیاری از کتاب‌ها دیده می‌شود.

این بخش را با توضیحاتی درباره تلاش‌های لژاندر درباره اثبات قانون تقابل مربعی به پایان می‌بریم:

لژاندر این قانون را به ۸ قسمت تقسیم کرد. او سعی می‌کند هر قضیه را جداگانه ثابت کند. درباره قضیه ۱ و ۲ و ۷ او موفق می‌شود اثبات‌های کامل ارائه دهد اما اثبات او برای قضایای دیگر نیازمند فرض‌های دیگری بود که او نتوانست آن فرض‌ها را به طور کامل ثابت کند. سوالی که مطرح می‌شود این است که آیا اثبات لژاندر «به طرز مناسبی» قابل کامل شدن بود یا نه؟

برای پاسخ به این سوال ابتدا بایست توضیحاتی درباره اثبات او ارائه شود. ابزار اصلی او برای اثبات قانون تقابل مربعی، نتیجه‌ای است که خود آن را به دست آورد که امروزه به قضیه‌ی لژاندر معروف است:

> فرض کنید r، s و t خالی از مربع، صحیح مثبت و دو به دو نسبت به هم اول باشند. در این صورت معادله‌ی $rx^2+sy^2=tz^2$ جواب ناصفر در اعداد صحیح دارد، اگر و تنها اگر اعداد صحیح λ ، μ و ν موجود باشد به طوری که $\frac{t\nu^2-r}{s}$ و $\frac{t\mu^2-s}{r}$ و $\frac{r\lambda^2+s}{t}$ همه صحیح باشند.

او برای شروع اثبات نشان می‌دهد که این شرایط بر حسب نماد لژاندر به صورت زیر قابل نمایش است:

$$\left(\frac{rt}{s}\right)=1 \qquad \left(\frac{st}{r}\right)=1 \qquad \left(\frac{-rs}{t}\right)=1$$

هم‌چنین او از تعریف نماد خود نتیجه می‌گیرد که این نماد ضربی است و هم‌چنین $\left(\frac{-N}{a}\right)=\left(\frac{N}{a}\right)$ و $\left(\frac{-N}{b}\right)=-\left(\frac{N}{b}\right)$ زمانی که $\frac{a-1}{2}$ زوج و $\frac{b-1}{2}$ فرد باشد. این‌ها خواص مهمی از این نماد هستند که به وفور از آن‌ها استفاده می‌شود.

حال به اثبات او از قضیه ۱ توجه کنید:

فرض کنید $a\overset{4}{\equiv}1$ و $b\overset{4}{\equiv}3$ و $\left(\frac{b}{a}\right)=1$ اما $\left(\frac{a}{b}\right)=-1$. چون $b\overset{4}{\equiv}3$ پس $\left(\frac{-a}{b}\right)=1$. پس طبق قضیه‌ی لژاندر معادله‌ی $x^2+ay^2-bz^2=0$ جواب نابدیهی دارد. با حذف مقسوم‌علیه مشترک x و y و z از معادله، می‌توان فرض کرد یکی فرد است. اما $0\equiv x^2+ay^2-bz^2\overset{4}{\equiv}x^2+y^2+z^2$ نتیجه می‌دهد x و y و z همگی زوج‌اند که متناقض است.

پس او برای قضیه ۱ و ۲ اثبات کاملی دارد. اثبات‌های او به طور کلی با ایده‌ی مشابه فرض خلف و استفاده از قضیه‌ی لژاندر ارائه می‌شوند و بدین صورت ثابت می‌کنند که معادله جواب دارد و به تناقض می‌رسند. اما بر خلاف قضیه‌های ۱ و ۲ و ۷ باقی اثبات‌های او نیازمند فرض‌هایی بودند. برای مثال به اثبات قضیه ۸ توجه کنید:

> فرض کنید (پیمانه ۴) $b \equiv B \equiv 3$ به طوری که $\left(\frac{b}{B}\right) = \left(\frac{B}{b}\right) = -1$. حال فرض کنید می‌توانیم عدد اول $p \overset{4}{\equiv} 1$ را طوری بیابیم که $\left(\frac{p}{B}\right) = \left(\frac{p}{b}\right) = -1$. حال با استفاده از قضیه ۲ داریم که $\left(\frac{B}{p}\right) = \left(\frac{b}{p}\right) = -1$ پس معادله‌ی $Bx^2 + by^2 - py^2 = 0$ جواب نابدیهی دارد که مشابه قضیه ۱ می‌شود به تناقض رسید.

این‌جا فرض اضافه‌ای که نیاز بود، فرض وجود p اول است که هم $p \overset{4}{\equiv} 1$ و هم $\left(\frac{p}{B}\right) = \left(\frac{p}{b}\right) = -1$. در واقع چنین p باید به پیمانه‌ی $4Bb$ باقی‌مانده‌ای از بین تعدادی باقی‌مانده خاص داشته باشد. در واقع لژاندر نیازمند استفاده از گزاره زیر بود. همان‌طور که گاوس در کتاب «تجسسات حسابی» اشاره می‌کند که اثبات لژاندر می‌تواند به نوعی فرمول‌بندی شود که قانون تقابل مربعی را به عنوان نتیجه‌ای از این گزاره اثبات کند:

> اگر a و b دو عدد طبیعی و نسبت به هم اول باشند، عددی اول موجود است که به هنگ b برابر با a باشد.

گزاره‌ای که به هیچ عنوان نمی‌توان آن را بدیهی فرض کرد. در خیلی از تحقیقات افراد دیگری مثل گاوس نیز، استفاده از این گزاره می‌توانست کارها را راحت تر کند اما آن زمان این گزاره به عنوان یک حدس مطرح بود و هنوز ثابت نشده بود. گاوس دشوار بودن اثبات این گزاره را درک می‌کرد و به همین دلیل، در تحقیقات خویش از به کار بردن آن پرهیز می‌نمود. این گزاره بعدها توسط ریاضی‌دانی به نام دیریکله[۱۴]، که یکی از بزرگ‌ترین ریاضی‌دانان قرن ۱۹ است، اثبات شد و امروزه به قضیه‌ی دیریکله مشهور است.

لژاندر برای ادعای خود، یعنی وجود چنین عدد اولی، نمی‌توانست اثبات محکمی ارائه کند. همان‌طور که اشاره شد، به غیر از اثبات قضایای ۱و ۲ و ۷ بقیه اثبات‌های لژاندر شامل فرض‌هایی مشابه درباره وجود چنین اعداد اولی بود. در حقیقت خود او هم، نابدیهی بودن این فرض‌ها را باور داشت و در تلاش‌های بعدی، او سعی کرد قضیه ۸ را با استفاده از «معادله‌ی پل» و فرض دیگری درباره وجود عددی اول ثابت کند:

[۱۴] Dirichlet

لم لژاندر : برای هر عدد اول $a \overset{4}{\equiv} 1$، عدد اول $b \overset{4}{\equiv} 3$ موجود است که $\left(\frac{a}{b}\right) = -1$.

اما لژاندر برای این لم هم نتوانست اثبات محکمی ارائه دهد و اثبات‌های او ناقص باقی ماند. در واقع اثبات‌های لژاندر بدون استفاده از قضیه‌ی دیریکله قابل کامل شدن نبود. هم‌چنین لم لژاندر هم قابل ثابت شدن نیست مگر این‌که برای اثبات آن از خود قانون تقابل مربعی و قضیه‌ی دیریکله استفاده کرد. توجه کنید که قضیه‌ی دیریکله خود قضیه‌ای بسیار دشوارتر از قانون تقابل مربعی است و شاید استفاده از آن برای اثبات قانون تقابل چندان معقول نباشد. بر همین اساس بعید به نظر می‌رسد که کامل کردن اثبات لژاندر، «به طرز مناسبی» امکان‌پذیر بوده باشد و تاریخ نظریه‌ی اعداد نیازمند فرد دیگری بوده تا اثبات مناسبی برای این قانون ارائه دهد.

۶ گاوس

بدون تردید، صحبت از گاوس در این مقاله را، باید از کتاب *Disquisitiones Arithmeticae* (به فارسی «تجسسات حسابی») آغاز کرد. با نگاهی تاریخی می‌توان کتاب «تجسسات حسابی» را از جهات مختلفی حائز اهمیت شمرد. به هر صورت کتاب فوق، که در ۲۱ سالگی گاوس نگاشته شد و سال ۱۸۰۱ و در ۲۴ سالگی او به چاپ رسید، یکی از تاثیرگذارترین و انقلابی‌ترین آثار تاریخ علم حساب است و بسیاری از مهم‌ترین قوانین ریاضیات، اول بار در این کتاب طرح شده (به عنوان نمونه، اولین اثبات کامل قضیه‌ی اساسی حساب در این کتاب آورده شده است.) و هم‌چنین نمادگذاری‌های بسیار تاثیرگذاری، مانند همنهشتی، نیز برای اولین بار در این کتاب آمده است. دلیل برجسته‌ی دیگری که این کتاب را نه تنها به عنوان یک اثر ریاضی باارزش، بلکه به عنوان یک اثر تاریخی مشخص می‌کند، ساختار آن است؛ ساختار «قضیه/اثبات قضیه/نتیجه/مثال» اولین بار در کتاب گاوس بود که پیشنهاد شد و تاکنون نیز مورد استفاده‌ی تمام مولفین کتب ریاضی قرار گرفته است [۱۷]. اما در کنار این تاثیرگذاری تاریخی، این کتاب را می‌توان نقطه‌ی عطف تاریخچه‌ی قانون تقابل مربعی دانست.

اولین اثبات کامل قانون تقابل مربعی که به‌وسیله‌ی استقرا صورت می‌پذیرد، در این کتاب آورده شده است. اثبات اول قانون تقابل مربعی، شاید امروزه به نظر طولانی برسد و خواندن آن خارج از حوصله‌ی بسیاری از افراد باشد. در واقع نکته‌ی جالبی که درباره‌ی بسیاری از اولین اثبات‌های تاریخ ریاضی به چشم می‌خورد این است که معمولاً این اثبات‌ها بسیار طولانی‌اند یا فهم مشکلی دارند. یکی

از دلایل چنین اتفاقی را می‌توان چنین در نظر گرفت که زمان اثبات این قضایا، هنوز بسیاری از تئوری شکل نگرفته بود و اگر هم مطرح شده بوده، هنوز به تکامل امروزی نرسیده بود. به علاوه بسیاری از نمادگذاری‌ها هنگام اولین اثبات مفهومی نداشته است. در نمونه‌های بسیاری می‌توان چنین مشاهده کرد که اثبات‌های اولیه، در طی زمان کوتاه‌تر می‌شوند و اجزا و ساختار دقیق و منظم‌تری به خود می‌گیرند. در حقیقت، بخش بزرگی از روندی که اثبات‌های کوتاه، موجز و زیباتری در اختیار ما قرار می‌دهد، در هنگام ارائه‌ی اثبات اول تکمیل نگشته و همین امر موجب می‌شود که خواندن این اثبات‌ها کمی خسته‌کننده و گیج‌کننده باشند. به طور مثال اثبات یکی از مشهورترین قوانین نظریه‌ی اعداد به نام «قانون اعداد اول»، اولین بار در حجمی نظیر چند ده صفحه آمده است در حالی که امروزه اثبات‌هایی در حدود ۲ یا ۳ صفحه از آن در دسترس است. اثبات اول قانون تقابل نیز، امروزه توسط براون بازنویسی شده و به اثبات خواندنی و کوتاه‌تری تبدیل شده است.

در نخستین اثبات گاوس، علی‌رغم تفاوت‌های بسیار، شباهتی میان اثبات او و اثبات ناکامل لژاندر به چشم می‌خورد؛ زیرا هر دو اثبات نیازمند یک نتیجه‌ی اضافی‌ست که بر خلاف لژاندر، گاوس موفق به اثبات آن می‌شود:

قضیه ۱۲۹ از کتاب تجسسات حسابی : اگر a یک عدد اول از فرم $8n+1$ باشد، لزوما تعدادی عدد اول کم‌تر از $2\sqrt{a}+1$ موجود است که a به پیمانه‌ی آن‌ها یک نامانده است.

در این کتاب، گاوس تنها به همین یک اثبات بسنده نمی‌کند و در ادامه، اثبات دیگری با استفاده از فرم‌های مربعی ارائه می‌دهد. او در فصل ۵ کتاب به بیان فرم‌های مربعی دوتایی می‌پردازد و قضایای بسیاری راجع به آن‌ها بیان می‌کند و به همین جهت بخش اعظم کتاب، به این موضوع اختصاص یافته است. او پس از اثبات بسیاری از قضایای مربوط به فرم‌های مربعی دوتایی، نهایتاً اثباتی از قانون تقابل مربعی بوسیله‌ی آن ارائه می‌دهد.

گاوس هم‌چنین در این کتاب، از اثبات دیگری نیز سخن می‌گوید، اما بخش مربوط به این اثبات، پیش از انتشار به دلایلی نه چندان مشخص، حذف می‌شود. این اثبات‌ها که با استفاده از «معادلات تناوبی مربعی» صورت گرفته است، تا مرگ گاوس انتشار پیدا نکرد. به همین دلیل، با این‌که از نظر زمانی، این دو اثبات، اثبات‌های سوم و چهارم قانون تقابل مربعی هستند، اما با توجه به زمان انتشار، در کتب رسمی، آن‌ها را به اثبات‌های هفتم و هشتم گاوس می‌شناسند. اما اثبات‌های سوم تا ششم به چه طریق صورت گرفت؟

اثبات سوم گاوس، با استفاده از «لم گاوس» صورت گرفت که بیان آن به صورت زیر است:

لم گاوس: اگر $p > 2$ عددی اول باشد و a نسبت به p اول باشد. اعداد صحیح

$$a, 2a, \dots, \frac{p-1}{2}a$$

و باقی‌مانده‌ی تقسیم آن‌ها بر pرا در نظر بگیرید. اگر n را تعداد باقی‌مانده‌های بیشتر از $\frac{p}{2}$ در نظر بگیریم آن‌گاه

$$(\frac{a}{p}) = (-1)^n$$

او در اثبات پنجم خود نیز از این لم استفاده می‌کند. در واقع بیش‌تر اثبات‌هایی که در آینده نیز از قانون تقابل ارائه شد، با استفاده از همین لم بوده است(۴۶ اثبات از کلیه‌ی ۲۳۳ اثبات این قانون). اثبات چهارم و ششم با استفاده از مفهومی به عنوان «جمع گاوسی» صورت گرفت، که ایده‌ی کلیدی گاوس در اثبات تعمیم قانون تقابل درجات بالاتر بود.

اما سوالی که می‌کوشیم با کاوشی جزئی به آن نگاهی بیندازیم، درباره‌ی چرایی ارائه‌ی هشت اثبات مختلف توسط گاوس است. عملی که با استناد به مراجع معتبر، ریاضی‌دانان بسیاری را تحت تاثیر قرار داده است. در وهله‌ی اول، شاید بتوان گفت که آوردن چند اثبات کمک به استحکام قانون می‌نماید، اما بعید به نظر می‌رسد که گاوس مساله‌ی استحکام یک قضیه را در ارائه‌ی چند اثبات یافته باشد. وی در نامه‌ای به هاینریش ویلهلم ماتیاس اولبرز چنین می‌نویسد:

> مقصود من از اثبات، آن معنای حقوق‌دانان نیست که در نگاه‌شان، دو اثبات نیمه برابر یک اثبات کامل است، بلکه در نگاه یک ریاضی‌دان، اثبات نیمه=۰ است و اثباتی مورد قبول است که هر شبهه‌ای را غیرممکن سازد.

به همین دلیل چندان معقول به نظر نمی‌رسد که ریاضی‌دانی چون او، که هر کدام از اثبات‌هایش کامل است و خود به درستی آن‌ها واقف است و چنین تعریفی از اثبات ریاضی ارائه می‌دهد، اثبات‌های دیگری را به منظور استحکام‌بخشی و اطمینان از درستی حکم خود آورد. البته هم‌چنان می‌توان این دلیل را در نظر گرفت، اما همان‌طور که گفته شد، بعید است دلیل اصلی گاوس برای ارائه‌ی هشت اثبات کامل باشد.

دلیل دیگری که می‌تواند به ذهن بیاید، علاقه‌مندی به ارائه‌ی یک اثبات زیباست، زیرا او قانون تقابل را «قضیه‌ی طلایی» و «جواهر حساب متعالی» می‌نامید [۲۰،ص۲۰۰]. چنین استدلالی، با توجه به آن‌که هیچ نقل قولی از خود گاوس در این زمینه در اختیار نیست

و عملاً هیچ نوشته‌ای ما را از آن آگاه نمی‌سازد، به دلیل عدم ابطال‌پذیری، با علم به آن‌که ممکن است درست باشد، مورد نظر ما نیست. می‌توان از دو منظر علمی‌تر به این مساله نگاه کرد:

آن‌چه در بسیاری از کتب و مقالات راجع به قانون تقابل و چرایی ارائه‌ی چند اثبات توسط گاوس مطرح می‌شود، اشتیاق او در به دست آوردن قانون تقابل برای درجات بالاتر است. پیگیری ایده‌ی تعمیم، می‌تواند دلیل مناسبی برای اثبات‌های مختلف او باشد، زیرا گاوس در ادامه‌ی فعالیت‌های خود توانست بیان قانون تقابل درجه سوم و چهارم را ارائه دهد و برای آن‌ها نیز اثباتی ارائه دهد، اما پیش از او، شاگرد مورد علاقه‌ی او، گاتولد آیزنشتاین[15]، توانست اثبات این قوانین را به نام خود ثبت کند. در حقیقت توجه گاوس در ادامه‌ی فعالیت‌ها به تعمیم قانون تقابل به درجات بالاتر(او ذکر می‌کند که تقابل درجات بالاتر به مراتب سخت‌تر از تقابل مربعی است)، می‌تواند این گمان را تقویت کند که انگیزه‌ی او از اثبات‌های متعدد این قانون، ایده‌ی تعمیم است.

اما می‌توان، انگیزه‌ی گاوس را، از دریچه‌ای کمی متفاوت و هم‌چنین کلی‌تر از ایده‌ی تعمیم نگریست. یوری مانین[16] در مصاحبه‌ای، تصور خود را از انگیزه‌ی گاوس، چنین بیان می‌کند:

«وقتی خیلی جوان بودم، به شدت به این مساله علاقه‌مند شدم که گاوس هفت یا هشت اثبات برای قانون تقابل مربعی ارائه کرده است. آن‌چه مایه‌ی پریشانی من می‌شد آن بود که او به چه دلیل نیازمند هفت یا هشت اثبات بوده است. هر بار که درک وسیع‌تری از نظریه‌ی اعداد به دست می‌آوردم، ذهن گاوس را بیشتر درک می‌کردم. بی‌تردید او به دنبال یک اثبات قانع‌کننده‌تر نبود، زیرا یک اثبات نیز به اندازه‌ی کافی قانع‌کننده است. نکته این‌جا بود که اثبات راهی‌ست که ما قلمروهای جدید کشف می‌کنیم، ویژگی‌های جدیدی از چشم‌انداز ریاضیات.» گفته‌ی مانین از این نظر قابل توجه است که در انگیزه‌خوانی او، به علم ریاضیات و مساله‌ی «اثبات» کمی ژرف‌تر و کلی‌تر از ایده‌ی تعمیم نگاه می‌شود و در مدل او، تعمیم نیز جای می‌گیرد.

همان‌طور که در بخش‌های قبل صورت‌بندی ریاضی‌دانان پیش از گاوس از قانون تقابل مربعی را بیان کردیم، لازم است این‌جا صورت‌بندی گاوس را هم بیان کنیم. این بخش را با بیان این صورت‌بندی به پایان می‌بریم.

گاوس در کتاب «تجسسات حسابی»، ابتدا مثال‌هایی می‌آورد و بررسی می‌کند که ۳- و ۵+ و ۷- چه زمانی به هنگ عدد اول p، مانده مربعی هستند. او مشاهده می‌کند که این، زمانی اتفاق می‌افتد که دقیقاً p به هنگ آن‌ها مانده مربعی باشد. سپس او آماده است که قانون تقابل مربعی را به صورت مورد نظر خود بیان کند:

> اگر p یک عدد اول از فرم $4n+1$ باشد، $+p$ یک مانده یا نامانده از هر عدد اولی که مثبت و آن هم یک مانده یا نامانده از p باشد، خواهد بود. اگر p از فرم $4n+3$ باشد، $-p$ همان خواص را خواهد داشت.

به زبان امروزی این معادل است با

$$\left(\frac{(-1)^{\frac{p-1}{2}}p}{q}\right)=\left(\frac{q}{p}\right)$$

که این هم معادل است

$$(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)=\left(\frac{q}{p}\right)$$

که همان صورت امروزی قانون تقابل مربعی است.

۷ موخره

درباره‌ی قوانین تقابل چنین گفته می‌شود که «تاریخ قوانین تقابل همان تاریخ نظریه‌ی جبری اعداد است». در این مقاله اما، بیشتر با تاکید بر قانون تقابل مربعی، و نه مراتب بالاتر، به بررسی روند پیشرفت و تکامل آن پرداختیم. در حقیقت، در مقاله‌ی حاضر تلاش‌مان بر این بود که خط اولیه‌ی پیدایش مساله را بیابیم و سپس، چالش‌ها و بن‌بست‌های ریاضی‌دانان مختلف را تا رسیدن به یک اثبات کامل نشان دهیم. در نمایش داستان تحول یک ایده و مساله، شاید قانون تقابل یکی از بهترین مثال‌ها باشد؛ از این نظر که قانون تقابل در گستره‌ی تاریخ ریاضیات مدرن ریشه دوانده است و بر همین مبنا می‌توان آن را مدلی کامل از تحول ریاضیاتی در نظر گرفت.

همان‌طور که گفته شد، ماجرا از حدس‌های فرما آغاز شد. کارهای اویلر بر روی نتایج فرما و تکمیل و اثبات بسیاری از آنان، پیشرفت بزرگی در این سیر ایجاد کرد. اویلر در مقالات خود توانست اولین بیان قانون تقابل را ارائه دهد و در ادامه، لژاندر با یک نمادگذاری هوشمندانه، صورت قضیه را ساده و دقیق‌تر صورت‌بندی کرد و تلاش مهمی در راستای اثبات آن انجام داد، که به دلیل نرسیدن به یک نتیجه‌ی اضافی، از اثبات کامل آن باز ماند و نهایتاً گاوس توانست اثباتی از این قانون ارائه دهد. اولین اثبات گاوس، مانند سایر مدل‌های پیشروی در ریاضیات، نه تنها پایان کار نبود، بلکه دریچه‌های بسیاری میان ریاضی‌دانان گشود. گاوس توانست هشت اثبات از این قانون ارائه دهد و کارهای او منجر به تعمیم قضیه شد. اگر به بررسی ادامه‌ی روند قانون تقابل نیز نگاه بیندازیم، می‌توان تلاش بزرگ‌ترین ریاضی‌دانان، همچون آیزنشتاین، کومر[17]،

[15] Gotthold Eisenstein
[16] Yuri Manin
[17] Kummer

هیلبرت[۱۸]، آرتین[۱۹] و بسیاری دیگر را نظاره کرد. در آن دست تلاش‌ها می‌توان تاثیر پیشرفت ارتباطات و هم‌چنین آکادمی‌ها را نیز مورد بررسی قرار داد. اما آن‌چه در ادامه مورد بررسی قرار می‌دهیم بازه‌ی فعالیت‌های فرما تا گاوس است؛ بازه‌ای از تلاش‌های تاریخی که با دانش دبیرستانی از نظریه‌ی اعداد نیز قابل فهم است.

یک نکته‌ی بسیار جالب در روند پیشرفت قانون تقابل مربعی، آن است که دانشمندان به طور دقیق از فعالیت‌های یکدیگر اطلاعی نداشتند و بیشتر تلاش‌های هر کدام به گونه‌ای مستقل صورت گرفته است. به عنوان نمونه نه لژاندر و نه گاوس از صورت‌بندی اویلر از قانون تقابل مربعی اطلاعی نداشته‌اند. در کنار این نکته، همان‌طور که در این نوشته نشان داده شد، صورت‌بندی هر کدام از این ۳ دانشمند از قانون تقابل متفاوت است. اما در این روند چه تفاوتی میان دانشمندان و نوع نگاه آن‌ها به مساله به چشم می‌خورد؟

فرما و اویلر: در حقیقت این فرما بود که اویلر را به نظریه‌ی اعداد علاقه‌مند کرد، اما فرما علی‌رغم تاثیر بسیار در علم نظریه‌ی اعداد، به اثبات دقیق حدس‌های خود نرسید. همان‌طور که در بخش فرما اشاره شد، او بایست دلایل نسبتاً محکمی برای ادعاهای خود داشته باشد، اما نه دلایلی که در ریاضیات امروزی محکمه‌پسند باشد. شاید بتوان گفت مثالی که گاوس در نقل قول خود راجع به اثبات در علم حقوق می‌آورد بیشتر به کار فرما مشابهت دارد. دلایل فرما به‌صورت شهودی قابل قبول هستند، اما صورت‌بندی ریاضیاتی دقیقی ندارند. در بررسی رابطه‌ی اویلر و فرما، این نکته حائز اهمیت است که اویلر توانست بسیاری از قضایای فرما را اثبات نماید، به گونه‌ای که آن اثبات‌ها در ریاضیات امروزی نیز قابل استناد باشند. در حقیقت تفاوت میان این دو ریاضی‌دان را می‌توان ناشی از پیشینه‌ی ریاضیاتی هر کدام دانست؛ فرما هیچ‌گاه ریاضیات را به صورت حرفه‌ای و آکادمیک فرا نگرفت و شاید بر همین اساس باشد که قضایا و حدس‌های او نیازمند مکملی چون لئونارد اویلر بود. کسی که به طور حرفه‌ای ریاضیات فراگرفته و توانایی صورت‌بندی دقیق اثبات‌ها را داشت.

اویلر و لژاندر: همان‌طور که گفته شد، لژاندر از مقاله‌ی نهایی اویلر که در آن به قانون تقابل رسیده بود، خبر نداشت و در مسیر رسیدن به این قانون مستقل عمل کرده بود. یک اختلاف رویکرد بارز که میان لژاندر و اویلر به چشم می‌خورد این است که لژاندر در حل این مساله، ایده‌ی نمادگذاری را برمی‌گزیند، در حالی که اویلر، در عباراتی طولانی‌تر و پیچیده‌تر به بیان خود رسید. در حقیقت، استفاده‌ی لژاندر از نمادگذاری در حل مساله، نکته‌ای بود که به ساده‌تر شدن صورت‌بندی مساله کمک کرد و از فرمالیسم بیشتری یاری می‌جست.

[۱۸] Hilbert
[۱۹] Artin

نکته‌ای که در تفاوت نحوه‌ی فکر کردن این دو دانشمند قابل بررسی است.

لژاندر و گاوس: اولین نقد بر برهان لژاندر توسط گاوس بیان شده است. گاوس قضایای لژاندر را به ۵ حالت تقسیم می‌کند.

حالت ۱. *همان قضیه ۷ لژاندر است. گاوس نشان می‌دهد با استفاده از تکنیک‌های لژاندر می‌توان نتیجه گرفت که امکان ندارد bRB و BRb (برای به یاد آوردن شرایط B، A، b و a به بخش لژاندر مراجعه کنید). توجه کنید که منظور گاوس از pRq و pNq به ترتیب این است که p به هنگ q مانده و نامانده مربعی باشد.*

حالات دیگری که امکان ندارد رخ دهند:

حالت ۲. aNb و bRa *(قضیه ۱ و ۲ لژاندر)*.

حالت ۳. ANa و aRA *(قضیه ۳ و ۴ لژاندر)*.

حالت ۴. bNa و aRb *(قضیه ۵ و ۶ لژاندر)*.

حالت ۵. BNb و bNB *(قضیه ۸ لژاندر)*.

گاوس بعد از اذعان به این‌که حالت ۱ و ۲ «جای هیچ ایرادی ندارد» نشان می‌دهد که فرض‌های مربوط به حالات دیگر وابسته به قضیه‌ای درباره‌ی تصاعدهای حسابی (همان‌طور که در فصل لژاندر گفتیم این قضیه امروز به قضیه‌ی دیریکله مشهور است) است و از این طریق ایراد خود بر اثبات لژاندر را مطرح می‌کند.

«ریاضی‌دانان بر شانه‌های یکدیگر ایستاده‌اند». این جمله منتسب به کارل فریدریش گاوس است. در واقع، یکی از مهم‌ترین ثمرات مطالعه‌ی تاریخ ریاضیات، درک بهتر این گفته است و این‌که ریاضی‌دانان چگونه بر یکدیگر تاثیر می‌گذاشتند. در بررسی قانون تقابل مربعی نیز می‌توان تاثیر هر کدام از بزرگ‌ترین ریاضی‌دانان را بر یکدیگر و هم‌چنین بر موضوع، در یک سیر تاریخی گسترده، مشاهده کرد. شاید بد نباشد این بخش را با قسمتی از مقدمه‌ی کتاب «تجسسات حسابی» گاوس به پایان ببریم: «سهم بیشتر، متعلق نویسندگان مدرن است که از بین آن‌ها نویسندگان اندکی که مفتخر به افتخاری جاودانه بودند، نظیر فرما، اویلر، لاگرانژ، لژاندر(و تعداد اندک دیگری) راه ورود به معبد این دانش الهی را گشودند و غنای وافر درون آن را آشکار نمودند.»

۸ بد نیست انجام شود!

در تمام مدت بررسی سیر تکامل تاریخی قانون تقابل مربعی، سوالات بزرگی برای نگارندگان پیش می‌آمد که تعمیق در آن را لازمه‌ی

تحقیقات گسترده‌ی دیگری می‌دیدند و این بررسی دقیق در حوصله‌ی مطلب نمی‌گنجید. در انتها به عنوان کارهایی که بد نیست انجام شود، به چند مورد کوتاه اشاره می‌کنیم:

بررسی سیر تحول مفهوم «اثبات» در نظرگاه ریاضی‌دانان:

در حقیقت اکثر مطالعات پژوهشی تاریخ ریاضی به نوعی با تحولات مفهوم «اثبات» در ارتباط هستند، بالاخص زمانی که حاصل بازه‌ی زمانی گسترده‌ای فعالیت و هم‌چنین گستره‌ی مکانی وسیعی باشند. از این رو، بررسی این موضوع می‌تواند به طور مستقل و با استفاده از مطالعات علوم انسانی، به شکلی پربار صورت گیرد.

رویکرد مقاله‌نویسی اویلر:

بسیاری اویلر را صاحب یکی از درست‌ترین نگاه‌ها به ریاضیات می‌دانند و هم‌چنین پروژه‌ای در حال انجام است که تمامی مقالات او ترجمه شود تا به ایده‌های جدیدی برسند. در این میان، رویکرد مقاله‌های اویلر از نکات حائز اهمیت است. در بسیاری از مسائل، رویکرد اویلر رنگ و بوی تجربی‌تری به خود می‌گیرد و ساز و کار این رویکرد، می‌تواند حاوی نکات جدیدی باشد.

قانون تقابل؛ قبل از گاوس و بعد از گاوس:

همان‌طور که مشاهده شد، تاثیر گاوس در پیشرفت قانون تقابل بسیار عظیم است. پس از او سمت و سوی قانون تقابل به سمت تعمیم درجات بالاتر می‌رود و نظریه‌ی جبری اعداد شکل می‌گیرد و وقایع پیش از او، به طور مختصر در این نوشته ارائه شد. مطالعه‌ی تطبیقی دوران پیش و پس از گاوس، با در نظر گرفتن پارامترهای تغییر آکادمی‌ها و تالیفات، می‌تواند بسیار راه‌گشا باشد.

مراجع

[1] C. F. Gauss. Disquisitiones Arithmeticae: English Edition. Translated from the Latin, by A. A. Clarke. Yale University Press, Yale, CT, 1966.

[2] http://www.rzuser.uni-heidelberg.de/~hb3/rchrono.html

[3] K. Ireland and M. Rosen. A Classical Introduction to Modern Number Theory . Springer, New York City, NY, 1990.

[4] F. Lemmermeyer. Reciprocity Laws: From Euler to Eisenstein. Springer, New York City, NY, 2000.

[5] A.Weil. Number Theory: An Approach Through History from Hammurapi to Legendre. Hamilton Printing Company, Castleton, NY, 1987.

[6] L. Euler. Observationes circa divisionem quadratorum per numeros primes. Opera Omnia 1-3. Birkhäuser Basel, Basel, 1783.

[7] L. Euler. Theoremata circa divisores numerorum in hac forma paa $\pm$ qbb contentorum. Opera Omnia: Series 1, Volume 2, pp. 194 - 222

[8] L. Euler. Theorematica circa divisores numerorem. Opera Omnia: Series 1, Volume 2, pp. 50-61

[9] C. F. Gauss. Theoria motus corporum coelestium in sectionibus conicis solem ambientium. Sumtibus F. Perthes, 1809.

[10] E. Brown. The First Proof of the Quadratic Reciprocity Law, Revisited. The American Mathematical Monthly, 88:257-264, 1981.

[11] K. Barner. How Old Did Fermat Become? Internationale Zeitschrift für Geschichteund Ethik der Naturwissenschaften, Technik und Medizin, 9:209-228, 2001.

[12] E. Sandifer. (2005, December). http://www.maa.org/editorial/euler/HowEulerDidIt26factorsofforms.pdf, How Euler Did It

[13] R. Denomme. (2009, june, 8), https://kb.osu.edu/dspace/bitstream/handle/1811/37232/Stickelberger.pdf, A History of Stickelberger's Theorem

[14] R. Laubenbacher and D. Pengelley.(1994), http://math.nmsu.edu/~history/schauspiel/schauspiel.htmlGauß, Eisenstein, and the "third" proof of the Quadratic Reciprocity Theorem: Ein kleines Schauspiel

[15] S. Weintraub.(2011, march) On Legendre's Work on the Law of Quadratic Reciprocity. The American Mathematical Monthly, Vol. 118, No. 3,March 2011

[16] O. Baumgart. (2010, december, 4), http://www.rzuser.uniheidelberg.de/~hb3/publ/gauss4.pdf, The Quadratic Reciprocity Law. A Comparative Presentation of its Proofs

[17] J. R. Newman. The Harper Encyclopedia of Science, Volume 3. Harper & Row, New York City, NY, 1963.

[18] A. H. Beiler. Recreations in the Theory of Numbers: The Queen of Mathematics Entertains. Dover Publications, Mineola, NY, 1964.

[19] D. Cox. Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts)

[20] S. Singh. Fermat's Last Theorem: The Story of a Riddle that Confounded the World's Greatest Minds for 358 Years. Fourth Estate Ltd., London, England, 1997.

چند حقیقت درباره‌ی گروه‌ها

دکتر امیر جعفری

۱ قضیه‌ی فروبینیوس

این قضیه در مورد تعداد جواب‌های معادله‌ی

$$x^m = ۱$$

در گروه متناهی G است. این عدد را با $N_m(G)$ نشان می‌دهیم.

مشاهده‌ی ۱. *اگر ب.م.م. m و مرتبه‌ی گروه -که آن را n می‌نامیم- d باشد؛ آنگاه*

$$N_m(G) = N_d(G)$$

اثبات. واضح است که اگر $x^d = ۱$ آنگاه $x^m = ۱$. حال اگر بنویسیم $d = rn + ms$؛ آنگاه اگر $x^m = ۱$، و با توجه به اینکه طبق قضیه‌ی لاگرانژ $x^n = ۱$ داریم:

$$x^d = (x^n)^r \cdot (x^m)^s = ۱$$

پس $x^d = ۱$ و $x^m = ۱$ باهم معادلند. □

مشاهده‌ی ۲. *اگر اعداد صحیح m و m' نسبت به هم اول باشند، یک تناظر ۱-۱ بین دو مجموعه‌ی زیر وجود دارد:*

$$A = \left\{x \in G | x^{mm'} = ۱\right\}$$
$$B = \left\{(a,b) \in G \times G | a^m = b^{m'} = ۱, ab = ba\right\}$$

اثبات. نگاشتی از B به A که (a,b) را به ab می‌فرستد در نظر بگیرید. این نگاشت یک به یک است: اگر در

$$a_۱ b_۱ = a_۲ b_۲$$

طرفین را به توان m برسانید، خواهید داشت:

$$b_۱^m = b_۲^m$$

از طرف دیگر $mr + m's = ۱$ و چون $b_۱^{m'} = b_۲^{m'} = ۱$ به ازای اعداد صحیحی مانند r و s، نتیجه می‌شود $b_۱ = b_۲$ و بنابراین $a_۱ = a_۲$. همچنین این نگاشت پوشاست: اگر $x \in A$، آنگاه به ازای همان r, s در بالا $(x^{m's}, x^{mr}) \in B$ تحت این نگاشت به x می‌رود. □

مشاهده‌ی ۳. *اگر $N_m(G)$ را تعداد جواب‌های معادله‌ی $x^m = ۱$ در گروه G بنامیم، آنگاه بیان دیگر مشاهده‌ی ۲ این است که:*

$$N_{mm'}(G) = \sum_{\{b|b^{m'}=۱\}} N_m\left(C\left(b\right)\right)$$

که جمع بر روی تمام عناصر $b \in G$ که $b^{m'} = ۱$ است، زده می‌شود. $C\left(b\right)$ زیرگروه از تمام عناصری در G است که با b جابه‌جا می‌شوند یا به عبارت دیگر مرکزساز b. توجه کنید که اگر $b' = gbg^{-۱}$ آنگاه:

$$N_m\left(C\left(b\right)\right) = N_m\left(C\left(b'\right)\right)$$

در واقع $C\left(b'\right) = gC\left(b\right)g^{-۱}$ و مرتبه‌ی عناصر تحت مزدوج‌گیری تغییر نمی‌کند. بنابراین:

$$N_{mm'}(G) = \sum_b N_m\left(C\left(b\right)\right) \cdot \left[G : C\left(b\right)\right]$$

که این بار جمع روی نماینده‌های تزویجی عنصری مانند b است که $b^{m'} = ۱$. (توجه: تعداد مزدوج‌های یک عنصر $b \in G$ برابر است با اندیس $C(b)$ در G.)

$$N_{mm'}(G) = \sum_{b \in S} N_m\left(C\left(b\right)\right) \cdot \left[G : C\left(b\right)\right]$$

که S یک خانواده از نماینده‌های کامل کلاس‌های تزویجی جواب‌های معادله‌ی $b^{m'} = ۱$ است: هر جوابی از این معادله با عضو یکتایی از S مزدوج است.

حال صورت قضیه‌ی فروبینیوس را بیان می‌کنیم:

قضیه‌ی ۴. *اگر d ب.م.م. m و مرتبه گروه G -که با n نشان داده می‌شود- باشد، آنگاه:*

$$d | N_m(G)$$

در ادامه قضیه‌ی فروبینیوس را با استقرا بر مرتبه‌ی گروه ثابت می‌کنیم؛ یعنی، حکم فوق را برای هر زیرگروه سره‌ی H از G و هر m معتبر فرض می‌کنیم.

مشاهده‌ی ۵. *اگر قضیه‌ی فروبینیوس برای اعداد نسبت به هم اول m و m' برقرار باشد، آنگاه برای mm' نیز برقرار است.*

اثبات. کافی است تنها حالتی را در نظر بگیریم که m و m' مرتبه گروه G را عاد می‌کنند (به مشاهده‌ی ۱ مراجعه کنید). لذا باید ثابت کرد که اگر m مرتبه‌ی گروه را بشمارد آنگاه تعداد جواب‌های $x^m = ۱$ مضربی از m است. در این‌صورت برای هر زیرگروه H از G داریم:

$$m | N_m(H) \cdot [G : H]$$

زیرا با فرض آنکه H سره باشد؛ اگر قرار دهیم $d_۱ = (m, |H|)$ آنگاه $m = d_۱ \cdot k$ و داریم $d_۱ | N_m(H)$ (از فرض استقرا) و $k | [G : H]$ (زیرا $m||H|.[G : H]$) و نتیجه حاصل می‌شود. در شرایطی هم

که $H = G$ باز ادعای فوق معتبر است زیرا فرض کرده‌ایم قضیه‌ی فروبینیوس برای گروه G و عدد m برقرار است. پس از رابطه‌ی مشاهده‌ی ۳ نتیجه می‌شود $m|N_{mm'}(G)$ با تغییر نقش m و m' بدست می‌آید $m'|N_{mm'}(G)$ و بنابراین $mm'|N_{mm'}(G)$. □

بنابراین اثبات قضیه‌ی فروبینیوس، به اثبات آن برای حالاتی که m توانی از یک عدد اول باشد تقلیل می‌یابد.

مشاهده‌ی ۶. *اگر p عددی اول و قضیه‌ی فروبینیوس برای $m = p^k$ که مرتبه گروه را عاد می‌کند برقرار باشد؛ برای $m = p^{k-1}$ نیز برقرار است.*

اثبات. N تعداد عناصر را G از مرتبه دقیقاً p^k بگیرید. آنگاه:

$$N_{p^k} = N_{p^{k-1}} + N$$

ولی در یک گروه متناهی G می‌دانیم تعداد عناصر از مرتبه‌ی m ضربی از $\phi(m)$ است. پس $N|\ \phi\left(p^k\right)$ و چون $\phi\left(p^k\right) = p^{k-1}(p-1)$ نتیجه حاصل می‌شود. □

پایان اثبات: *حال فرض کنید مرتبه‌ی گروه به صورت $n = q \cdot m$ است که $q = p^k$ و $(m,p) = 1$. بنابراین با استقرا روی مرتبه‌ی گروه‌ها می‌دانیم اگر $C(b) \neq G$ آنگاه: $q|N_q\left(C\left(b\right)\right) \cdot \left[G : C\left(b\right)\right]$ (به استدلال بیان شده در ابتدای مشاهده‌ی ۵ مراجعه کنید.) پس از*

$$N_n(G) = \sum_{b\in S} N_m\left(C\left(b\right)\right) \cdot \left[G : C\left(b\right)\right]$$

و اینکه $N_n(G) = n$ (قضیه لاگرانژ) نتیجه می‌شود که:

$$N_q(G)\cdot|S\cap Z| \equiv 0 \text{ mod } q$$

که Z مرکز گروه است. حال چون عناصر S از مرتبه اول نسبت به p هستند (زیرا در $x^m = 1$ صدق می‌کنند و $(m,q) = 1$ بود.) $p \nmid |S\cap Z|$ بنابراین به ناچار باید $q|N_q(G)$. ولی $q = p^k$ بزرگترین توانی از p است که مرتبه‌ی G را می‌شمارد و حال مشاهده‌ی ۶ حکم را برای سایر توان‌های p نتیجه می‌دهد.

۲ برای چه n هایی هر گروه n عضوی دوری است؟

اگر بتوان یک گروه m عضوی غیر دوری یافت، برای هر مضرب m نیز گروه غیر دوری با آن مرتبه وجود دارد؛ کافی است که گروه غیردوری از آن مرتبه را در گروهی دلخواه ضرب دکارتی کنیم. از آنجایی که گروه غیر دوری از مرتبه p^2 ($\mathbb{Z}_p \times \mathbb{Z}_p$) وجود دارد، پس شرط لازم برای اینکه هر گروه n عضوی دوری باشد، آن است که n خالی از مربع باشد. مشابهاً اگر p و q دو عدد اول باشند که $q \mid p-1$، می‌توان یک گروه ناآبلی از مرتبه pq یافت:

$$G = \langle a,b | a^p = b^q = 1, bab^{-1} = a^q\rangle$$

در واقع این گروه ضرب نیمه مستقیم $\mathbb{Z}_q$ با $\mathbb{Z}_p$ - که با $\mathbb{Z}_p \rtimes \mathbb{Z}_q$ نشان داده می‌شود- با عمل $\mathbb{Z}_q \to \mathbb{Z}_p^* = \mathbb{Z}_{p-1}$ است. پس شرط لازم برای این‌که هر گروه n عضوی دوری باشد این است که $(n,\phi(n)) = 1$ (توجه: اگر $n = p_1 \cdots p_k$ که p_i ها متمایزند، آنگاه می‌دانیم که $\phi(n) = (p_1-1)\cdots(p_k-1)$.)

قضیه‌ی ۷. *این شرط کافی نیز می‌باشد.*

حکم را با استقرا روی مرتبه‌ی گروه G ثابت می‌کنیم. بنابراین می‌توان فرض کرد که هر زیرگروه سره‌ی G و هر خارج قسمتی از G که تعداد اعضایش کمتر از تعداد اعضای G باشد؛ دوری است. توجه کنید که اگر m مقسوم علیه‌ای از n باشد آنگاه شرط $(n,\phi(n)) = 1$ نتیجه می‌دهد $(m,\phi(m)) = 1$. همچنین از آنجایی که هر گروه آبلی از مرتبه‌ی خالی از مربع دوری است، پس می‌توان فرض کرد G غیرآبلی است. می‌خواهیم به تناقض برسیم.

مشاهده‌ی ۸. *از آنجایی که دوری بودن مرکز G، $Z(G)$ و گروه خارج قسمتی $G/Z(G)$ نتیجه می‌دهد G دوری است، بنابراین فرض می‌کنیم $Z(G) = 1$.*

لازم به ذکر است که اینجا از این نکته استفاده کردیم که همواره $G/Z(G)$ در صورت دوری بودن باید بدیهی باشد.

مشاهده‌ی ۹. *اگر H یک زیرگروه ماکسیمال G باشد (زیرگروه سره‌ای که هر زیرگروه بزرگتر از آن کل گروه است) و $1 \neq x \in H$. آنگاه $H = C(x)$.*

اثبات. چون H دوری است پس $H \subseteq C(x)$ و چون از مشاهده‌ی ۸ $Z(G) = 1$ پس $C(x) \neq G$ در نتیجه از ماکسیمال بودن H حکم نتیجه می‌شود. □

مشاهده‌ی ۱۰. *اگر H و H' دو زیرگروه متمایز و ماکسیمال از G باشند آنگاه $H \cap H' = 1$.*

اثبات. اگر $1 \neq x \in H \cap H'$ آنگاه بنابر مشاهده‌ی ۹

$$H = H' = C(x)$$

پس اگر H و H' متمایز باشند باید $H \cap H'$ بدیهی باشد. □

مشاهده‌ی ۱۱. *اگر H یک زیرگروه ماکسیمال G باشد آنگاه نرمال‌ساز H، $N(H)$ برابر با H است.*

اثبات. فرض کنید H دوری از مرتبه m باشد و $g \in N(H)$ بنابراین $\phi : H \to H$ که $\phi(x) = gxg^{-1}$ مرتبه‌ای دارد که هم $|Aut(H)|$ و هم $|G|$ را عاد می‌کند، ولی این دو عدد به ترتیب $\phi(m)$ و n هستند که نسبت به هم اولند. پس g با عناصر H جابه‌جا می‌شود. اگر $g \notin H$ آنگاه گروه تولید شده توسط H و g که کل گروه است گروهی آبلی خواهد شد. این زیرگروه به دلیل ماکسیمال بودنِ H یا خود H است یا کل گروه که حالت دوم بنابر فرض‌خلف رخ نمی‌دهد. $\square$

پایان اثبات: *فرض کنید H یک زیرگروه ماکسیمال k عضوی باشد. تعداد مزدوج‌های متمایز H طبق مشاهده‌ی ۱۱ برابر است با*

$$[G : N(H)] = [G : H]$$

و بنابر مشاهده‌ی ۱۰ این زیرگروه‌ها اشتراک بدیهی دارند. پس تعداد اعضای همه‌ی این مزدوج‌ها برابر است با $\frac{n}{k}(k-1)+1$. این عدد با n نمی‌تواند مساوی باشد زیرا در غیر این صورت $n(k-1) = k(n-1)$ که هیچ‌گاه به ازای اعداد طبیعی $n > k$ برقرار نیست [1] . بنابراین می‌توان یک عنصر y خارج از این اعضا در نظر گرفت. حال اگر H' زیرگروه ماکسیمالی از G باشد که y را دربردارد؛ هر مزدوج H' نیز از مزدوج‌های H متمایز است (چون مزدوج y خارج از تمامی مزدوج‌های H است) . پس با استفاده‌ی مجدد از مشاهده‌ی ۱۰؛ اگر H' ، k عضوی باشند تعداد عناصر متمایز این مزدوج‌ها برابر است با:

$$\frac{n}{k}(k-1) + \frac{n}{k'}(k'-1) + 1$$

این عدد از n بزرگتر است :

$$2n - \frac{n}{k} - \frac{n}{k'} + 1 \geq n + 1 \Leftrightarrow$$
$$\frac{n}{k} + \frac{n}{k'} \leq n \Leftrightarrow \frac{1}{k} + \frac{1}{k'} \leq 1$$

که چون $k, k' \geq 2$ برقرار است. لذا به تناقض می‌رسیم.

۳ شرط لازم وکافی برای نرمال بودن

هرگاه H زیرگروه G با اندیس متناهی n باشد، آنگاه اگر $H \trianglelefteq G$ برای هر $x \in G$ خواهیم داشت:

$$x^n \in H$$

زیرا که G/H یک گروه n عضوی است و بنابر قضیه‌ی لاگرانژ $(xH)^n = H$ پس $x \in H$. حال می‌خواهیم عکس این حکم را بررسی کنیم:

سؤال ۱۲. *آیا اگر برای هر $x \in G$ داشته باشیم $x^n \in H$ آنگاه H زیرگروه نرمال G خواهدبود؟*

این حکم در حالت کلی غلط است. برای مثال به گروه ۲۷ عضوی ناآبلی G فکر کنید [2] که هر عضو آن به توان ۳ برابر ۱ می‌شود. می‌خواهیم ثابت کنیم اگر n عددی اول باشد، حکم صحیح است.

مشاهده‌ی ۱۳. *اگر $[G : H] = n$ آنگاه برای هر $x \in G$ داریم $x^{n!} \in H$.*

اثبات. هم‌دسته‌های $H, xH, x^2H, \cdots, x^nH$ را در نظر بگیرید، چون تعداد آن‌ها از n بیشتر است، پس $i < j$ وجود دارد که:

$$x^i \in H = x^jH \Rightarrow x^{j-i} \in H$$

چون $j - 1 \leq n$ پس $j - i \mid n!$ پس $x^{n!} \in H$. $\square$

مشاهده‌ی ۱۴. *اگر $x \notin H$ آنگاه هم‌دسته‌های $H, xH, x^2H, \cdots, x^{n-1}H$ با هم متمایزند.*

اثبات. در غیر این صورت $x^iH = x^jH$ پس $k < n$ وجود دارد که $x^k \in H$. چون طبق فرض $x^n \in H$ پس $x^{nr+ks} \in H$. حال r و s را با استفاده از $(n, k) = 1$ طوری بگیرید که $nr + ks = 1$. پس $x \in H$ که تناقض است. $\square$

پایان اثبات: *فرض کنید H نرمال نباشد. پس $x \notin H$ وجود دارد که برای یک $y \in H$ ، $xyx^{-1} \notin H$ را برآورده می‌کند. بنابراین از مشاهده‌ی ۱۴ هم‌دسته‌های*

$$H, xyx^{-1}H, xy^2x^{-1}H, \cdots, xy^{n-1}x^{-1}H$$

متمایزند. لذا همه‌ی هم‌دسته‌ها باید اینجا ظاهر شده باشند. فرض کنید $x \in xy^kx^{-1}H$. پس $h \in H$ وجود دارد که:

$$x = xy^kx^{-1}h \Rightarrow x^{-1} = y^{-k}h^{-1}$$
$$\Rightarrow x = hy^k \in H$$

که تناقض است.

سؤال ۱۵. *برای چه n های دیگری می‌توان حکمی مشابه قبل را ثابت کرد؟*

[1] این حالت خاصی است از یک حکم کلی‌تر: یک گروه متناهی نمی‌تواند با مزدوج‌های یک زیرگروه سره‌اش پوشانده شود.

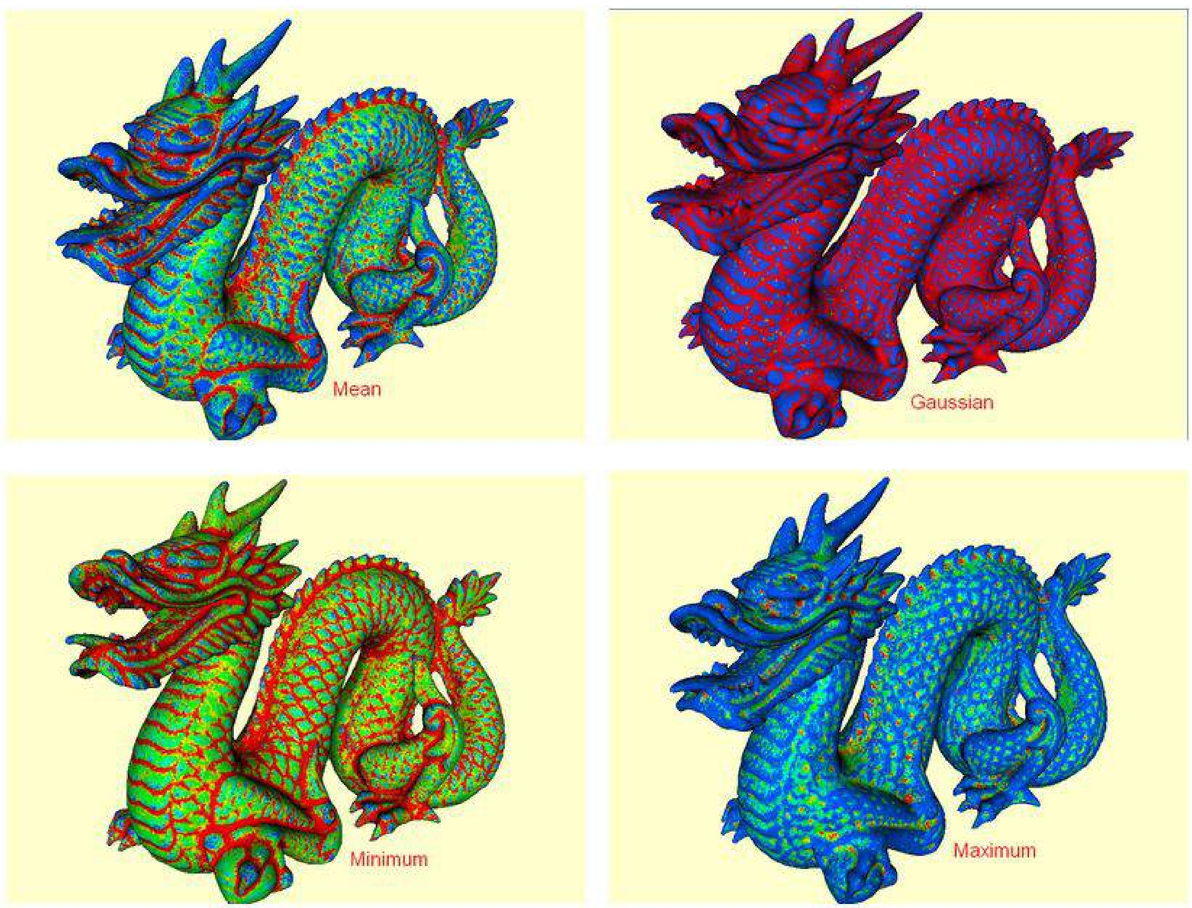
[2] گروه ماتریس‌های بالامثلثی و 3×3 با درایه‌های در میدان $\mathbb{Z}_3$ که در آنها درایه‌های روی قطر یک باشند.

کمی دورتر از فضای اقلیدسی

ابوالفضل طاهری

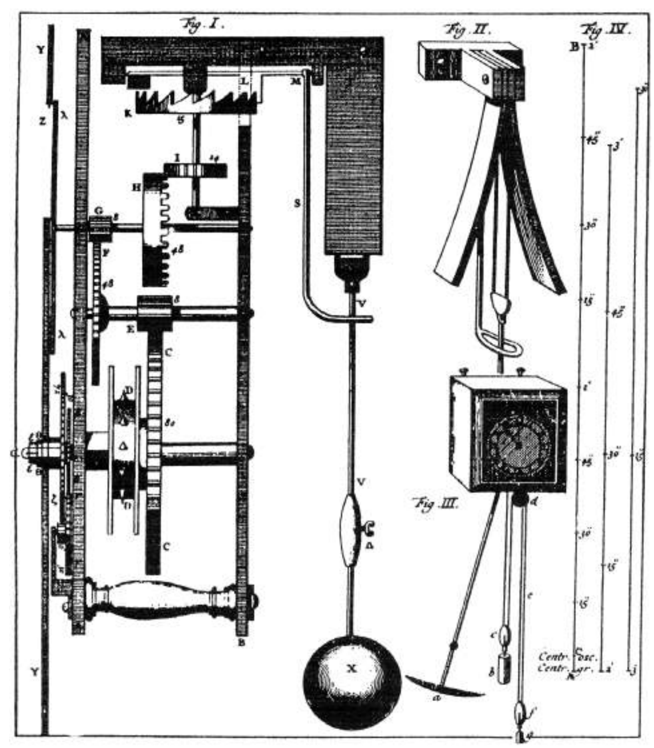
چکیده

احتمالاً همگی در دنیای واقعی تصوری شهودی از خمیدگی (انحنا) داریم؛ پیچ‌هایی از جاده که سریعتر می‌پیچند خطرناک‌ترند! اما در دنیای ریاضیات، "انحنا"های زیادی وجود دارد: انحنای اصلی، انحنای میانگین، انحنای گاوسی، انحنای ریمانی، انحنای مقطعی، انحنای ریچی، انحنای اسکالر و ... . به دست آوردن شهود نسبت به برخی از این انحناها، هم‌چون انحنای اصلی، کار چندان دشواری نیست اما فهم اغلب آن‌ها هم‌چون انحنای ریچی به سادگی میسر نمی‌شود. در این مقاله قصد بر آن است تا برخی انحناهای معروف را مطرح و آن‌ها را به صورتی شفاف بیان کنیم. اما قبل از آن باید دید هدف از تعریف انحنا چیست و چه انتظاراتی از آن داریم؟



مقدمه

انحنا[۱] یکی از بنیادی‌ترین مفاهیمی است که در ریاضیات مطرح شده است. شاید اولین کسی که به طور رسمی از مفهوم انحنا در کارهای خود استفاده کرد، کارل فردریش گاوس[۲] بود. گاوس در مطالعه‌ی رویه‌ها این مفهوم را وارد کرد و فرمولی برای انحنا به دست آورد. اما سال‌ها پیش از گاوس نیز افرادی از این مفهوم در کارهای خود استفاده کرده بودند. هیوژن[۳]، منجم، فیزیک‌دان و ریاضی‌دان هلندی قرن هفدهم، یکی از افراد است. هیوژن توانست با کمک این مفهوم ساعت پاندولی طراحی کند که زمان را با دقت بسیار خوبی نمایش می‌داد. جزئیات کارهای او را می‌توانید در [۸] ببینید.



شکل ۱: ساعت پاندولی هیوژن

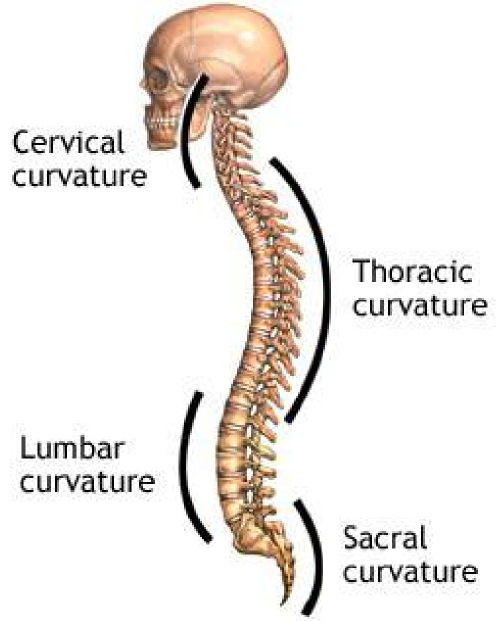
کپلر[۴]، نیوتن[۵]، لایب‌نیتز[۶] و اویلر[۷] از دیگر افرادی هستند که انحنا را به نحوی در کارهای خود به کار برده‌اند. ([۹]، [۳])

اما قبل از هر چیز باید ببینیم هدف ما از تعریف انحنا (خمیدگی) چیست؟ و در ابتدا سعی کنیم برأساس همان شهودی که در ذهن از خمیدگی داریم آن را تعریف کنیم. از نامی که بر این مفهوم نهاده‌اند -خمیدگی- پیداست قبل از تعریف آن نیاز به خم داریم! در قسمت اول سعی می‌کنیم خم و مفاهیمی از آن را که مورد نیاز است به صورت مختصر مطرح کنیم.

در ادامه به دنبال بیانی ریاضی از خمیدگی که در ذهن داریم خواهیم پرداخت و در بخش‌های بعد به دنبال گسترش مفهوم انحنا به فضاهای دیگر و با اهداف دیگر خواهیم بود.

خم‌های هموار

چهار منحنی طبیعی در ستون فقرات وجود دارد؛ گردن[۸]، قفسه‌ی سینه[۹]، کمر[۱۰] و ساکروم[۱۱]. این خم‌ها به همراه دیسک‌های بین مهره‌ای، کمک می‌کنند که فشارها و تنش‌هایی که از فعالیت‌های روزمره مانند راه رفتن و یا فعالیت‌های شدید مانند دویدن و پریدن رخ می‌دهد، پخش شده و فشار کمتر بر بدن و ستون فقرات وارد آید.[۱۲]



شکل ۲: خم‌های ستون فقرات که باعث کاهش فشارهای ناشی از فعالیت‌های ما می‌شوند.

برای بررسی چنین خم‌هایی به کمک ریاضیات ابتدا لازم است آن‌ها را به زبان ریاضیات نمایش دهیم به همین خاطر در این بخش به معرفی

[۱] Curvature
[۲] Karl Friedrich Gauss
[۳] Christiaan Huygens
[۴] Johannes Kepler
[۵] Sir Isaac Newton
[۶] Gottfried Wilhelm Leibniz
[۷] Leonhard Euler

[۸] Cervical
[۹] Thoracic
[۱۰] Lumbar
[۱۱] Sacral
[۱۲] http://www.nlm.nih.gov/medlineplus/ency/imagepages/19463.htm

اجمالی چنین خم‌هایی می‌پردازیم. اغلب مطالب این بخش از [۱۸] و [۱] انتخاب شده است. برای جزئیات بیشتر می‌توانید به همین مراجع، مراجعه کنید.

مقصود از خم پارامتری[۱۳] (یا خم پرمایش شده) در $\mathbb{R}^n$ -اغلب به اختصار خم- تابعی هم‌چون $\gamma : I \to \mathbb{R}^n$ از یک بازه‌ی I به $\mathbb{R}^n$ است. اغلب می‌توانیم I را بازه‌ای زمانی تلقی کنیم که در این صورت $\gamma(t)$ را به صورت "مکان یک ذره در زمان t" تعبیر می‌کنیم. نماد r نیز در $r = \gamma(t)$ به معنای "بردار مکان" برای نمایش مقدار تابع γ مرسوم است. ([۱۸])

بنابراین خم، یک تابع است و می‌توان از قضایا و تعاریف حساب دیفرانسیل و انتگرال استفاده کرد. بر همین اساس تعاریف و قضایای زیر را داریم:

تعریف ۱. *یک خم هموار پارامتری یک نگاشت $\gamma : I \to \mathbb{R}^n$ است که γ مشتق‌پذیر است و مشتق آن پیوسته است بعلاوه مشتق γ همواره ناصفر است.*

اگر $\gamma(t) = (x_1(t), \ldots, x_n(t))$ را به این شکل نمایش دهیم، در این صورت منظور از مشتق‌پذیر بودن γ، مشتق‌پذیر بودن $x_1, \ldots, x_n$ است و داریم $\gamma'(t) = (x_1'(t), \ldots, x_n'(t))$.

مثال ۲. *خم فضایی $\gamma : \mathbb{R} \to \mathbb{R}^3$ را به صورت $\gamma(t) = (a\cos t, a\sin t, bt)$ تعریف می‌کنیم که در آن a و b دو عدد حقیقی داده شده‌اند. این خم مارپیچ دوار[۱۴] نام دارد. تصویر خم، یعنی مجموعه‌ی*

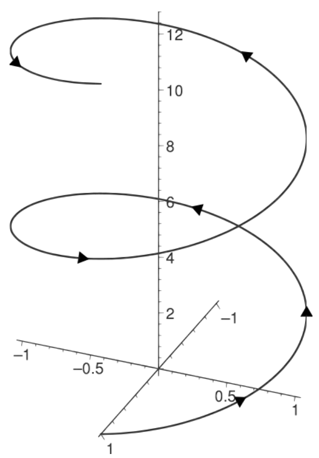
$$\{(x,y,z) = (a\cos t, a\sin t, bt) : t \in \mathbb{R}\}$$

روی استوانه‌ای قرار دارد که بر دایره‌ی $x^2 + y^2 = a^2$ در صفحه‌ی (x,y) بنا شده است و با آهنگ ثابت b در امتداد محور z روی استوانه می‌پیچد. مقدار z نسبت به t به ازای $b > 0$ صعودی و به ازای $b < 0$ نزولی است. ([۱۸])

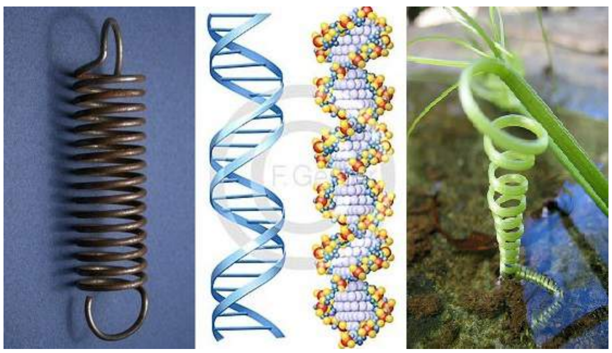
مارپیچ دوار در طبیعت به شکل‌های مختلفی ظاهر می‌شود که نمونه‌هایی از آن را در شکل ۴ می‌بینید. (شکل سمت چپ یک فنر، شکل میانی ساختار مولکول DNA و شکل سمت راست بخشی از یک گیاه است).

فرض کنید $\gamma : I \to \mathbb{R}^n$ یک خم هموار باشد. با در نظر گرفتن تعبیر متداول $\gamma'(t)$ در حکم سرعت در زمان t، می‌توانیم طول این بردار $|\gamma'(t)|$ را تندی حرکت در زمان t تصور کنیم. از آن‌جا که تندی حرکت را معمولاً آهنگ پیمودن مسافت نیز تعبیر می‌کنیم، تعریف

[۱۳] Parametrized Curve
[۱۴] Helix



شکل ۳: مارپیچ دوار به ازای $a = 1$ و $b = 1$.



شکل ۴: اشکال مختلف مارپیچ دوار

مناسبی برای طول خم به صورت زیر به دست می‌آوریم. نقطه‌ای مانند t_0 در I را به مثابه‌ی مبدا اندازه‌گیری طول تثبیت و تابع طول خم $l : I \to \mathbb{R}$ را به صورت زیر تعریف می‌کنیم:

$$s = l(t) = \int_{t_0}^{t} |\gamma'|$$

به طور مثال در مورد مارپیچ دوار داریم:

$$\gamma'(t) = (-a\sin t, a\cos t, b)$$

$$|\gamma'(t)| = \sqrt{a^2 + b^2}$$

بنابراین طول خم برابر می‌شود با:

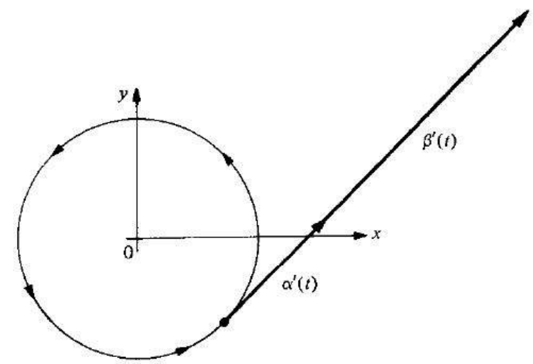
$$\int_0^T \sqrt{a^2+b^2}\,\mathrm{d}t = T\sqrt{a^2+b^2}$$

دو خم

$$\alpha(t) = (\cos t, \sin t), \beta(t) = (\cos 2t, \sin 2t)$$

را در نظر بگیرید. تصویر این دو خم یکی و همان دایره‌ی واحد ($x^2+y^2=1$) است اما تندی حرکت β دو برابر خم α است. بنابراین یک خم را می‌توان به شکل‌های مختلفی پرمایش کرد. اما دقت کنید

که با یک تابع پیوسته $f : [0, 2\pi] \to [0, \pi]$، $f(t) = \frac{t}{2}$ می‌توان پرمایش α را به β تبدیل کرد. به این تغییر پرمایش، بازپرمایش خم می‌گوییم.



شکل ۵: پرمایش‌های مختلف یک خم

گزاره ۳. ([۱۸]) *طول یک خم به پرمایش خاص تصویر یک خم وابسته نیست. دقیق‌تر، فرض کنیم $\tilde{\gamma} : [a', b'] \to \mathbb{R}^n$ یک بازپرمایش $\gamma : [a, b] \to \mathbb{R}^n$ باشد. اگر $\alpha : [a', b'] \to [a, b]$ تناظر بین دو بازه‌ی تعریف شده باشد، $\tilde{\gamma} = \gamma \circ \alpha$، و l و $\tilde{l}$ توابع طول مربوط به γ و $\tilde{\gamma}$ باشد، آن‌گاه:*

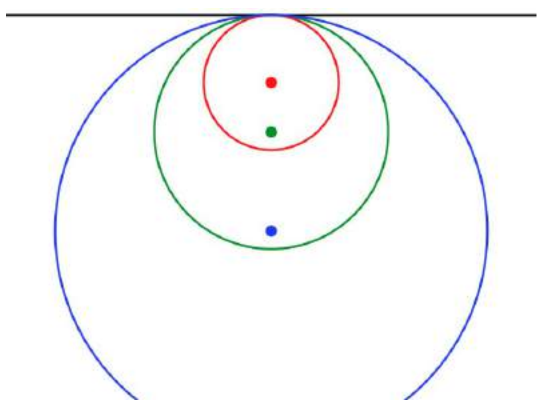
$$|l(\alpha(\tau))| = |\tilde{l}(\tau)|$$

انحنای خم

انحنای خم در ریاضیات، در واقع همان تصوری است که از انحنا در ذهن داریم. میزانی خمیدگی یک خط راست که در واقع تفاوت خم با یک خط راست است، و این می‌تواند همان معیار ما برای سنجیدن میزان انحنای یک خم باشد؛ میزان تفاوت در هر نقط با یک خط راست. بنابراین انتظار می‌رود که انحنای خط راست صفر باشد.

اولین نکته‌ای که از این تعریف به نظر می‌رسد شاید این باشد که اگر در هر نقطه میزان تفاوت از یک خط راست را داشته باشیم، باید بتوان خم به صورت یکتا مشخص کرد (با در نظر گرفتن انتقال و دوران در صورت لزوم). نگاهی به شکل ۲ شاید این مسأله را واضح‌تر کند. ستون فقرات خط راستی بوده است که هر نقطه‌ی آن به میزان مشخصی از این خط راست فاصله گرفته است!

اگر به یک دایره نگاهی بیندازیم، چیزی که به نظر می‌رسد این است که میزان تفاوت آن با خط راست در هر دو نقطه‌ای یکسان است. بنابراین انتظار می‌رود دایره انحنای ثابتی در تمامی نقاط داشته باشد، بنابراین با کمک نکته‌ی اول باید بتوان یک دایره را تنها با یک عدد که انحنای آن است به صورت یکتا (با در نظر گرفتن انتقال) مشخص کرد. و می‌دانیم که یک دایره به کمک شعاع آن به صورت یکتا مشخص می‌شود! بنابراین باید ارتباط نزدیکی بین شعاع دایره و انحنای آن وجود داشته باشد. نکته‌ی دیگری که در مورد دایره به نظر می‌رسد این است که هر چه شعاع آن بیشتر باشد هر قسمت کوچک آن به خط راست نزدیکتر است!

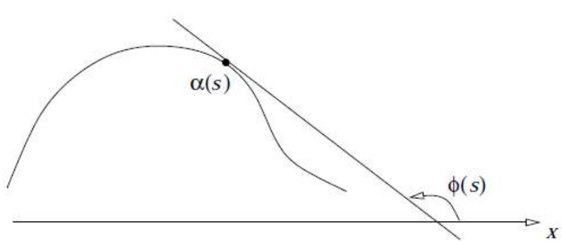


شکل ۶: دایره‌های بزرگتر به صورت موضعی به خط راست نزدیکترند.

مسأله‌ی دیگری که به چشم می‌آید این است که انحنا تنها به شکل خم بستگی دارد و مستقل از نحوه‌ی پرمایش خم است. بنابراین انتظار می‌رود که نسبت به بازپرمایش خم ناوردا باشد.

با این اهداف و انتظارات سعی می‌کنیم انحنا را به زبان ریاضی برای خم‌های در صفحه تعریف کنیم و از این پس آن را با κ نمایش می‌دهیم.

فرض کنید α یک خم باشد که بر حسب طول پرمایش شده است. هم‌چنین ϕ را تابعی بگیرید که در هر نقطه زاویه‌ی خم مماس بر خم را با محور xها مشخص می‌کند.



میزان تفاوت از خط راست در واقع همان آهنگ تغییر $\phi(s)$ نسبت به تغییرات طول خم است. به عبارتی داریم:

$$\kappa = |\frac{\mathrm{d}\phi}{\mathrm{d}s}|$$

حال بیایید انحنای دایره را حساب کنیم. فرض کنید α دایره‌ای به شعاع r است که به صورت پادساعتگرد آن را پرمایش کرده‌ایم. در

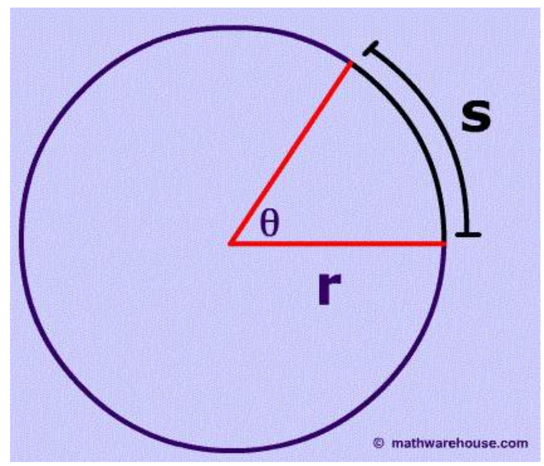
این صورت

$$\phi(s) = \frac{\pi}{۲} + \theta$$

که در آن θ زاویه نسبت به محور x است. بنابراین داریم $\theta = \frac{s}{r}$ و در نتیجه به دست می‌آید:

$$\kappa = \frac{\mathrm{d}\phi}{\mathrm{d}s} = \frac{۱}{r}$$

که همان چیزی است که انتظارش را داشتیم.



شکل ۷: $s = r\theta$

با انتخاب یک جهت می‌توانیم قدرمطلق را حذف کنیم و انحنا را با علامت آن در نظر بگیریم. اما دقت کنید که این علامت به نحوه‌ی پرمایش خم وابسته است. به طور مثال اگر در مورد دایره، آن را به صورت پادساعتگرد پرمایش می‌کردیم انحنای آن $-\frac{۱}{r}$ به دست می‌آمد. بنابراین از این به بعد جهت مثبت محور x را به عنوان جهت استاندارد در نظر می‌گیریم.

اگر فرض کنیم $T(s) = \alpha'(s)$، در این صورت چون α پرمایش برحسب طول است، بنابراین واحد طول در واحد زمان طی می‌شود پس داریم $||T||^۲ = T.T = ۱$ و در نتیجه به دست می‌آید:

$$T'.T + T.T' = ۰ \Rightarrow T.T' = ۰$$

بنابراین تغییرات $T(s)$ باید در جهت بردار نرمال، N باشد. انحنا نیز میزان چرخش $T(s)$ در طول بردار نرمال $N(s)$ است وقتی که $||T(s) \times N(s)|| = ۱$. بنابراین تعریف دیگری برای انحنا به این شکل به دست می‌آید:

$$T' = \frac{\mathrm{d}T}{\mathrm{d}s} = \kappa N$$

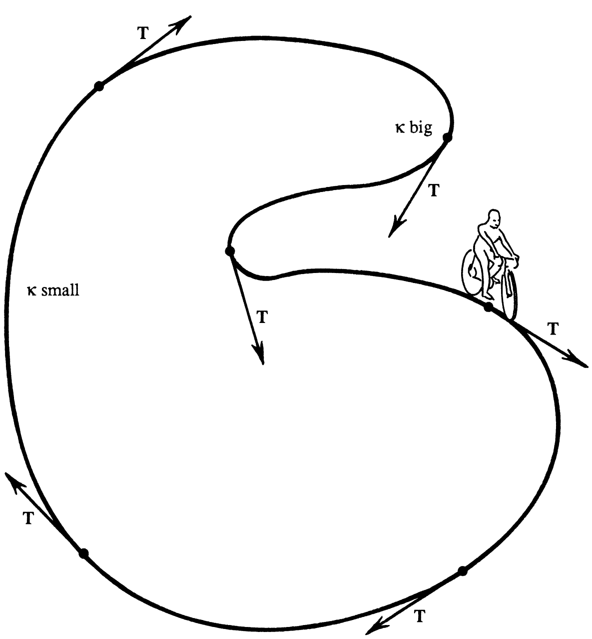
که به سادگی نتیجه می‌شود این دو تعریف معادل‌اند:

$$\begin{aligned}
\kappa &= T'.N \\
&= \frac{\mathrm{d}N}{\mathrm{d}s}.N \\
&= \lim_{\Delta s \to ۰} \frac{T(s+\Delta s) - T(s)}{\Delta s}.N \\
&= \lim_{\Delta s \to ۰} \frac{\Delta\phi.||T||}{\Delta s} \\
&= \lim_{\Delta s \to ۰} \frac{\Delta\phi}{\Delta s} \\
&= \frac{\mathrm{d}\phi}{\mathrm{d}s} = \kappa
\end{aligned}$$

قضیه‌ی زیر یکی دیگر از انتظارات ما را از تعریف انحنا برآورده می‌کند. اثبات این قضیه را می‌توانید در [۱۴] ببینید.

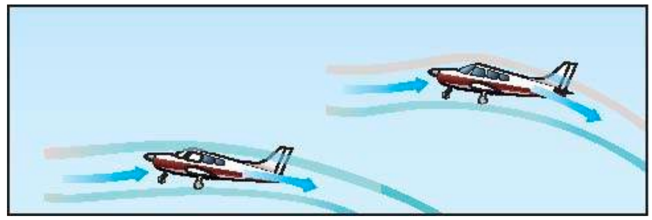
قضیه ۴. *فرض کنید $\kappa : (a,b) \to \mathbb{R}$ یک تابع انتگرال‌پذیر باشد. در این صورت خم $\alpha : (a,b) \to \mathbb{R}^۲$ که برأساس طول پرمایش شده است وجود دارد به طوری که انحنای آن κ باشد. بعلاوه اگر $\tilde{\alpha} : (a,b) \to \mathbb{R}^۲$ خم دیگری با همین خصوصیات باشد آن‌گاه با انتقال و دوران می‌توان آن را به α تبدیل کرد.*

حال با این نتایج، می‌توانید بگویید که چرا پیچ‌هایی از جاده که سریعتر می‌پیچند، خطرناک‌ترند؟!



رویه‌های هموار

رویه[۱۵]ها یکی دیگر از طبیعی‌ترین اشیاء ریاضی هستند و کاربردهای وسیعی در فیزیک، مهندسی، گرافیک کامپیوتری و بسیاری دیگر از علوم دارند.



شکل ۸: برای بررسی خواص آیرودینامیک هواپیما، مسأله‌ی اصلی بررسی جریان هوا در طول رویه است.

شکل ۹: برای ساخت پویانمایی‌ها و بازی‌های کامپیوتری نیاز به شناسایی و کار با رویه‌ها داریم.

رویه‌ها موجودات عجیبی نیستند. تقریباً تمام آن چیزی که در واقعیت می‌بینیم را می‌توان به عنوان یک رویه در نظر گرفت. در واقع هر قسمت کوچک از این موجودات را می‌توان با کمی تغییرات به صفحه ($\mathbb{R}^۲$) تبدیل کرد. این واقعیت را در مورد کره‌ی زمین به سادگی می‌توان دید. و اینکه می‌توانیم نقشه‌ی کره‌ی زمین را روی کاغذ داشته باشیم از همین واقعیت حاصل می‌شود.

هم‌چون خم‌ها، در این‌جا نیز به بررسی رویه‌های هموار می‌پردازیم. آن‌هایی که شکستگی و تیزی ندارند و در واقع رفتارهایی طبیعی از خود نشان می‌دهند. بنابراین تعریف زیر را از [۱] برای رویه داریم.

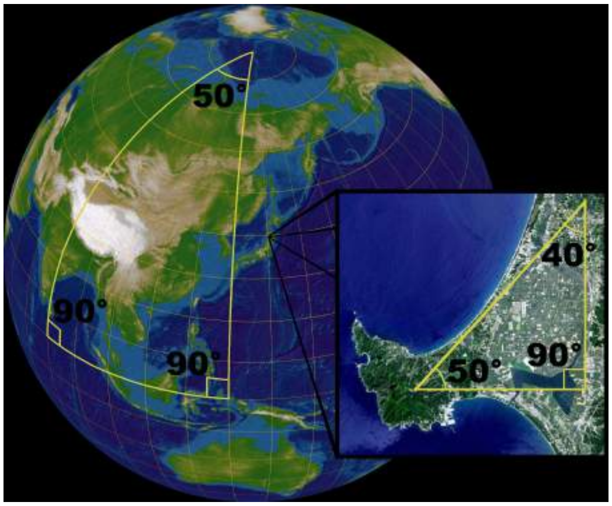
تعریف ۵. *زیرمجموعه‌ی $S \subset \mathbb{R}^۳$ یک رویه‌ی هموار است اگر برای هر $p \in S$ همسایگی V در $\mathbb{R}^۳$ و نگاشت $\mathbf{x} : U \to V \cap S$ از باز $U \subset \mathbb{R}^۲$ به توی $V \cap S \subset \mathbb{R}^۳$ وجود داشته باشد به طوری که:*

- $\mathbf{x}$ *مشتق‌پذیر باشد. به عبارتی اگر*

$$\mathbf{x}(u,v) = (x(u,v), y(u,v), z(u,v)), (u,v) \in U$$

نمایش دهیم، توابع $x(u,v)$، $y(u,v)$ و $z(u,v)$ از هر مرتبه‌ای در U مشتق پاره‌ای پیوسته داشته باشند.

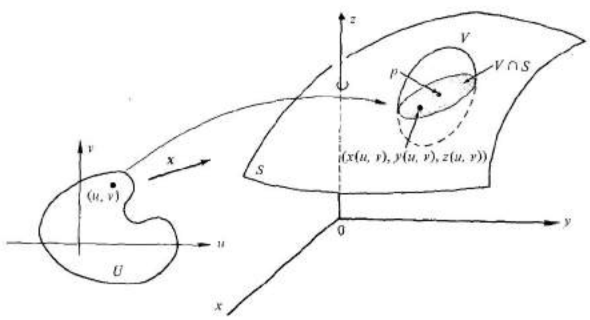
[۱۵] Surface



شکل ۱۰: رویه‌ها به صورت موضعی بخشی از صفحه‌اند.

- $\mathbf{x}$ *یک همسان‌ریختی باشد. چون بنابر خاصیت اول $\mathbf{x}$ پیوسته است این بدان معناست که $\mathbf{x}^{-۱} : V \cap S \to U$ وجود داشته باشد و پیوسته باشد.*

- *برای هر $q \in U$ مشتق $\mathrm{d}\mathbf{x}_q : \mathbb{R}^۲ \to \mathbb{R}^۳$ یک‌به‌یک است.*



شکل ۱۱: رویه فضایی توپولوژیک، هاسدورف و شمارای نوع دوم است که برای هر نقطه‌ی آن همسایگی همسان‌ریخت با زیرمجموعه‌ای باز از صفحه‌ی اقلیدسی وجود دارد.

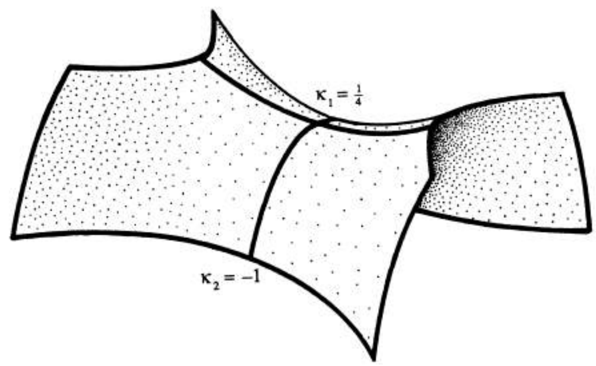
در این قسمت به تعریف رویه بسنده می‌کنیم و در ادامه می‌خواهیم همان کاری را که برای خم‌ها انجام دادیم برای رویه‌ها انجام دهیم: تعریف ابزاری به نام انحنا که ما را در شناسایی و مطالعه‌ی رویه‌ها کمک کند. برای آشنایی بیشتر با رویه‌ها می‌توانید از [۱]، [۱۱] و [۱۲] استفاده کنید.

اما برای تعریف انحنای رویه، کاندیداها و ایده‌های مختلفی وجود دارد که هر کدام را به طور جداگانه بررسی می‌کنیم.

انحنای رویه‌ها

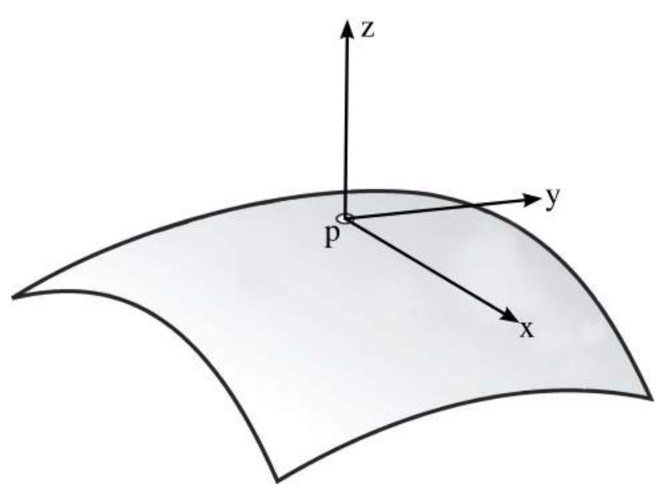
اولین ایده‌ای که به ذهن می‌رسد، استفاده از انحنای خم و گسترش آن برای رویه‌هاست. فرض کنید S یک رویه و p نقطه‌ای از آن باشد. می‌خواهیم انحنای S را در p به کمک انحنای خم تعریف کنیم. اما در S بی‌نهایت خم وجود دارد که از نقطه‌ی p می‌گذرد. بنابراین نیاز داریم که خم‌هایی خاص را انتخاب کنیم تا تعداد انحناهای به دست آمده را به مقدار خوبی کاهش دهیم و بتوانیم برأساس آن نتایج معقولی بگیریم.

در انتخاب خم‌ها نیز اولین ایده‌ای که به ذهن می‌رسد، خم‌هایی است که کمترین و بیشترین انحنا را در نقطه‌ی p دارند.



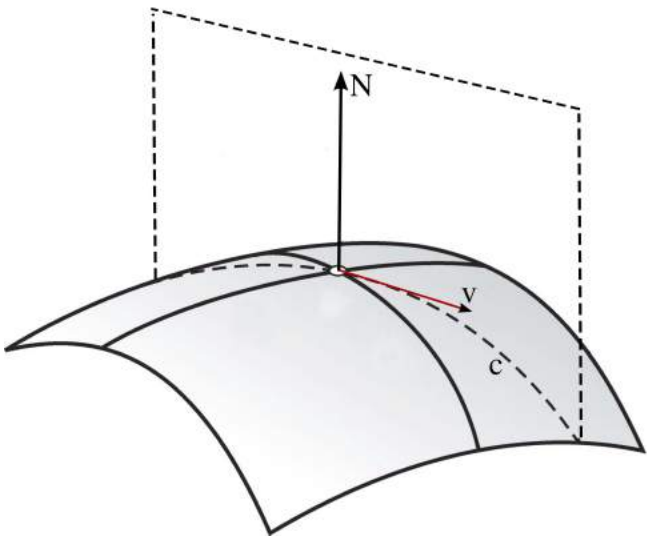
شکل ۱۲: خم‌های با انحنای کمینه و بیشینه در رویه‌ی زین اسبی.

بیایید کمی بیشتر این خم‌ها را بررسی کنیم. با کمی انتقال و دوران روی رویه‌ی S می‌توانیم فرض کنیم نقطه‌ی p در مرکز مختصات قرار دارد و صفحه‌ی مماس بر آن صفحه‌ی $z = 0$ است.[۱۶] بنابراین $N = (0, 0, 1)$ بردار نرمال واحد بر S در نقطه‌ی p است.



[۱۶] بر هر رویه S می‌توان در هر نقطه p یک صفحه مماس کرد که آن را با T_pS نمایش می‌دهیم.

حال فرض کنید $v = (v_1, v_2, 0)$ یک بردار (سرعت) واحد در صفحه‌ی T_pS باشد. c را خم پارامتری شده‌ای که از تقاطع صفحه‌ی حاصل از v و N با رویه‌ی S به دست می‌آید، در نظر بگیرید.

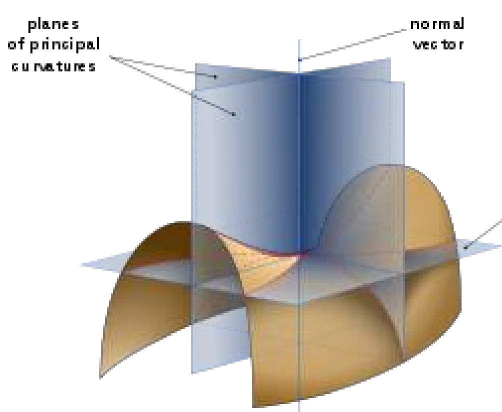


در این صورت c به صورت زیر تعریف می‌شود:

$$c(t) = (v_1 t, v_2 t, f(v_1 t, v_2 t))$$

که در آن f نگاشتی است که نمودار آن در $\mathbb{R}^3$ همان رویه‌ی S خواهد بود.[۱۷] حال می‌توانیم انحنای c را به دست آوریم. با کمی محاسبات می‌توانیم نتیجه بگیریم:

گزاره ۶. *خم‌هایی در نقطه‌ی p که بردار سرعت آن‌ها، V_p، در جهت (خلاف جهت) $\frac{\mathrm{d}N(t)}{\mathrm{d}t}$ باشد –که $N(t)$ بردار نرمال واحد بر S در نقاط خم است– همان خم‌هایی روی S هستند که انحنای کمینه (بیشینه) را در نقطه‌ی p دارند.*



[۱۷] هر رویه را می‌توان موضعاً به عنوان نمودار تابعی به دامنه‌ی بازی از صفحه گرفت.

بنابراین به کمک این گزاره می‌توانیم انحنای بیشینه و کمینه را که با κ_1 و κ_2 نمایش می‌دهیم به دست آوریم. این انحناها را انحنای اصلی می‌نامیم.

انحنای گاوسی را برای رویه‌ی S در نقطه‌ی p، $K = \kappa_1\kappa_2$ و انحنای میانگین را $H = \kappa_1 + \kappa_2$ [۱۸] تعریف می‌کنیم. بنابراین در شکل ۱۲ داریم $K = -\frac{1}{4}$ و $H = -\frac{3}{4}$.

اما این تعاریف انتظارات ما را برآورده می‌سازد؟! در اولین گام انحنای کره را محاسبه کنیم. با توجه به خصوصیاتی که برای انحنای دایره داشتیم، انحناهای اصلی برای کره به شعاع r به دست می‌آید $\kappa_1 = \kappa_2 = \frac{1}{r}$ و در نتیجه $K = \frac{1}{r^2}$ و $H = \frac{2}{r}$. که مستقل از نقطه‌ی انتخاب شده است و بنابراین کره انحنای ثابت دارد. قضیه‌ای که لیبمان[۱۹] در سال ۱۹۰۰ ارائه کرد این مطلب را کامل می‌کند:

قضیه ۷. *اگر S یک رویه (بسته) با انحنای گاوسی ثابت K باشد در این صورت S یک کره به شعاع $\frac{1}{\sqrt{K}}$ است.* ([۱۶])

به همان شکلی که برای کره بررسی شد، دیده می‌شود که انحنای میانگین و گاوسی برای صفحه همواره صفر است. بنابراین تنها این سؤال باقی می‌ماند:

انحنایی که تعریف شد تحت چه شرایطی بدون تغییر خواهد ماند؟ در واقع آیا می‌توان گفت که این انحنا (هم‌چون انحنای خم) تنها به خود رویه بستگی دارد و یکی از خواص ذاتی آن است؟

پاسخ این سؤال یکی از بنیادی‌ترین نتایج هندسه دیفرانسیل[۲۰] است که توسط کارل فردریش گاوس تحت عنوان قضیه‌ی چشمگیر[۲۱] اثبات شد. گاوس ثابت کرد که انحنای گاوسی نسبت به خمش ناوردا می‌ماند. منظور از خمش نیز تغییراتی روی رویه است که تحت آن‌ها طول خم‌ها و زوایای بین آن‌ها بدون تغییر می‌ماند. در واقع به طور دقیق، انحنای گاوسی تحت نگاشت‌های موضعاً ایزومتریک[۲۲] ناورداست.

تعریف ۸. *نگاشت $f : S_1 \to S_2$ را یک وابرسانی[۲۳] می‌گوییم اگر f یک همسان‌ریختی مشتق‌پذیر، با مشتق پیوسته باشد که وارون آن نیز مشتق‌پذیر با مشتق پیوسته است. و آن را موضعاً وابرسانی می‌گوییم اگر برای هر نقطه از S_1 همسایگی بازی وجود داشته باشد که تحدید f به آن وابرسانی باشد.*

[۱۸] در برخی مراجع برای انحنای میانگین یک ضریب یک‌دوم نیز در نظر گرفته می‌شود.

[۱۹] Liebmann

[۲۰] Differential Geometry

[۲۱] Gauss's Theorema Egregium (Latin: "Remarkable Theorem")

[۲۲] Local Isometry

[۲۳] Diffeomorphism

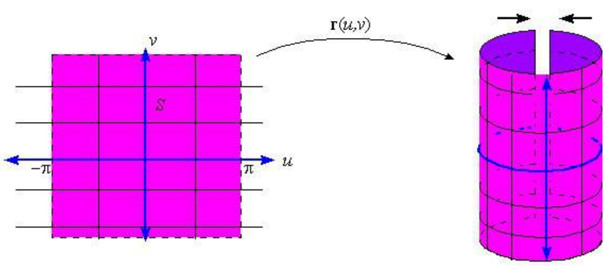
نگاشت f را ایزومتری موضعی می‌گوییم اگر یک وابرسانی موضعی باشد که تحت آن فاصله‌ها حفظ می‌شود.

بنابراین قضیه‌ی گاوس به شکل زیر بیان می‌شود:

قضیه ۹. *انحنای گاوسی تحت ایزومتری‌های موضعی ناورداست. به عبارتی اگر $f : S_1 \to S_2$ یک ایزومتری موضعی باشد در این صورت:*

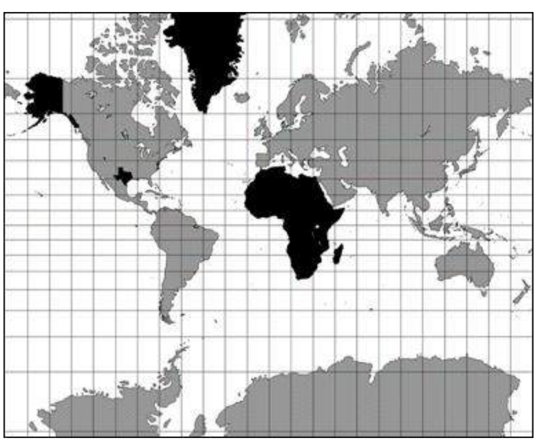
$$K_{S_1}(p) = K_{S_2}(f(p)), \forall p \in S_1$$

در مقاله‌ی [۱۷] می‌توانید استدلالی شهودی بر این قضیه را ببینید. به سادگی می‌توان از این قضیه نتایج مفیدی را به دست آورد. به طور مثال یک استوانه را در نظر بگیرید. استوانه با یک صفحه موضعاً ایزومتریک است بنابراین باید انحنای آن صفر باشد.



شکل ۱۳: موضعاً ایزومتریک بودن صفحه و استوانه

یکی از نتایج دیگر که به سادگی حاصل می‌شود: هیچ بخشی از کره را نمی‌توان بدون تغییر طول‌ها روی صفحه قرار داد. چون انحنای گاوسی صفحه صفر است ولی کره انحنای مثبت دارد.



شکل ۱۴: نقشه‌ی کره‌ی زمین را نمی‌توان با حفظ مقیاس‌ها روی صفحه‌ی مسطح ترسیم کرد.

یک مثال از رویه‌های موضعاً ایزومتریک که چندان واضح به نظر

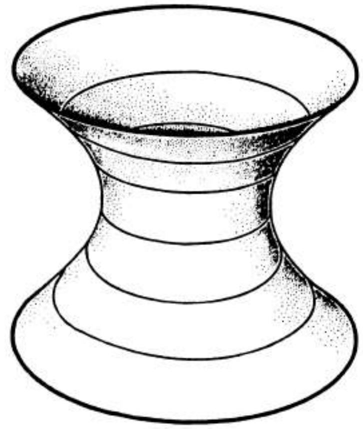
نمی‌رسد، هذلولی‌وار[24] و حلزونی شکل[25] است. برای هذلولی‌واری که با

$$F:(z,\theta)\mapsto(\cos\theta\cosh z,\sin\theta\cosh z,z)$$

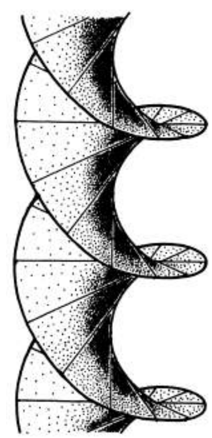
$$-\infty<z<\infty, 0<\theta<2\pi$$

تعریف می‌شود به دست می‌آوریم:

$$K=\frac{-1}{\cosh^4 z}, H=0$$



شکل ۱۵: هذلولی‌وار



شکل ۱۶: حلزونی شکل

هم‌چنین برای حلزونی شکلی که با

$$G:(u,v)\mapsto(u\cos v,u\sin v,v)$$

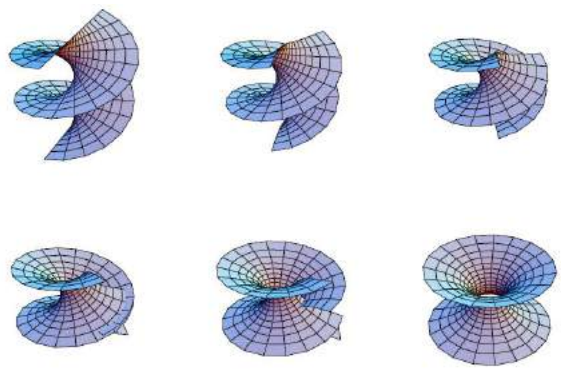
$$u>0, -\infty<v<\infty$$

تعریف می‌شود به دست می‌آوریم:

$$K=\frac{-1}{(1+u^2)^2}, H=0$$

[24] Catenoid
[25] Helicoid

محاسبات لازم برای به دست آوردن انحنای میانگین و گاوسی برای این دو رویه را می‌توانید در [۷] ببینید. در این مقاله هم‌چنین نگاشت ایزومتری بین این دو رویه ارائه شده است که تصویر آن را در شکل ۱۷ می‌بینید.



شکل ۱۷: ایزومتری تبدیل حلزونی شکل به هذلولی‌وار

قضیه‌ی زیر از ریمان، حالت خاصی از عکس قضیه‌ی گاوس را نشان می‌دهد.

قضیه ۱۰. *فرض کنید S یک رویه با انحنای گاوسی ثابت صفر باشد. در این صورت برای هر $p\in S$ همسایگی از p وجود دارد که با بازی از فضای اقلیدسی ایزومتریک است.*

در واقع قضیه‌ی کامل‌تری به شکل زیر وجود دارد که اثبات آن را می‌توانید در [۶] ببینید. مقاله‌ی [۱۰] نیز نتایج جالبی در مورد رویه‌های با انحنای صفر مطرح کرده است.

قضیه ۱۱. *برای رویه‌ی S داریم:*

- *اگر انحنای $K=0$ باشد در این صورت S به صورت موضعی با صفحه ایزومتریک است.*
- *اگر $K=1$ باشد در این صورت S به صورت موضعی با کره‌ی واحد ایزومتریک است.*
- *اگر $K=-1$ باشد در این صورت S با نیم‌صفحه‌ی بالایی[26] مجهز به متریک پوآنکاره[27] ایزومتریک است.*

[26] Upper Half Space

[27] منظور از نیم‌صفحه‌ی بالایی، مجموعه‌ی اعداد مختلط

$$\mathbb{H}=\{x+iy|y>0; x,y\in\mathbb{R}\}$$

است که به متریک پوآنکاره مجهز شده است. متریک پوآنکاره نیز به این شکل تعریف می‌شود:

$$(\mathrm{d}s)^2=\frac{(\mathrm{d}x)^2+(\mathrm{d}y)^2}{y^2}$$

رویه‌های مینیمال

در بخش قبل، بیشتر در مورد انحنای گاوسی صحبت شد. در این بخش کمی در مورد انحنای میانگین صحبت می‌کنیم و به یکی از مهمترین نتایج آن می‌پردازیم.

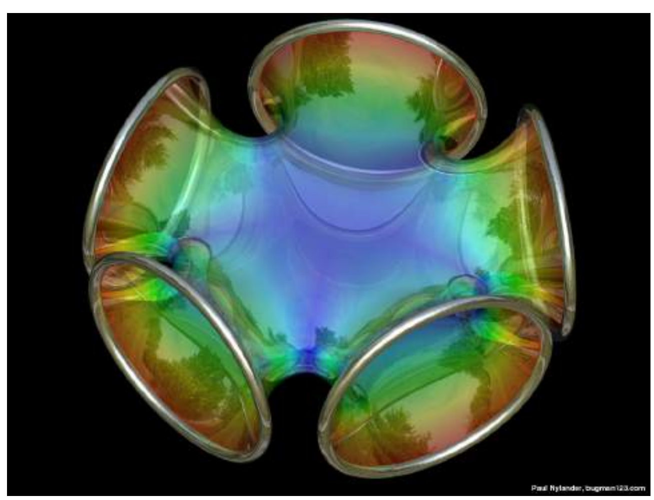
انحنای میانگین برخلاف انحنای گاوسی، جزو خواص ذاتی فضا نیست و تحت نگاشت‌های ایزومتری تغییر می‌کند. به طور مثال در مورد استوانه می‌توان این مسأله را مشاهده کرد که انحنای میانگین آن ناصفر و وابسته به شعاع مقطع استوانه است. اما همین متغیر بودن انحنای میانگین ابزاری برای پیدا کردن رویه‌های مینیمال شده است.

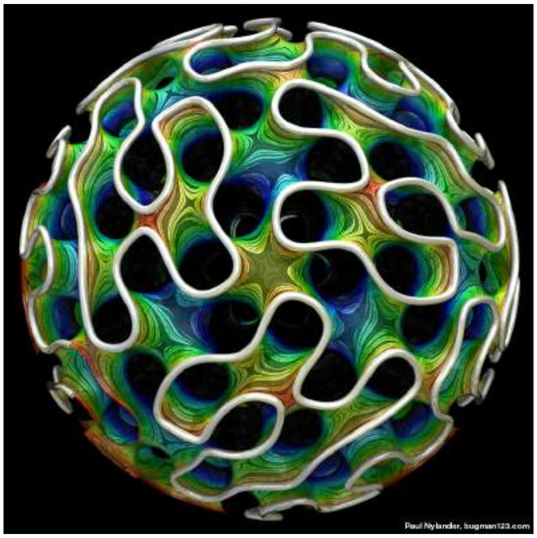
رویه‌ی $S \subset \mathbb{R}^3$ مینیمال می‌گوییم اگر برای هر نقطه‌ی $p \in S$ همسایگی از p در S وجود داشته باشد که سطح آن نسبت به مرز آن کمینه باشد. شاید این تعریف در ظاهر چندان ارتباطی به موضوع نداشته باشد اما قضیه‌ی زیر مطلب دیگری را نشان می‌دهد:

قضیه ۱۲. *رویه‌ی $S \subset \mathbb{R}^3$ مینیمال است اگر و تنها اگر انحنای میانگین آن ثابت و برابر صفر باشد.*

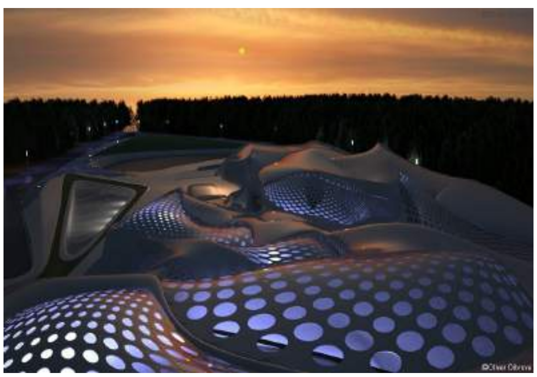
یکی از سؤالات مطرح در این حوزه، به دست آوردن یک رویه با کمترین سطح که مرز آن مشخص است، می‌باشد. از قضیه‌ی فوق کمک می‌گیرند و با تغییرات روی یک رویه‌ی اولیه که مرز آن مشخص است، آن را به رویه‌ای تبدیل می‌کنند که انحنای میانگین آن همه‌جا صفر است. با این حال جوابی که به دست می‌آید یکتا نیست و برای یک مرز مشخص ممکن است رویه‌های متفاوتی با انحنای میانگین صفر به دست آوریم.

با توجه به گفته‌های بخش قبل و قضیه‌ی ۱۱، هذلوی‌وار و حلزونی‌شکل دو رویه‌ی مینیمال هستند. شکل زیر نیز رویه‌ای مینیمال را نشان می‌دهد.





رویه‌های مینیمال در علوم مختلفی هم‌چون زیست‌شناسی[۲۸]، نظریه‌ی نسبیت[۲۹]، نظریه کوانتومی ریسمان[۳۰]، معماری[۳۱] و تکنولوژی پزشکی[۳۲] استفاده می‌شود.

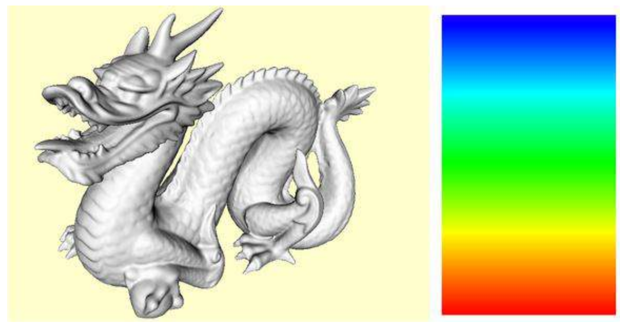


شکل ۱۸: استفاده از رویه‌های مینیمال در معماری: پروژه‌ی طراحی مرکز تلویزیون و ارتباطات برای شهر زوریخ با استفاده از رویه‌های مینیمال

[۲۸] Biology
[۲۹] Relativity Theory
[۳۰] Quantum String Theory
[۳۱] Architecture
[۳۲] Medical Technology

اژدهای استنفورد

تصویر ابتدای نوشتار خروجی برنامه‌ای است که روی اژدهای استنفورد انجام شده است و سعی در مشخص کردن انحنای این رویه دارد. رنگ قرمز انحنای منفی را نشان می‌دهد و هر چه از آن فاصله می‌گیریم و به رنگ آبی نزدیک می‌شویم انحنا افزایش می‌یابد. این کار توسط ایمان صادقی[۳۳] انجام شده است.



شکل ۱۹: طیف رنگی برای نمایش انحنای اژدهای استنفورد

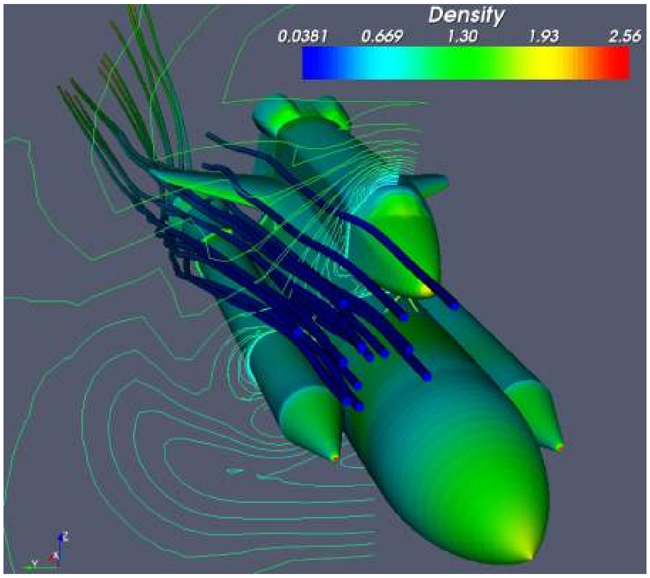
همان‌طور که قبلا نیز مطرح شد، یکی از کاربردهای انحنا در گرافیک کامپیوتری است. در این‌جا قصد نداریم به این مقوله بپردازیم و تنها به توضیح مختصری در مورد ابزاری که برای به دست آوردن این عکس استفاده شده است بسنده می‌کنیم.

ابزاری که برای این کار استفاده شده است نرم‌افزار متن بازی به نام VTK[۳۴] است که برای کارهای گرافیکی سه بعدی، پردازش تصویر و بصری‌سازی استفاده می‌شود. نرم‌افزاری که در پروژه‌های مختلفی، همچون پروژه‌ی تانگوی گوگل[۳۵] استفاده شده است. و روی سیستم عامل‌های مختلفی همچون ویندوز، لینوکس و مکینتاش قابل استفاده است.[۳۶]

برای آشنایی بیشتر از کاربردهای خم‌ها و رویه‌ها در گرافیک کامپیوتری می‌توانید از مراجع [۱۵] و [۲] استفاده کنید.

خم روی رویه

خم‌ها را در فضای اقلیدسی بررسی کردیم و دیدیم که تنها توابعی از $\mathbb{R}$ به $\mathbb{R}^n$ هستند که خواصی مانند پیوستگی را برای آن‌ها فرض می‌کنیم. با این دیدگاه می‌توانیم خم‌ها را روی رویه‌ها داشته باشیم به این شکل



شکل ۲۰: جریان سیال در اطراف شاتل فضایی. رنگ‌ها چگالی جریان را در هر نقطه مشخص می‌کند. یکی دیگر از کارهایی که توسط VTK انجام شده است.

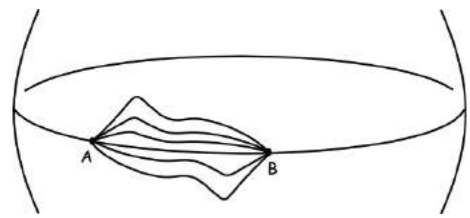
که یک خم روی رویه‌ی S تابعی پیوسته همچون $\alpha : (a, b) \to S$ است.

شکل ۲۱: نوشته‌های روی ظروف سفالی را می‌توان نمونه‌ای از خم‌های روی رویه‌ها در نظر گرفت.

حال آیا می‌توان برای این خم‌ها نیز نوعی انحنا نسبت به رویه تعریف کرد؟ دید ما نسبت به انحنای خم در فضای اقلیدسی، در واقع میزان تفاوت از خط راست یعنی کوتاه‌ترین فاصله‌ی بین دو نقطه بود. بنابراین اگر در این‌جا نیز کوتاه‌ترین فاصله‌ی بین نقاط را بدانیم، می‌توانیم برای خم روی رویه، انحنا در هر نقطه را تفاوت از این خطوطی که کوتاه‌ترین فاصله را توصیف می‌کنند، در نظر بگیریم. اما این خطوط چه هستند؟ روی یک رویه بین دو نقطه را می‌توان

[۳۳] دانشجوی دکترای علوم کامپیوتر، دانشگاه San Diego و از فارغ‌التحصیلان مهندسی کامپیوتر دانشگاه صنعتی شریف.
[۳۴] Visualization Toolkit
[۳۵] Tango Project
[۳۶] http://vtk.org/
https://www.google.com/atap/projecttango/

تنها به کمک خم‌هایی روی آن طی کرد. بنابراین کوتاه‌ترین فاصله به وسیله‌ی یکی از خم‌هایی که آن دو نقطه را به هم متصل می‌کنند حاصل می‌شود. به طور غیردقیق، به خم‌هایی که کوتاه‌ترین فاصله بین دو نقطه را دارند ژئودزیک[۳۷] می‌گوییم (واژه‌ی ژئودزیک از ژئودزی، دانش اندازه‌گیری اندازه و شکل زمین برگرفته شده است. در ابتدا، منظور از یک ژئودزیک، کوتاه‌ترین مسیر بین دو نقطه روی سطح زمین بود اما بعدها این واژه تعمیم داده شد تا اندازه‌گیری‌ها در فضاهای ریاضیاتی عمومی‌تری را نیز شامل شود[۳۸]).



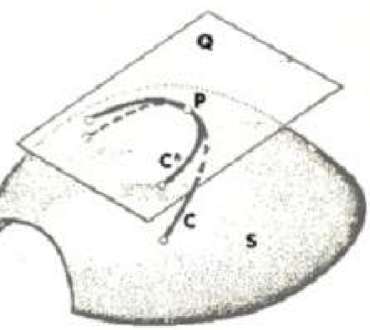
شکل ۲۲: خطی که روی دایره‌ی عظیمه A را به B وصل می‌کند ژئودزیک است چون کمترین طول را دارد. در واقع ژئودزیک‌های کره، همان دایره‌های عظیمه هستند.

از آن‌جایی که رویه‌ها را می‌توان به طور موضعی با $\mathbb{R}^2$ یکی در نظر گرفت، پس در مورد ژئودزیک‌ها می‌توان گفت که به طور موضعی اگر رویه را به $\mathbb{R}^2$ تصویر کنیم، باید بخشی از خطی راست شوند. همین تصور را می‌توان به عنوان تعریف برای انحنای خم روی رویه در نظر گرفت.

تعریف ۱۳. *فرض کنید C یک خم روی رویه‌ی S باشد. انحنای ژئودزیکی[۳۹] C در نقطه‌ی P را انحنای تصویر C روی صفحه‌ی مماس بر رویه در نقطه‌ی P تعریف می‌کنیم و آن را با κ_g نمایش می‌دهیم.*

اما آیا همواره بین دو نقطه‌ی به اندازه‌ی کافی نزدیک به هم روی رویه، ژئودزیک وجود دارد؟ پاسخ این سؤال، در واقع تعریف ما را کامل می‌کند. قضیه‌ای وجود دارد که در مورد رویه‌ی S نشان می‌دهد که به طور موضعی همواره ژئودزیکی یکتا وجود دارد. در این‌جا فقط به صورت قضیه اشاره می‌کنیم و برای جزئیات آن می‌توانید به [۱] مراجعه کنید.

قضیه ۱۴. *فرض کنید S رویه‌ای هموار باشد در $\mathbb{R}^3$ باشد. نقطه‌ی P را روی S و بردار مماس $w \neq 0$ را در نقطه‌ی P در نظر بگیرید. در این صورت عدد حقیقی $\epsilon > 0$ و ژئودزیک یکتای $\gamma : (-\epsilon, \epsilon) \to S$*



شکل ۲۳: صفحه‌ی Q صفحه‌ی مماس بر رویه در نقطه‌ی P است و C' تصویر خم C روی این صفحه. انحنای C روی S را انحنای C' در P تعریف می‌کنیم.

وجود دارد به طوری که

$$\gamma(0) = P, \gamma'(0) = w$$

روش‌ها و قضایایی برای به دست آوردن انحنای ژئودزیکی وجود دارد که در این‌جا از مطرح کردن آن صرف‌نظر می‌کنیم و می‌توانید آن‌ها را در مراجع مختلفی هم‌چون [۱] ببینید. در پایان این بخش، یکی از مهمترین قضایای هندسه دیفرانسیل را در رابطه با رویه‌ها مطرح می‌کنیم که ارتباطی بین هندسه و توپولوژی برقرار می‌کند. این قضیه به نام گاوس-بونه[۴۰] نام‌گذاری شده است. گاوس از شکلی خاص از این قضیه آگاه بود اما هیچ‌گاه آن را منتشر نکرد و بونه[۴۱] حالتی خاص از این قضیه را در سال ۱۸۴۸ منتشر کرد. در ادامه به صورت مختصر مطالب لازم برای این قضیه آمده است.

یک مثلث‌بندی برای یک رویه، پوشاندن رویه با مثلث‌هاست به طوری که این مثلث‌ها با هم تنها در یک رأس یا یک یال می‌توانند اشتراک داشته باشند. مثلث‌بندی، کاربردهای فراوانی دارد، از نمونه‌های آن می‌توان به استفاده برای ذخیره‌ی رویه‌ها در کامپیوتر اشاره کرد.

قضیه‌ای وجود دارد که می‌گوید هر رویه، یک مثلث‌بندی دارد. حال بر این اساس مشخصه‌ی اویلر[۴۲] را برای رویه‌ی S به این صورت زیر تعریف می‌کنیم.

تعریف ۱۵. *فرض کنید S یک رویه باشد و یک مثلث‌بندی برای آن ارائه شده است که این مثلث‌بندی شامل V رأس، E یال و F وجه است. در این صورت مشخصه‌ی اویلر S تعریف می‌شود:*

$$\chi(S) = V - E + F$$

[۳۷] Geodesic

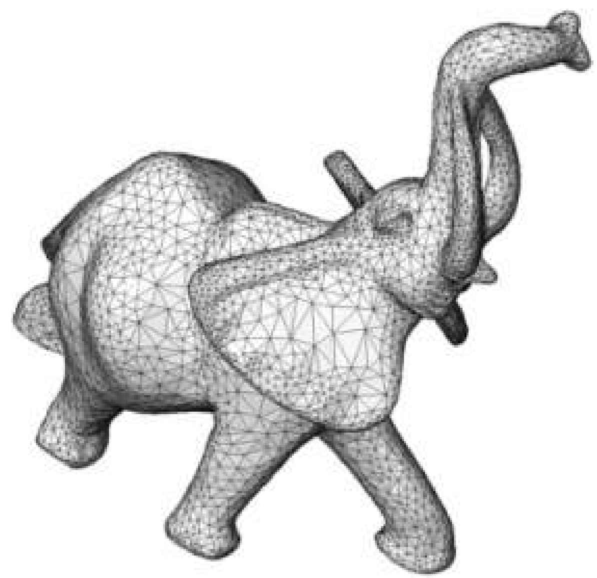
[۳۸] ویکی‌پدیا

[۳۹] Geodesic Curvature

[۴۰] Gauss-Bonnet Theorem

[۴۱] Pierre Ossian Bonnet

[۴۲] Euler Characteristic



شکل ۲۴: نمونه‌ای از مثلث‌بندی رویه

برای اینکه تعریف ابهام نداشته باشد، باید نشان داد که اگر رویه‌ای دو مثلث‌بندی متفاوت داشت آن‌گاه مشخصه‌ی اویلر هر دو مثلث‌بندی یکسان است. این مطلب نیز به صورت قضیه‌ای در توپولوژی ثابت می‌شود. هم‌چنین ثابت می‌شود که مشخصه‌ی اویلر یک ناوردای توپولوژیک است، به عبارتی تحت همسان‌ریختی‌ها ثابت باقی می‌ماند.

قضیه‌ی گاوس-بونه این دو مفهوم -انحنا (هندسه) و مشخصه‌ی اویلر(توپولوژی)- را به هم پیوند می‌دهد.

قضیه ۱۶. *فرض کنید S یک رویه‌ی فشرده با مرز ∂S باشد. اگر K را انحنای گاوسی S و κ_g را انحنای ژئودزیک ∂S در نظر بگیریم، در این صورت داریم*

$$\int_S K\mathrm{d}A + \int_{\partial S} \kappa_g \mathrm{d}s = ۲\pi\chi(S)$$

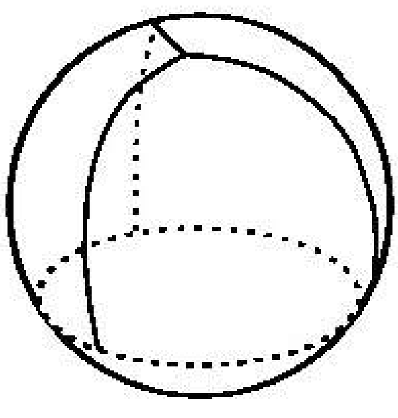
که در آن $\mathrm{d}A$ عنصر سطح و $\mathrm{d}s$ عنصر طول است.

اگر فرض کنیم S مرز نداشته باشد در این صورت $\int_{\partial S} \kappa_g \mathrm{d}s = ۰$ و بنابراین داریم:

$$\int_S K\mathrm{d}A = ۲\pi\chi(S)$$

نکته‌ای که این مطلب را جذاب‌تر می‌کند این است که برای تمامی رویه‌های فشرده و جهت‌پذیر، مشخصه‌ی اویلر به دست آمده است (در واقع تمامی این رویه‌ها دسته‌بندی شده‌اند). بنابراین برای چنین رویه‌هایی $\int_S K\mathrm{d}A$ مشخص است! و از آن‌جایی که مشخصه‌ی اویلر یک ناوردای توپولوژیک است، بنابراین با تغییر شکل‌های توپولوژیک هم‌چون پیچاندن رویه، مقدار آن ثابت و در واقع مقدار $\int_S \kappa \mathrm{d}A$ ثابت باقی می‌ماند. این مقدار را انحنای تام[۴۳] می‌گوییم. برای خم‌ها نیز انحنای تام به همین شکل تعریف می‌شود. برای خم‌های در صفحه نیز مقدار این انتگرال همواره ضریب صحیحی از ۲π است که اندیس خم[۴۴] یا عدد چرخش[۴۵] می‌گویند.

یکی از ساده‌ترین کاربردهای این قضیه را می‌توان در محاسبه‌ی مساحت سطح کره دید. شکل زیر یک مثلث‌بندی ساده از کره را نشان می‌دهد و بنابراین مشخصه‌ی اویلر برای کره ۲ = ۴ + ۶ − ۴ است.



شکل ۲۵: مثلث‌بندی کره

بنابراین چون می‌دانیم انحنای گاوسی کره ثابت و برابر با $\frac{۱}{r^۲}$ است که در آن r شعاع کره است، به دست می‌آید:

$$\int_S \frac{۱}{r^۲}\mathrm{d}A = ۴\pi$$
$$\Rightarrow \frac{۱}{r^۲}S = ۴\pi$$
$$\Rightarrow S = ۴\pi r^۲$$

اثباتی شهودی و جالب از قضیه‌ی گاوس-بونه را می‌توانید در [۱۷] ببینید.

انحنای رویه ۲

به تعریف انحنای خم بازگردیم. هر چه انحنای خم بیشتر باشد، از خط راست فاصله‌ی بیشتری دارد و بنابرین فاصله‌ی دو نقطه روی خم از فاصله‌ی خط راست بین آن دو به نسبت بیشتر می‌شود. این دیدگاه و با

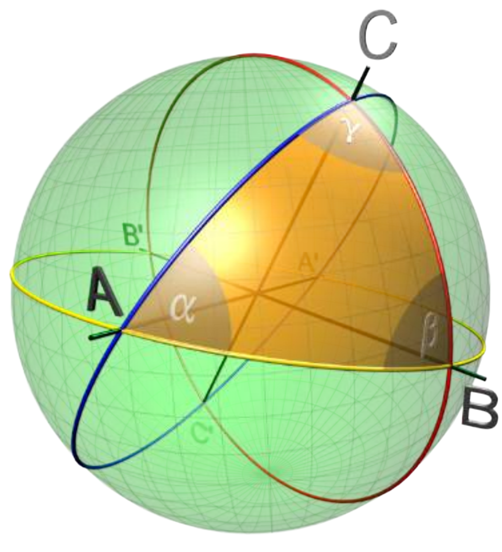
[۴۳] Total Curvature
[۴۴] Index of Curve
[۴۵] Winding Number

در نظر داشتن رویه‌های مینیمال (اینکه با یک مرز مشخص، کمترین سطح را داشته باشیم) ایده‌ی دیگری را برای انحنا در ذهن می‌آورد؛ دخیل کردن مساحت در تعریف انحنا. در واقع مقایسه‌ی مساحت اشیاء موجود روی رویه با فضای اقلیدسی. اما مساحت چه اشیائی را مقایسه کنیم؟! اولین ایده را قضیه‌ی گاوس-بونه معرفی می‌کند، مثلث! پس سعی می‌کنیم مثلث‌ها را به صورت مشخص‌تری روی رویه تعریف کنیم. نحوه‌ی انجام این کار را هم می‌توان به سادگی از فضای اقلیدسی به دست آورد و مثلث‌ها را عینا تعریف کرد با این تفاوت که در رویه با ژئودزیک‌ها سر و کار داریم و خطوط راست دیگر معنای خاصی ندارند.

تعریف ۱۷. *فرض کنید S یک رویه باشد و $\gamma_1, \gamma_2, \gamma_3$ سه ژئودزیک روی S باشند به طوری که*

$$\gamma_1(1) = \gamma_2(0), \gamma_2(1) = \gamma_3(0), \gamma_3(1) = \gamma_1(0)$$

در این صورت سه‌تایی $(\gamma_1, \gamma_2, \gamma_3)$ را با $\triangle_S(\gamma_i)_{i=1}^{3}$ نمایش می‌دهیم و آن را مثلث ژئودزیک[۴۶] *در S با اضلاع γ_1، γ_2 و γ_3 می‌گوییم.*



شکل ۲۶: یک مثلث ژئودزیک روی کره

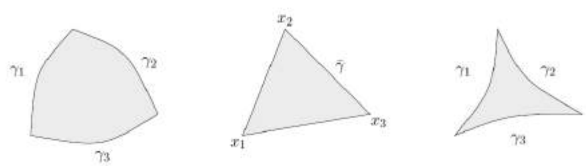
اینکه مثلث را به ژئودزیک‌ها نمایش می‌دهیم به این خاطر است که در فضاهای کلی‌تر بین دو نقطه ممکن است بیشتر از یک ژئودزیک داشته باشیم اما در فضاهایی که یکتایی برقرار است هم‌چون فضای اقلیدسی می‌توان مثلث را با نقاط نمایش دهیم.

حال که مثلث‌ها ژئودزیک را تعریف کردیم باید معیاری برای مقایسه داشته باشیم. اولین مسأله این است که آیا هر مثلث روی رویه را می‌توانیم با هر مثلث دیگر در $\mathbb{R}^2$ مقایسه کنیم؟ مسلما اگر مساحت برای ما مطرح باشد، چنین چیزی درست نیست. پس هر مثلث تنها با مثلث‌هایی خاص قابل مقایسه است که آن‌ها را مثلث‌های مشابه می‌گوییم. برای تعریف این مثلث‌ها نیز ساده‌ترین راه را پیش می‌گیریم:

تعریف ۱۸. *فرض کنید S یک رویه باشد و $\triangle_S(\gamma_i)_{i=1}^{3}$ یک مثلث ژئودزیک روی آن باشد. در این صورت مثلث $\triangle_{\mathbb{R}^2}(x_i)_{i=1}^{3}$ را در فضای اقلیدسی مشابه $\triangle_S(\gamma_i)_{i=1}^{3}$ می‌گوییم اگر*

$$d(\gamma_i(0), \gamma_i(1)) = d_{\mathbb{R}^2}(x_i, x_{i+1}), i \in \{1, 2, 3\}, x_4 = x_1$$

حال می‌توانیم به سادگی انحنای موردنظرمان را براساس این تعاریف ارائه دهیم. رویه‌ی S را با انحنای نامنفی می‌گوییم اگر مثلث‌های ژئودزیک روی S از مثلث‌ها مشابه آن مساحت بیشتری داشته باشند و دارای انحنای نامثبت می‌گوییم اگر مساحت کمتری داشته باشند.



شکل ۲۷: انحنای الکساندر-مثلث سمت چپ با انحنای نامنفی، مثلث سمت راست با انحنای نامثبت و مثلث میانی، یک مثلث مشابه با این دو مثلث در فضای اقلیدسی را نشان می‌دهد.

در واقع اگر بخواهیم تعریف فوق را به صورت دقیق‌تر بیان کنیم، به صورت زیر مطرح می‌شود:

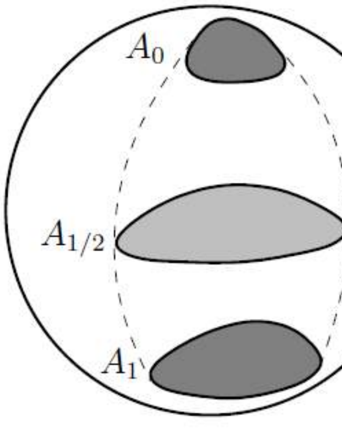
تعریف ۱۹. *می‌گوییم رویه‌ی S دارای انحنای الکساندر نامنفی است اگر برای هر مثلث ژئودزیک $\triangle_X(\gamma_i)_{i=1}^{3}$ و مثلث مشابه با آن مانند $\triangle_{\mathbb{R}^2}(x_i)_{i=1}^{3}$ داشته باشیم:*

$$\forall t \in [0, 1]; d(\gamma_1(0), \gamma_2(t)) \geq d_{\mathbb{R}^2}(x_1, \bar{\gamma}(t))$$

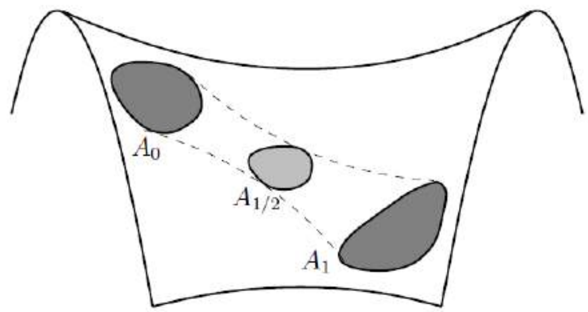
که در آن $\bar{\gamma}$ ژئودزیک بین x_2 و x_3 در فضای اقلیدسی است. مشابهاً S دارای انحنای الکساندر نامثبت است اگر نامساوی فوق جهت نامساوی عوض شود.

تعریف فوق چندان دقیق به نظر نمی‌رسد و در واقع خواسته‌های ما از انحنا را برآورده نمی‌سازد اما هم‌چنان مطالب جالبی از آن قابل استنتاج است. در واقع اگر یک سطح را بین دو ژئودزیک حرکت دهیم، از مثبت یا منفی بودن انحنا این مطلب به دست می‌آید که سطح موردنظر چگونه تغییر می‌کند.

[۴۶] Geodesic Triangle



شکل ۲۸: انحنای نامنفی.



شکل ۲۹: انحنای نامثبت.

اما مطلب مهم دیگری که در مورد انحنای الکساندر می‌توان گفت، سادگی تعریف آن و قابل گسترش بودن آن به دسته‌ی بزرگی از فضاهای ریاضی که فضاهای ژئودزیک نامیده می‌شود و حالت خاصی از فضاهای متریک هستند، می‌باشد. فضاهای ژئودزیک در واقع فضاهایی هستند که بین هر دو نقطه‌ی آن حداقل یک ژئودزیک وجود دارد.

علاوه بر بحث بالا، مثلث‌های ژئودزیک روی رویه‌ها ارتباط نزدیکی با انحنای گاوسی دارند. ارتباطی که توسط گاوس به دست آمد. گاوس ثابت کرد اگر $\triangle$ یک مثلث ژئودزیک روی رویه‌ی S با زاویه‌های α، β و γ و رئوس A، B و C باشد (شکل ۲۶)، در این صورت

$$\int_{\triangle} \kappa \mathrm{d}A = \alpha + \beta + \gamma - \pi$$

برقرار است که κ انحنای گاوسی است.

اگر بخواهیم به کمک این قضیه انحنای گاوسی در یک نقطه‌ی P را به دست آوریم، می‌توانیم مثلث ژئودزیک حول نقطه‌ی P را در هر مرحله کوچک‌تر کنیم و در نهایت با حدگیری انحنا در نقطه‌ی P به دست می‌آید. و به این شکل می‌توان گفت که انحنای الکساندر با انحنای گاوسی هم‌علامت است. در واقع قضیه‌ی زیر برقرار است:

قضیه ۲۰. *فرض کنید S یک رویه‌ی هموار باشد. در این صورت S دارای انحنای الکساندر نامنفی است اگر و تنها اگر انحنای گاوسی آن نامنفی باشد.*

در شماره‌ی بعدی به تعمیم رویه‌ها و انحنا در بعدهای بالاتر می‌پردازیم. برای این‌کار روش‌های مختلفی وجود دارد که یکی از ساده‌ترین آن‌ها ادامه‌ی همین روش و استفاده از حجم به جای مساحت است. ما نیز از همین روش استفاده می‌کنیم و با مقایسه‌ی حجم اشیاء با فضای اقلیدسی، انحنا را تعریف خواهیم کرد.

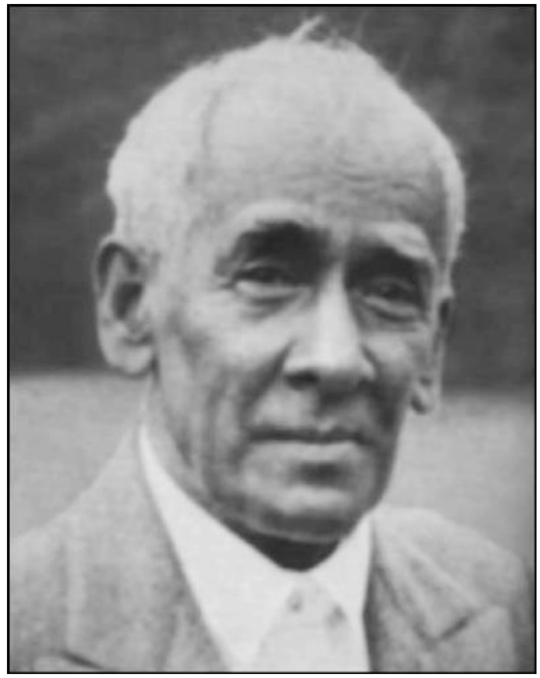
در ادامه به عنوان مطلب آخر، قضیه‌ی جذاب و زیبای چهار رأس[۴۷] را می‌آوریم و به بررسی آن می‌پردازیم. اغلب مطالب این بخش، ترجمه‌ی قسمت‌هایی از [۴] است و توصیف کاملی از این قضیه را می‌توانید در همین مقاله دنبال کنید.

قضیه‌ی چهار رأس

قضیه‌ی چهار رأس، یکی از اولین نتایج هندسه دیفرانسیل سرتاسری است که می‌گوید هر خم ساده‌ی بسته در صفحه، به جز دایره، حداقل چهار رأس -نقاطی که انحنا در آن‌ها بیشینه یا کمینه‌ی موضعی است- دارد. در سال ۱۹۰۹ سیاماداس[۴۸] این مطلب را برای خم‌های اکیدا محدب در صفحه ثابت کرد و پس از آن در ۱۹۱۲، آدولف کنزر[۴۹] قضیه را برای تمامی خم‌های ساده‌ی بسته در صفحه به دست آورد.

وارون قضیه‌ی چهار رأس به این شکل بیان می‌شود که هر تابع پیوسته حقیقی مقدار روی دایره که حداقل دو بیشینه و دو کمینه‌ی موضعی داشته باشد، تابع انحنای یک خم ساده‌ی بسته در صفحه است. در سال ۱۹۷۱، هرمان گلاک[۵۰] قضیه را برای انحنای اکیدا مثبت ثابت کرد و در نهایت، دالبرگ[۵۱] صورت کلی آن را در سال ۱۹۹۷ نتیجه گرفت اما انتشار آن به دلیل مرگ وی در سال ۱۹۹۸ به تعویق افتاد و سپس با ویرایش‌های ویلهلم[۵۲] و کرملین[۵۳] در سال

[۴۷] Four-Vertex Theorem
[۴۸] Syamadas Mukhopadhyaya
[۴۹] Adolf Kneser
[۵۰] Herman Gluck
[۵۱] Bjorn Dahlberg
[۵۲] Vilhelm Adolfsson
[۵۳] Peter Kumlin



شکل ۳۰: پروفسور سیاماداس ۱۹۳۷-۱۸۶۶

شکل ۳۱: پروفسور کنزر ۱۹۳۰-۱۸۶۲

۲۰۰۵ منتشر شد.

بیایید سعی کنیم این قضیه را نقض کنیم! برای این کار باید خم ساده‌ی بسته‌ای در صفحه مثال بزنیم که تابع انحنای آن ناثابت است و یک کمینه و یک بیشینه موضعی دارد و بنابراین بین دو کمان یکنواست (چرا حداقل یک بیشینه و یک کمینه داریم؟). بنابراین سعی می‌کنیم این‌چنین خمی را به کمک تعدادی کمان دایره بسازیم. اما این کار ممکن نیست! ببینیم چرا؟

به دایره‌های شکل ۳۳ به عنوان ریل‌های قطار نگاه کنید. خم موردنظر ما اثر قطاری است که در امتداد این ریل‌های دایره‌ای حرکت می‌کند و در نقاط مشخص شده از یکی به دیگر منتقل می‌شود.

فرض کنید قطار از نقطه‌ی ۱ در پایین بزرگترین دایره شروع می‌کند که انحنای کمینه را دارد. سپس در طور دایره‌ی بزرگ به سمت راست شروع به حرکت می‌کند و به نقطه‌ی ۲ می‌رسد و به دایره‌ی کوچک‌تر منتقل شده، انحنای آن افزایش می‌یابد.

نکته‌ی قابل توجه این است که قطر عمودی دایره‌ی کوچک‌تر در سمت راست قطر عمودی دایره‌ی بزرگ‌تر قرار می‌گیرد.

برای کوتاه شدن بحث، به قطار اجازه دهیم تا روی دایره‌ی دوم حرکت کند تا زمانی که به نقطه‌ی شماره ۳ در بالای این دایره برسد که در سمت راست قطر عمودی دایره‌ی اصلی است.

حال بازگردیم و یک بار دیگر اجازه دهیم قطار از نقطه‌ی پایین دایره‌ی اصلی که این‌بار به آن برچسب ′۱ را داده‌ایم، حرکت کند و این‌بار به سمت چپ برود. در این صورت در نقطه‌ی ′۲ به دایره‌ی کوچک‌تر منتقل می‌شود و تا نقطه‌ی ′۳ بالای آن ادامه می‌دهد که در سمت چپ قطر عمودی دایره‌ی اصلی است.

اگر خم مسیر قطار به شکل خم ساده‌ی محدب بسته باشد در این صورت باید نقاط ۳ و ′۳ منطبق شوند ... اما این اتفاق رخ نمی‌دهد! بنابراین چنین خمی وجود ندارد!

مهم نیست که از چه تعداد دایره استفاده می‌کنیم، در هر صورت به همین تناقض دچار خواهیم شد: خم نمی‌تواند بسته شود مادامی که بخواهیم ساده بماند، و بنابراین وجود نخواهد داشت.

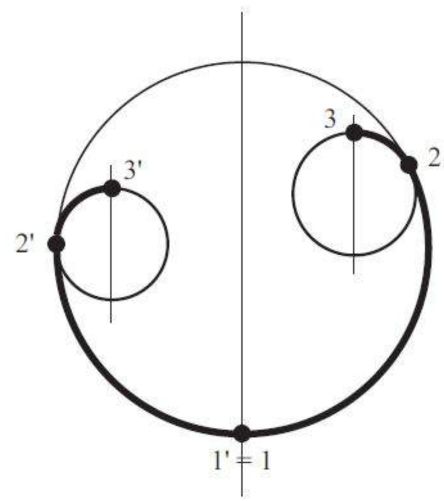
اما اگر اجازه دهیم خم خودش را قطع کند، در این‌صورت به سادگی می‌توان خمی به دست آورد که تابع انحنای آن تنها یک کمینه و یک بیشینه داشته باشد.

شکل ۳۴ خمی با معادله‌ی $r = ۱ - ۲\sin\theta$ در مختصات قطبی را نمایش می‌دهد. انحنای این خم مقدار کمینه‌ی خود در پایین دایره‌ی بزرگ دارد و مقدار بیشینه را در نقطه‌ی پایینی دایره‌ی کوچک‌تر دارد و بین این دو یکنواست.

در سال ۱۹۸۵ رابرت اوسرمان[۵۴] اثبات ساده‌ای برای قضیه‌ی چهار رأس در حالت کلی ارائه کرد. در ابتدای مقاله‌ای که اوسرمان ارائه کرد، آمده است:

[54] Robert Osserman

شکل ۳۲: پروفسور دالبرگ ۱۹۹۸–۱۹۴۹



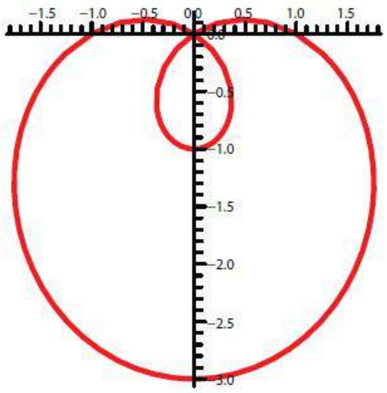
شکل ۳۳: تلاش برای به دست آوردن مثال نقض برای قضیه چهار رأس

ماهیت اثبات در یک عبارت خلاصه می‌شود: دایره‌ی محیطی را در نظر بگیرید.

فرض کنید E یک مجموعه‌ی ناتهی و فشرده (بسته و کران‌دار) در صفحه باشد. تمام دایره‌هایی که E را دربرمی‌گیرند، در نظر بگیرید. یک دایره‌ی یکتای C وجود دارد که کوچکترین دایره است. به این دایره، دایره‌ی محیطی می‌گوییم.
در مورد $E \cap C$ چه می‌توان گفت؟

۱) C باید با E اشتراک داشته باشد، زیرا در غیر این صورت می‌توان C را اندکی کوچک کرد و هم‌چنان شامل E باشد.

۲) $E \cap C$ نمی‌تواند به طور کامل در یک نیم‌دایره‌ی باز قرار بگیرد زیرا در این صورت با کمی حرکت دادن C بدون تغییر اندازه،



شکل ۳۴: یک خم با دو رأس

شکل ۳۵: مجموعه‌ی فشرده‌ی E و دایره‌ی محیطی آن.

هم‌چنان C شامل E هست اما هیچ نقطه‌ی اشتراکی ندارند و این با (۱) در تناقض است.

۳) بنابراین $E \cap C$ باید حداقل شامل دو نقطه باشد و اگر دقیقا شامل دو نقطه باشد، در این صورت آن دو نقطه روی یک قطر قرار دارند. این زمانی رخ می‌دهد که E یک بیضی باشد.

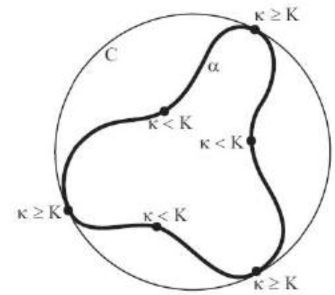
با این مقدمات، صورت قضیه‌ی اوسرمان به شکل زیر است:

قضیه ۲۱. *فرض کنید α یک خم ساده‌ی بسته از کلاس C^2 (دوبار مشتق‌پذیر با مشتق‌های پیوسته) در صفحه باشد که C دایره‌ی محیطی آن است. اگر $\alpha \cap C$ حداقل n مؤلفه داشته باشد، در این صورت α حداقل $2n$ رأس دارد.*

مراحل اصلی اثبات به شکل زیر است:

۱) فرض کنید R شعاع دایره‌ی محیطی باشد و $K = \frac{1}{R}$ انحنای آن. اگر $\alpha \cap C$، حداقل n مؤلفه‌ی داشته باشد در این صورت

α، حداقل n رأس دارد که در آن‌ها $\kappa \geq K$ و حداقل n رأس دارد که $\kappa < K$ است.



شکل ۳۶: خم α و دایره‌ی محیطی آن.

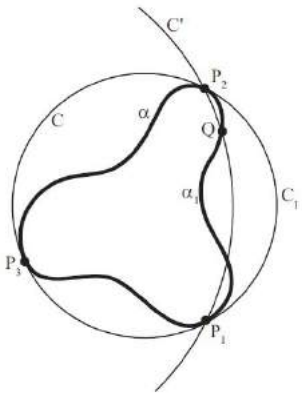
دقت کنید که فرض ما این نیست که ۶ نقطه‌ای که روی شکل ۳۶ مشخص شده‌اند، رأس‌های α هستند. اما اگر به ترتیب شش نقطه روی خم α پیدا کنیم که انحنای آن همان‌طور که دیده می‌شود زیاد و کم می‌شود، در این صورت می‌توانیم آن‌ها را با شش رأس از α که سه‌تای آن بیشینه‌ی موضعی و سه‌تای آن کمینه‌ی موضعی است، عوض کنیم.

همان‌طور که از شکل مشخص است، به سادگی می‌توان n نقطه را روی α مشخص کرد که در آن‌ها $\kappa \geq K$ است. حال برای مشخص کردن بین آن‌ها که $\kappa < K$ باشد به شکل زیر عمل می‌کنیم.

فرض کنید $P_1, P_2, \ldots, P_n$، n نقطه در خلاف جهت عقربه‌های ساعت، روی دایره‌ی C باشند که هر کدام از یک مؤلفه‌ی $C \cap \alpha$ انتخاب شده است. چون $C \cap \alpha$ نمی‌تواند به تمامی در یک نیم‌دایره‌ی باز قرار بگیرد، می‌توانیم این نقاط را به گونه‌ای انتخاب کنیم که فاصله‌ی هر دو نقطه‌ی متوالی روی C از یک نیم‌دایره بیشتر نباشد.

حال کمان C_1 از C بین P_1 و P_2 و کمان α_1 متناظر با آن را روی خم α در نظر بگیرید. برای ساده‌تر شدن، فرض می‌کنیم که خطی که P_1 و P_2 را به هم وصل می‌کند، خطی عمودی است.

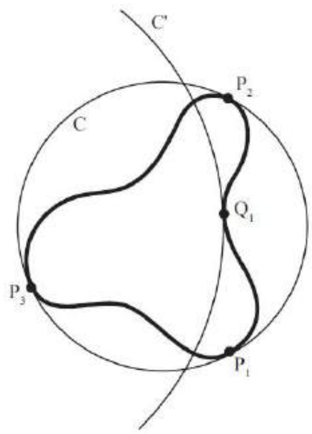
کمان α_1 از α در نقاط P_1 و P_2 بر C مماس است اما در تمام طول C این اتفاق رخ نمی‌دهد چون P_1 و P_2 در دو مؤلفه‌ی جدا از هم‌اند. بنابراین نقطه‌ی Q خارج از C_1 روی α_1 وجود دارد که سمت راست خط عمودی است که P_1 و P_2 را به هم وصل می‌کند. حال دایره‌ی را که از نقاط P_1، Q و P_2 می‌گذرد C' می‌نامیم.



شکل ۳۷: دایره‌ی C'.

چون Q درون C قرار دارد و C_1 از نیم‌دایره بزرگتر نیست، نتیجه می‌شود که دایره‌ی C' از C بزرگتر و بنابراین $K' < K$ است.

حال دایره‌ی C' را به سمت چپ به آرامی حرکت می‌دهیم تا زمانی که بر آخرین نقطه‌ی Q_1 از α_1 مماس شود.



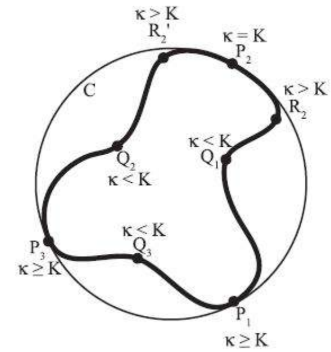
شکل ۳۸: انتقال C' به سمت چپ.

در Q_1 انحنای α، κ در شرط $\kappa(Q_1) \leq K' < K$ صدق می‌کند و این دقیقا همان چیزی است که به دنبال آن بودیم.

۲) مؤلفه‌های $\alpha \cap C$ می‌تواند تک‌نقطه‌ای یا کمانی بسته باشد. برای هر مؤلفه که به جای یک نقطه یک کمان است می‌توانیم بیشتر از دو رأس به دست آوریم بنابراین می‌توانیم بیشتر از $2n$ رأسی که قضیه‌ی اوسرمان وعده داده، داشته باشیم.

تصویر خم α که در شکل ۳۹ می‌بینید همان خم قبلی است با اندکی تغییرات که باعث شده است در نقطه‌ی P_2 بخشی از آن با دایره‌ی C یکی شده است. بنابراین انحنای α در نقطه‌ی P_2 در این‌جا دقیقا K است. بعلاوه نقاطی از α در دو طرف خم

وجود دارد، که انحنا آن از K بیشتر است. دو نمونه از این نقاط در شکل ۳۹ با R_2 و R_2' مشخص شده است. بنابراین نتیجه‌ای که حاصل می‌شود این است که بیشتر از دو رأس می‌توان از این مؤلفه به دست آورد.



شکل ۳۹: رأس‌هایی که از مؤلفه‌های کمانی حاصل می‌شود.

۳) ممکن است $\alpha \cap C$ تنها شامل یک مؤلفه باشد و آن زمانی رخ می‌دهد که این مؤلفه کمانی بزرگتر یا مساوی یک نیم‌دایره از C باشد. قضیه‌ی اوسرمان تضمین می‌کند که α حداقل دو رأس دارد و از (۲) نتیجه می‌شود بیشتر از دو رأس داریم.

۴) قضیه‌ی اوسرمان و (۲) با یکدیگر قضیه‌ی ۴ رأس را نتیجه می‌دهند.

حال به وراون قضیه‌ی چهار رأس می‌پردازیم. صورت دقیق وارون این قضیه به شکل زیر قابل بیان است که در آن منظور از S^1 همان دایره‌ی واحد است. اثبات این قضیه را می‌توانید در [۴] ببینید.

قضیه ۲۲. *فرض کنید $\kappa : S^1 \to \mathbb{R}$ تابعی پیوسته باشد که یا ثابت ناصفر است یا حداقل دو مقدار کمینه‌ی موضعی و دو بیشینه‌ی موضعی دارد. در این صورت خم ساده‌ی بسته‌ای مانند $\alpha : S^1 \to \mathbb{R}^2$ وجود دارد که برای هر $t \in S^1$ انحنای α در نقطه‌ی t برابر با $\kappa(t)$ است.*

با ثابت شدن این قضیه، مسأله‌ی چهار رأس برای خم‌های در صفحه کامل شد. اما آیا برای رویه‌های نیز می‌توان قضیه‌ی مشابهی مطرح کرد؟! شکل خاصی از این سؤال توسط محمد قمی در [۵] اثبات شده است. در این مقاله در واقع، پروفسور قمی ارتباطی بین خواص توپولوژیک و هندسی برقرار کرده است. ایده‌ی آن نیز از قضیه‌ای از پینکال[۵۵] حاصل می‌شود.

[۵۵] Ulrich Pinkall

قضیه ۲۳. *هر خم بسته در $\mathbb{R}^2$ که مرز یک رویه‌ی غوطه‌ور شده[۵۶] باشد حداقل چهار رأس دارد.* ([۱۳])

عکس‌های استفاده شده در این نوشتار تماماً از اینترنت و مراجع ذکر شده استخراج شده است. یکی از منابع مهم مورد استفاده نیز سایت ویکی‌پدیا بوده است.

مراجع

[۱] Carmo, Manfredo P. Do. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.

[۲] Cipolla, Roberto. *Active Visual Inference of Surface Shape*. Springer, 1996.

[۳] Coolidge, J. L. The unsatisfactory story of curvature. 1952.

[۴] Dennis DeTurck, Herman Gluck, Daniel Pomerleano and Vick, David Shea. The four vertex theorem and its inverse. 2007.

[۵] Ghomi, Mohammad. A riemannian four vertex theorem for surfaces with boundary. 2010.

[۶] Hitchin, Nigel. *Geometry of Surfaces*. Lecture Note, 2004.

[۷] Kitagawa, Takuya. Information in curvature.

[۸] Lodder, Jerry. Curvature in the calculus curriculum. 2003.

[۹] Margalit, Dan. The history of curvature.

[۱۰] Massey, William S. Surface of gaussian curvature zero in euclidean 3-space. 1961.

[۵۶] Immersed Surface

[۱۱] Morgan, Frank. *Riemannian Geometry, A Beginner's Guide*. Jones and Bartlett Publishers, 1993.

[۱۲] P. A. Firby, C. F. Gardiner. *Surface Topology*. Ellis Horwood, 1991.

[۱۳] Pinkall, U. On the four-vertex theorem. 1987.

[۱۴] Pressley, A. *Elementary Differential Geometry*. Springer-Verlag London, 2001.

[۱۵] Salomon, David. *Curves and Surfaces for Computer Graphics*. Springer, 2006.

[۱۶] Thomas Banchoff, Shiing-Shen Chern and Pohl, William. *Differential Geometry of Curves and Surfaces*. Springer, 2002.

[۱۷] بحرینی، علیرضا. *معماری رویه‌های دوبعدی در* $\mathbb{R}^3$ *و اثباتی شهودی بر قضیه‌ی گاوس- بونه*. فرهنگ و اندیشه‌ی ریاضی، شماره ۵۳، پاییز ۱۳۹۲.

[۱۸] شهشهانی، سیاوش. *حساب دیفرانسیل و انتگرال، جلد دوم*. انتشارات فاطمی، ۱۳۸۷.

بُعد در ریاضیات

عرفان صلواتی

بعد یک شیء را چگونه باید تعریف کرد؟ چرا ما تصور می‌کنیم صفحه ۲ بعدی و فضا ۳ بعدی است؟ چرا می‌گوییم سطح کره دو بعدی و کره توپر سه بعدی است؟

۱ مفاهیم بعد در ریاضیات

بعد یکی از مفاهیمی است که در شاخه‌های مختلف ریاضی مورد بررسی قرار گرفته و تعاریف مختلفی برای آن ارائه شده است.

در واقع بعد مفهومی ذهنی است که مانند خیلی مفاهیم دیگر، تعاریفِ ریاضی‌ای برای آن ساخته شده که تا جای ممکن با انتظار ما از مفهوم ذهنیِ آن سازگار باشد. هر کدام از ویژگی‌های بعد را که به عنوان ویژگیِ اساسی بعد در نظر بگیریم، یک تعریف از بعد به دست می‌آید. به همین دلیل است که تعاریف متعددی برای بعد در ریاضیات وجود دارد. تلاش برای صورت‌بندی‌های مناسبی از مفهوم بعد منجر به پیشرفت‌های مهمی در ریاضیات شده است.

بعد جبر خطی

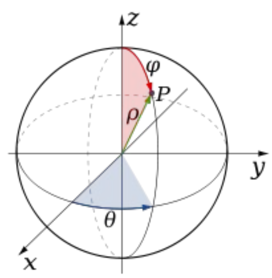
ساده‌ترین اشیائی که می‌توان از بعد آن‌ها سخن گفت، نقطه، خط، صفحه و به طور کلی زیرفضاهای آفین فضاهای اقلیدسی هستند (فرق زیرفضاهای آفین با زیرفضاهای خطی این است که زیرفضای آفین لزوماً از مبدأ نمی‌گذرد). در واقع تعریف ریاضی بعد هر چه باشد، باید طبق آن، نقطه موجودی صفر بعدی، خط موجودی یک بعدی و صفحه موجودی دو بعدی باشد.

با در نظر داشتن این ایده، تعریفی که از بعد یک شیء به ذهن می‌رسد بعدِ کوچک‌ترینِ زیرفضای آفینی است که شیء در آن قرار دارد. مثلاً مثلث، مربع و دیسک دو بعدی هستند چون در صفحه قرار دارند و چهاروجهی، مکعب و کره سه بعدی هستند چون در فضا قرار دارند. اگر چه مثلاً می‌توان مربع را نیز در فضا قرار داد اما به هر حال در یک صفحه قرار می‌گیرد. این بعد، بعد آفین نامیده می‌شود و به بیان دقیق‌تر جبر خطی، بعد آفین یک مجموعه عبارت است از بعد زیرفضای آفین تولید شده توسط آن.

این تعریف اگرچه بعضی انتظارات ما از بعد را برآورده می‌کند اما مشکلاتی هم دارد. طبق این تعریف سطح یک کره چه بعدی است؟ ما انتظار داریم سطح کره دوبعدی باشد اما با تعریف فوق سه بعدی می‌شود، زیرا سطح کره در یک صفحه قرار ندارد (در واقع به راحتی ثابت می‌شود بعد زیرفضای آفین تولید شده توسط سطح کره ۳ است). پس باید به دنبال تعریف دیگری از بعد باشیم که در آن، سطح کره دوبعدی باشد.

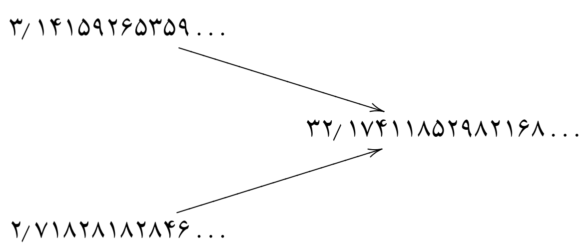
بعد خمینه‌ای

یکی دیگر از تصوراتِ شهودی ما از بعد یک مجموعه، تعداد پارامترهای مستقل لازم برای مشخص کردن نقاط آن مجموعه است. با این تعریف، سطح کره نیز دو بعدی است چون می‌توان نقاط سطح کره را با طول و عرض جغرافیایی آن در مختصات کروی مشخص کرد.



شکل ۱: مشخص کردن نقاط سطح کره با طول و عرض جغرافیایی

اما دقیق کردن این ایده و تبدیل آن به یک تعریف کارآمد اندکی کار دارد. در واقع با تعریف فعلی می‌توان نقاط سطح کره را حتی با یک عدد هم مشخص کرد، اگرچه این کار به طرز بسیار مصنوعی انجام می‌شود، اما روش کار این‌گونه است که شما ارقام نمایش اعشاری دو عدد را یکی در میان بین هم می‌چینید و یک عدد به دست می‌آورید و برعکس از روی عدد نهایی می‌توانید دو عدد اولیه را بازسازی کنید. شاید به نظر برسد که مشکل مختصات بالا، ناپیوسته بودن



شکل ۲: خلاصه کردن دو مختصه در یک مختصه

این نگاشت است، اما در واقع می‌توان حتی نگاشتی پیوسته از بازه $[۰,۱]$ به سطح کره ساخت به نحوی که همهٔ سطح کره را بپوشاند (چنین خمهایی را خمهای فضاپرکن می‌نامند).

پس برای ارائه تعریف مناسبی از بعد، باید تعریفی برای یک مختصات طبیعی ارائه کنیم. این مسیر منجر به تعریف خمینه[۱] می‌شود که در واقع تعمیم اشیائی چون سطح کره است. خمینه‌ها اشیاء اصلی مورد مطالعی در هندسه خمینه‌ها و توپولوژی دیفرانسیل هستند. برای کسب اطلاعات بیشتر در مورد خمینه‌ها، مطالعهٔ کتاب بسیار ارزشمند [۱] توصیه می‌شود.

بعد استقرایی

اکنون تعریف شهودی دیگری از بعد ارائه می‌کنیم و سعی می‌کنیم آن را دقیق کنیم. فرض کنید قطعه‌ای کاغذ را ببریم و دو قطعه کاغذ به دست بیاوریم. مرز بین دو قطعه، یک خم خواهد بود، یعنی یک مجموعهٔ یک بعدی. به طور کلی به نظر می‌رسد که برای بریدن یک مجموعه d بعدی به دو قطعه، مرز برش باید $d - ۱$ بعدی باشد.

اکنون این ایده را دقیق‌تر توضیح می‌دهیم. فرض کنید X یک مجموعه و x و y دو نقطه از آن باشند. یک زیرمجموعهٔ Y از X را جداکنندهٔ x و y می‌گوییم اگر هر مسیر پیوسته‌ای از x به y ناگزیر از Y عبور کند.

اکنون می‌توانیم به طور استقرایی بعد یک مجموعه را تعریف کنیم. ابتدا بعد مجموعه‌های متناهی از نقاط را برابر صفر تعریف می‌کنیم. سپس یک مجموعه را حداکثر d بعدی می‌گوییم اگر هر دو نقطهٔ آن یک جداکنندهٔ حداکثر $d - ۱$ بعدی داشته باشد و بعدِ مجموعه را کوچک‌ترین dای می‌گیریم که گزارهٔ بالا درست باشد.

این تعریف اگرچه به نظر کامل می‌رسد اما به مشکلاتی برمی‌خورد. مسأله زیر این موضوع را نشان می‌دهد.

مسأله ۱. ثابت کنید زیر مجموعه‌ای از صفحه وجود دارد که هر دو نقطه‌ای از صفحه را جدا می‌کند ولی شامل هیچ خم پیوسته‌ای نیست.

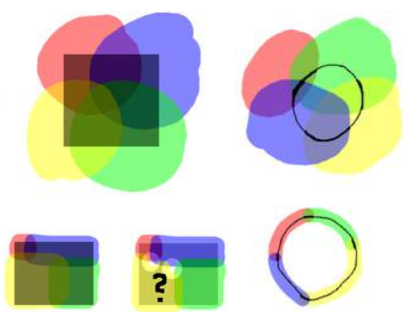
اگر چه نحوه ساخت مجموعه یاد شده خیلی غیر طبیعی است، ولی به هر حال از آنجا که این مجموعه شامل هیچ خم پیوسته‌ای نیست پس طبق تعریف ما، صفر بعدی است و در نتیجه صفحه حداکثر ۱ بعدی خواهد بود که این با انتظار ما سازگار نیست. تغییر کوچکی در تعریف فوق، منجر به تعریفی می‌شود که براوئر[۲] برای مفهوم بعد پیشنهاد کرد و بعد استقرایی[۳] نامیده می‌شود: مجموعهٔ X را حداکثر d بعدی می‌گوییم اگر برای هر دو زیر مجموعه بسته و مجزای A و B از آن، بتوان مجموعه‌های باز و مجزای U و V را یافت که مجموعهٔ $Y = X - (U \cup V)$ حداکثر $d - ۱$ بعدی باشد. در واقع Y مانند قبل یک جداکننده است با این تفاوت که اکنون زیرمجموعه‌ای بسته است. پایهٔ این تعریف استقرایی، مجموعهٔ تهی است که بعد آن -۱ تعریف می‌شود.

همانطور که مشاهده می‌شود در این تعریف فقط از مفهوم مجموعه‌های باز و بسته استفاده شده است، به همین دلیل می‌توان آن را به فضاهای توپولوژیک تعمیم داد.

بعد توپولوژیک

اکنون ایدهٔ مقدماتی دیگری در مورد بعد مطرح می‌کنیم که منجر به تعریف جدیدی از بعد می‌شود که توسط لبگ[۴] ارائه شده است. فرض کنید می‌خواهیم یک بازه از اعداد حقیقی را توسط بازه‌های کوچکتری بپوشانیم، در این صورت می‌توان دید که بازه‌های کوچکتر حتماً با یکدیگر همپوشانی خواهند داشت، اما می‌توان به نحوی این کار را انجام داد که هیچ نقطه‌ای در بیش از دو بازه قرار نگیرد.

حال فرض کنید می‌خواهیم یک مربع را با مربع‌های باز کوچکتری (یعنی مربع‌هایی که شامل لبهٔ خود نیستند) بپوشانیم. با اندکی تلاش خواهید دید که هر طور این کار را انجام دهید، نقطه‌ای یافت می‌شود که در حداقل سه تا از مربع‌های کوچک‌تر هست، و می‌توان به نحوی این کار را انجام داد که هر نقطه در حداکثر سه تا از مربع‌های کوچک‌تر باشد.



شکل ۳: راست: پوششی باز از مجموعهٔ سیاه‌رنگ و تظریفی از آن که هر نقطه در حداکثر دو باز هست، چپ: پوششی باز از مربع و تظریفی از آن که هر نقطه در حداکثر سه باز هست و تلاش ناموفق برای ساختن تظریفی که هر نقطه در حداکثر دو باز باشد

در حالت کلی، به نظر می‌رسد برای پوشاندن یک مجموعهٔ d بعدی توسط مجموعه‌های باز کوچک، ناگزیر نقطه‌هایی وجود خواهد

[۱] manifold
[۲] L. E. J. Brouwer (1881-1966)
[۳] inductive dimension
[۴] H. L. Lebesgue (1875-1941)

داشت که در حداقل $d + ۱$ تا از مجموعه‌ها هست. همین را می‌توان مبنای یک تعریف از بعد قرار داد: فضای متریک فشرده X را d بعدی می‌گوییم هرگاه برای هر δ بتوان آن را با مجموعه‌های باز به قطر حداکثر δ پوشاند به نحوی که هر نقطه در حداکثر $d + ۱$ تا از مجموعه‌ها قرار بگیرد. این تعریف قابلیت تعمیم به فضاهای توپولوژیک دلخواه را دارد، به همین دلیل آن را بعد توپولوژیک نیز می‌نامند. در بخش بعد صورت کلی این تعریف را خواهیم دید و نیز ثابت خواهیم کرد که فضای اقلیدسی $\mathbb{R}^d$ واقعاً d بعدی است.

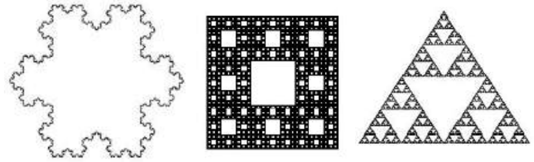
بعد فرکتالی و بعد هاوسدورف

اکنون به آخرین تعریف شهودی از بعد می‌رسیم. این تعریف با اندازهٔ اشیاء ارتباط دارد. اندازه و بعد ارتباط روشنی با هم دارند. اندازهٔ یک شیء یک بعدی عبارت است از طول آن، اندازهٔ یک شیء دو بعدی عبارت است از مساحت آن و اندازهٔ یک شیء سه بعدی عبارت است از حجم آن. ولی چگونه می‌توان از این ایده برای تعریف بعد یک شیء استفاده کرد؟ برای این کار از این واقعیت استفاده می‌کنیم که وقتی یک شیء را با ضریب ۲ تجانس بدهیم، اگر یک بعدی باشد، اندازه آن دو برابر، اگر دو بعدی باشد، اندازه آن ۴ برابر و اگر سه بعدی باشد، اندازهٔ آن ۸ برابر می‌شود.

به عنوان مثال، وقتی ضلع یک مربع را دو برابر کنیم، مربعی به دست می‌آوریم که به ۴ مربع به اندازهٔ مربع اولیه تقسیم می‌شود. با این ایده، اگر چه بُعد را تنها برای دسته محدودی از اشیاء می‌توان تعریف کرد، اما این ویژگی جالب را دارد که بُعد برخی مجموعه‌ها اعداد غیر صحیح می‌شود. با یک مثال توضیح می‌دهیم. مجموعه کانتور[۵] به این صورت تعریف می‌شود که از بازهٔ $[۰, ۱]$ شروع می‌کنیم و آن را $C_۰$ می‌نامیم سپس آن را به سه قسمت مساوی تقسیم کرده و قسمت وسط را حذف می‌کنیم و مجموعه باقی‌مانده را که متشکل از ۲ بازه به طول $\frac{۱}{۳}$ است $C_۱$ می‌نامیم. سپس هر یک از بازه‌های باقی‌مانده را مجدداً به سه قسمت برابر تقسیم می‌کنیم و قسمت‌های میانی را حذف می‌کنیم و مجموعهٔ حاصل را $C_۲$ می‌نامیم و این فرآیند را ادامه می‌دهیم. در نهایت تعریف می‌کنیم $C = \bigcap C_n$.

مجموعهٔ کانتور ویژگی‌های جالبی دارد، مجموعه‌ای ناشمارا و در عین حال با اندازهٔ صفر است. ویژگی جالب دیگر آن *خودمتشابه* بودن یا همان فرکتال بودنِ آن است. یعنی می‌توان آن را به دو قسمت افراز کرد که آن دو قسمت متشابه با خودِ مجموعهٔ کانتور هستند و فقط در ضریب $\frac{۱}{۳}$ ضرب شده‌اند. در نتیجه اگر ما مجموعهٔ کانتور را با ضریب ۳ تجانس دهیم، دو کپی از خودِ آن را به دست می‌آوریم، حال آنکه اگر مجموعهٔ کانتور، dبعدی می‌بود، باید ۳^d کپی از خودش به دست می‌آمد. پس می‌توان انتظار داشت $۲ = ۳^d$ که نتیجه می‌دهد بعد مجموعهٔ کانتور برابر است با $d = \log_۳ ۲ \approx ۰/۶۳$. با این ایده می‌توان برای بسیاری از فرکتال‌ها بعد فرکتالی را محاسبه کرد.

مسأله ۲. بعد فرکتال‌های زیر را بیابید:



شکل ۴: راست: مثلث سرپینسکی، وسط: فرش سرپینسکی، چپ: برف‌دانهٔ کخ

همانطور که گفته شد بعد فرکتالی محدودیت‌های زیادی دارد که استفاده از آن را برای مجموعه‌های غیر فرکتالی غیر ممکن می‌کند. تعریف دیگری از بعد که هم مبتنی بر مفهوم اندازه است و هم برای ردهٔ بسیار وسیعی از مجموعه‌ها قابل تعریف است، بعد هاوسدورف است که توسط هاوسدورف[۶] معرفی شده است و همانند بعد فرکتالی مقادیر غیر صحیح هم اختیار می‌کند. تعریف آن را در مسأله زیر می‌بینید:

مسأله ۳. فرض کنید X یک فضای متریک و S زیرمجموعه‌ای از آن باشد. برای هر عدد مثبت d، اندازهٔ هاوسدورف d-بعدیِ S تعریف می‌شود

$$C_d(S) := \lim_{\epsilon \to ۰} \inf_{(U_i)} \{\sum_{i=۱}^{\infty} \operatorname{diam}(U_i)^d\}$$

که inf روی همهٔ پوشش‌های باز (U_i) از S است که $\operatorname{diam} U_i < \epsilon$.

الف) ثابت کنید $۰ \leq \alpha \leq \infty$ وجود دارد که برای $d > \alpha$، $C_d(S) = ۰$ و برای $d < \alpha$، $C_d(S) = \infty$. این α را بعد هاوسدورف S می‌نامیم.

ب) ثابت کنید بعد هاوسدورف $\mathbb{R}^d$، d است.

ج) بعد هاوسدورف فرکتال‌های مسأله قبل را بیابید.

۲ بعد توپولوژیک

در این بخش مفهوم بعد توپولوژیک را که در بخش قبل معرفی شد به طور دقیق تعریف کرده و نشان می‌دهیم این مفهوم با مفهوم بعد فضاهای اقلیدسی سازگار است.

[۵] Cantor set

[۶] Hausdorff (1868-1942)

تعریف ۱. یک پوشش باز برای یک فضای توپولوژیک X یعنی خانواده‌ای از مجموعه‌های باز X مانند $\{U_\alpha\}_{\alpha\in I}$ که اجتماع آن‌ها کل فضای X بشود. منظور از یک تظریف باز از یک پوشش باز، پوششی باز مانند $\{V_\beta\}_{\beta\in J}$ است به طوری که هر V_β زیرمجموعه‌ی حداقل یکی از U_αها باشد.

تعریف ۲ (بعد توپولوژیک)**.** فضای توپولوژیک X حداکثر d بعدی گفته می‌شود اگر هر پوشش باز از X یک تظریف باز داشته باشد که هر نقطه از فضا در حداکثر $d+1$ تا از اعضای تظریف آمده باشد. بعد توپولوژیک X کوچکترین d ای است که X حداکثر d بعدی باشد.

در تعریف فوق، اگر هیچ d با خاصیت یاد شده موجود نباشد، فضا را بی‌نهایت بعدی گوییم. قضیه زیر نشان می‌دهد که بعد فضاهای اقلیدسی با تعریف بالا، با بعد طبیعی آن‌ها یکی است.

تعریف بالا برای اینکه تعریف خوبی باشد لازم است بعد فضای $\mathbb{R}^d$ را برابر با d به دست دهد. اکنون می‌خواهیم همین امر را ثابت کنیم. حکم را برای $[0,1]^d$ ثابت می‌کنیم. با تغییر کوچکی در اثبات، حالت $\mathbb{R}^d$ نیز ثابت می‌شود. اثبات را در دو طی دو قضیه زیر انجام می‌دهیم.

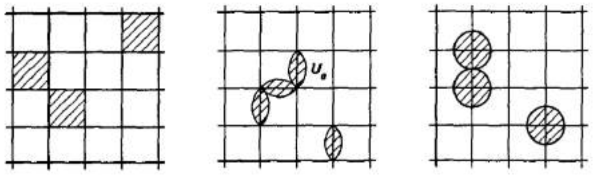
قضیه ۱. *$[0,1]^d$ حداکثر d بعدی است.*

اثبات. ابتدا حکم را برای $d=2$ ثابت می‌کنیم و سپس توضیح می‌دهیم که حالت کلی چگونه به طور مشابه به دست می‌آید. از لم زیر منسوب به لبگ استفاده می‌کنیم.

لم ۲. *اگر X فضایی فشرده و $\{U_\alpha\}_{\alpha\in I}$ پوششی باز برای آن باشد، آنگاه عدد مثبت λ وجود دارد که برای هر $x\in X$، گوی به مرکز x و شعاع λ کاملاً درون یکی از U_αها قرار می‌گیرد.*

برای دیدن اثبات این قضیه پرکاربرد رجوع کنید به [۳].

اکنون به اثبات قضیه باز می‌گردیم. فرض کنید پوشش باز $\{U_\alpha\}_{\alpha\in I}$ برای مربع $[0,1]^2$ داده شده باشد. چون مربع مجموعه‌ای فشرده است پس طبق لم قبل، عدد λ با خاصیت یاد شده وجود دارد. مربع را مطابق شکل ۲ به مربع‌های به قطر کمتر از $\frac{\lambda}{2}$ تقسیم می‌کنیم و سه خانواده از مجموعه‌های باز معرفی می‌کنیم که آن‌ها را $\mathcal{A}_1$، $\mathcal{A}_2$ و $\mathcal{A}_3$ می‌نامیم. خانوادهٔ $\mathcal{A}_1$ عبارت است از ناحیهٔ درون مربع‌ها. خانوادهٔ $\mathcal{A}_2$، عبارت است از همسایگی کوچکی از ضلع‌های مربع‌ها (به نحوی که همسایگی‌های اضلاع مجاور اشتراک ندارند) و خانوادهٔ $\mathcal{A}_3$، عبارت است از همسایگی‌های کوچکی از رئوس مربع‌ها. اجتماع همهٔ این بازها، کل $[0,1]^2$ را می‌پوشاند و با توجه به نحوه انتخاب λ هر کدام از این بازها کاملاً درون یکی از U_αها قرار می‌گیرد و از آن‌جا که اعضای هر خانواده از یک‌دیگر مجزا هستند پس هر نقطه متعلق به حداکثر ۳ باز (یکی از هر خانواده) است. پس $\mathcal{A}=\mathcal{A}_1\bigcup\mathcal{A}_2\bigcup\mathcal{A}_3$، یک تظریف باز با خاصیت مطلوب است.



شکل ۵: تظریف باز با این خاصیت که هر نقطه در حداکثر سه باز است.

پس حکم برای $d=2$ ثابت شد. برای $d>2$ نیز مشابهاً $[0,1]^d$ را به مکعب‌های به قطر کمتر از $\frac{\lambda}{2}$ تقسیم می‌کنیم و خانواده‌های $\mathcal{A}_1$ تا $\mathcal{A}_{d+1}$ را به طور مشابه تعریف می‌کنیم (در واقع $\mathcal{A}_i$ متشکل از همسایگی‌های کوچکی از وجوه $i-1$ بعدی مکعب‌ها خواهد بود). □

قضیه ۳. *$[0,1]^d$ حداقل d بعدی است.*

این بار نیز ابتدا حکم را برای $d=2$ ثابت می‌کنیم و سپس توضیح می‌دهیم که حالت کلی چگونه به طور مشابه به دست می‌آید.

از آنجا که $[0,1]^2$ از نظر توپولوژیک با مثلث هم‌ارز است پس کافی است حکم را تنها برای مثلث ثابت کنیم. مثلث متساوی‌الاضلاع به ضلع ۱ را T بنامید. پوشش دلخواهی از T با گوی‌های باز به قطر کم‌تر از $\frac{1}{2}$ در نظر بگیرید. نشان می‌دهیم این پوشش هیچ تظریفی ندارد که هر نقطهٔ آن در حداکثر دو باز باشد. در واقع ثابت می‌کنیم:

لم ۴. *برای هر پوشش از T با بازهای با قطر کمتر از $\frac{1}{2}$، نقطه‌ای از T وجود دارد که در حداقل سه باز قرار دارد.*

اثبات لم. با توجه به فشرده بودن T می‌توانیم پوشش را متناهی فرض کنیم. فرض کنید $\{U_1,\dots,U_n\}$ پوشش بازی از T باشد که قطر هر U_i کم‌تر از $\frac{1}{2}$ باشد و فرض کنید هر x در حداکثر دو تا از U_iها است. رئوس T را با ۱ و ۲ و ۳ شماره‌گذاری می‌کنیم. اکنون می‌خواهیم به هر U_i یکی از رئوس مثلث که آن را v_i می‌نامیم نسبت دهیم، این کار را به این طریق انجام می‌دهیم: اگر U_i با هیچ یک از اضلاع T اشتراک نداشت، v_i را یکی از رئوس مثلث به دلخواه می‌گیریم. اگر U_i فقط با یکی از اضلاع T اشتراک داشت، v_i را یکی از رئوس دو سر آن ضلع به دلخواه می‌گیریم و اگر U_i با دو تا از اضلاع T اشتراک داشت، v_i را رأس مشترک آن دو ضلع می‌گیریم. به راحتی می‌توان

مشاهده کرد که با توجه به فرض لم، هیچ یک از U_iها نمی‌توانند با هر سه ضلع اشتراک داشته باشند.

اکنون فرض کنید $\phi_1,\dots,\phi_n$ یک افراز واحد نسبت به $U_1,\dots,U_n$ باشند (یعنی ϕ_iها توابعی از T به $[0,1]$ هستند که $\operatorname{supp}\phi_i \subset U_i$ و $\phi_1+\cdots+\phi_n \equiv 1$. برای اثبات وجود افراز واحد رجوع کنید به [۳] بخش ۳۶).

اکنون تعریف می‌کنیم،

$$f(x)=\sum_{i=1}^{n}\phi_i(x)v_i$$

دقت کنید که f تابعی از T به T تعریف می‌کند. زیرا برای هر $x\in T$، $f(x)$ ترکیبی محدب از رئوس مثلث است و چون مثلث محدب است پس $f(x)$ نقطه‌ای از مثلث است. اما ادعا می‌کنیم که $f(x)$ در واقع روی اضلاع مثلث است. زیرا بنابر فرض، x در حداکثر دو تا از U_iها هست. اگر فقط در یکی از U_iها، مثلاً U_j باشد، آنگاه $\phi_j(x)=1$ و سایر $\phi_i(x)$ها صفر هستند. پس در این حالت $f(x)=v_j$، و اگر x در دو تا از U_iها باشد، مثلاً U_j و U_k، آنگاه سایر $\phi_i(x)$ها صفر هستند و $f(x)=\phi_j(x)v_j+\phi_k(x)v_k$. یعنی $f(x)$ ترکیبی محدب از دو تا از رأس‌های T است پس روی ضلع شامل آن دو رأس قرار دارد.

بنابراین ثابت شد که $f:T\to\partial T$ است (منظور از ∂S مرز مجموعهٔ S است). استدلال بالا حتی نشان می‌دهد که اگر x خودش روی یکی از اضلاع باشد، آنگاه $f(x)$ نیز نقطه‌ای روی همان ضلع است، زیرا U_iهایی که شامل x هستند v_i متناظر آن‌ها یکی از رئوس دو سر همان ضلع شامل x است. پس تا این‌جا ثابت شد که تابع f خواص زیر را دارد:

الف $f:T\to\partial T$ نگاشتی پیوسته است.

ب نقاط هر ضلع مثلث، توسط f به نقطه‌ای از همان ضلع نگاشته می‌شوند.

اکنون با استفاده از چنین fای یک شعبده‌بازی ترتیب می‌دهیم. می‌خواهیم لبهٔ مثلث را فقط با کشیدن و جمع کردن آن، بدون اینکه آن را پاره کنیم و بدون آینکه آن را داخل مثلث بیاوریم، در یک نقطه از خودش جمع کنیم!

به طور دقیق‌تر، قصد داریم با حرکت دادن نقاط لبهٔ مثلث به طور پیوسته و به نحوی که همواره روی اضلاع مثلث باقی بمانند همهٔ آن‌ها را در یک نقطه از لبهٔ مثلث جمع کنیم بدون اینکه لبهٔ مثلث در این فرآیند پاره شود. حرکت پیوستهٔ یک مجموعه، به زبان ریاضی با مفهوم هموتوپی بیان می‌شود.

یک هموتوپی عبارت است از یک نگاشت پیوستهٔ $F(t,x)\ :\ [0,1]\times X\ \to\ Y$ که مؤلفهٔ اول مؤلفهٔ زمان است. F را یک هموتوپی بین $g(x)=f(0,x)$ و $h(x)=f(1,x)$ می‌گوییم و می‌گوییم نگاشت‌های $g:X\to Y$ و $h:X\to Y$ با یک‌دیگر در Y هموتوپ هستند.

اکنون ما قصد داریم یک هموتوپی بین نگاشت همانی $id:\partial T\to\partial T$ و نگاشت ثابت ارائه کنیم و از این طریق به تناقض برسیم. مرکز مثلث را در مبدأ مختصات قرار می‌دهیم. تعریف می‌کنیم:

$$F(t,x)=\begin{cases} 2tf(x)+(1-2t)x & 0\leqslant t\leqslant \frac{1}{2}\\ f((2-2t)x) & \frac{1}{2}\leqslant t\leqslant 1\end{cases}$$

اولاً توجه کنید که F تابعی پیوسته است، زیرا هر ضابطهٔ آن به تنهایی پیوسته است و دو ضابطه در مقادیر مشترکشان، یعنی $t=\frac{1}{2}$ با هم برابرند.

دوماً توجه کنید که برد F در ∂T است. زیرا برای $0\le t\le\frac{1}{2}$، $F(t,x)$ ترکیب محدب x و $f(x)$ است و چون بنابر خاصیت ب در بالا، x و $f(x)$ متعلق به یک ضلع مثلث هستند، پس ترکیب محدب آن‌ها نیز متعلق به همان ضلع و در نتیجه متعلق به ∂T است. برای $\frac{1}{2}\le t\le 1$ نیز با توجه به این که برد f در ∂T است این ادعا درست است.

پس F یک هموتوپی در ∂T است. همچنین توجه کنید که $F(0,x)=x$ و $F(1,x)=f(0)$. پس نگاشت همانی در ∂T با نگاشت ثابت $g(x)=f(0)$ هموتوپ است. چنین چیزی همانطور که شهود ما می‌گوید ممکن نیست. قضیهٔ زیر این موضوع را بیان می‌کند:

قضیه ۵. *اگر D گوی بسته d بعدی باشد، نگاشت همانی $id:\partial D\to\partial D$ پوچ‌هموتوپ نیست.*

یک نگاشت را پوچ‌هموتوپ گوییم اگر با نگاشت ثابت هموتوپ باشد. دقت کنید که T از نظر توپولوژیک با گوی بستهٔ دو بعدی معادل است.

این قضیه به وسیلهٔ ابزارهای نسبتاً پیشرفتهٔ توپولوژی جبری ثابت می‌شود و نتایج بسیار جالبی دارد، از جمله قضیهٔ نقطه ثابت براوئر که بر اساس آن، هر نگاشت پیوستهٔ $f:D\to D$ دارای نقطه ثابت است. برای دیدن اثباتی از قضیه بالا در بعد ۲ با استفاده از گروه بنیادی به [۳] بخش ۵۵ رجوع کنید. همچنین برای دیدن اثباتی در حالت کلی با استفاده از گروه همولوژی به [۴] نتیجه ۲.۱۴ رجوع کنید.

پس وجود هموتوپی F ما را به تناقض می‌رساند، این تناقض نشان می‌دهد که فرض اولیهٔ ما مبنی بر وجود پوشش باز U_iها غلط بوده و به این ترتیب لم ثابت می‌شود و حکم قضیه برای $d = 2$ ثابت می‌شود. برای $d > 2$، اثبات مشابه است. تنها تفاوت این است که به جای مثلث، این بار باید T را سادک d-بعدی[7] در نظر بگیریم. □

نتیجه ۶. *$[0, 1]^d$، d بعدی است.*

مسأله ۴. بعد توپولوژیک مجموعه‌های زیر را بیابید:
مجموعهٔ کانتور، مثلث سرپینسکی، فرش سرپینسکی، برف‌دانهٔ کخ

مسأله ۵. اگر $X = Y \cup Z$ که Y و Z زیرفضاهای بسته و متناهی بعدِ X هستند آنگاه

$$\dim X = \max\{\dim Y, \dim Z\}$$

مسأله ۶. ثابت کنید گراف، یک فضای توپولوژیک ۱-بعدی است.

مسأله ۷. ثابت کنید بعد توپولوژیک یک خمینهٔ d بعدی، d است.

۳ قضایای نشاندن

قضایای نشاندن در ریاضیات اهمیت اساسی دارند. این قضایا عمدتاً حاکی از آن هستند که فضاهای با یک خاصیت مشخص را می‌توان در فضاهای ساده‌تری نشاند. نمونه‌هایی معروف از این نوع قضایا:

- قضیهٔ نشاندن ویتنی[8] در هندسه خمینه‌ها: هر خمینهٔ d بعدی را می‌توان در فضای $\mathbb{R}^{2d}$ نشاند.
- قضیهٔ یوریسون[9] در توپولوژی: هر فضای متریک جدایی‌پذیر را می‌توان در $[0, 1]^{\infty}$ نشاند.
- قضیهٔ نشاندن نش[10] در هندسه دیفرانسیل: هر خمینه ریمانی را می‌توان به طور ایزومتریک در یک فضای اقلیدسی نشاند.

در این فصل می‌خواهیم یک قضیهٔ نشاندن برای فضاهای توپولوژیک d بعدی بیان و اثبات کنیم.

قضیه ۷. *هر فضای متریک فشرده d بعدی را می‌توان در $\mathbb{R}^{2d+1}$ نشاند.*

[7] d-dimensional simplex
[8] Whitney embedding
[9] Urysohn
[10] Nash embedding

نکته. عدد $2d + 1$ در صورت این قضیه، بهترین ثابت ممکن است. یعنی فضاهای متریک فشرده d بعدی وجود دارند که نمی‌توان آنها را در $\mathbb{R}^{2d}$ نشاند. به عنوان مثال گراف‌ها فضاهای توپولوژیک ۱ بعدی هستند ولی گراف‌های غیر مسطح (مانند گراف کامل ۵ رأسی و گراف کامل دوبخشی $K_{3,3}$) را نمی‌توان در صفحه نشاند.

اثبات قضیهٔ ۷. فرض کنید X یک فضای متریک فشرده dبعدی با متر $d(x, y)$ باشد. قرار می‌دهیم $N = 2d + 1$. ایدهٔ کلی اثبات را می‌توان این‌گونه بیان کرد. هدف، به دست آوردن یک نگاشت پیوسته و یک‌به‌یک $f : X \to \mathbb{R}^N$ است (چون X فشرده است، بنابر قضیه‌ای مقدماتی در توپولوژی ([3] قضیهٔ ۲۶.۶) وارون آن نیز پیوسته است و در نتیجه یک نشاندن توپولوژیک است).

بدین منظور ابتدا نگاشت‌هایی می‌سازیم که *تقریباً* یک‌به‌یک هستند و بعد با حد گرفتن از چنین نگاشتهایی، نگاشتی یک‌به‌یک به دست می‌آوریم.

تعریف ۳. نگاشتی $f : X \to \mathbb{R}^N$ را ϵ-یک‌به‌یک گوییم است اگر $f(x) = f(y)$ نتیجه دهد $d(x, y) < \epsilon$.

فضای همهٔ نگاشت‌های پیوسته از X به $\mathbb{R}^N$ را با $\mathcal{C}(X, \mathbb{R}^N)$ نشان می‌دهیم. این فضا را به متر sup مجهز می‌کنیم. یعنی برای دو تابع f و g،

$$\rho(f, g) = \sup\{|f(x) - g(x)| : x \in X\}$$

می‌دانیم که $\mathcal{C}(X, \mathbb{R}^N)$ با این متر، یک فضای کامل است.

مجموعهٔ همهٔ نگاشته‌های پیوسته و ϵ-یک‌به‌یک $f : X \to \mathbb{R}^N$ را U_ϵ می‌نامیم. نشان خواهیم داد که U_ϵ زیرفضای باز و چگالی از $\mathcal{C}(X, \mathbb{R}^N)$ است، سپس از قضیهٔ بئر[11] نتیجه می‌شود که اشتراک

$$\bigcap_{n=1}^{\infty} U_{\frac{1}{n}}$$

ناتهی است. حال اگر f عضوی از این اشتراک باشد و $f(x) = f(y)$ آنگاه برای هر n، $d(x, y) < \frac{1}{n}$ و در نتیجه $x = y$، پس f یک‌به‌یک است و حکم ثابت می‌شود. اکنون ادعا را ثابت می‌کنیم.

الف) U_ϵ در $\mathcal{C}(X, \mathbb{R}^N)$ باز است. باید برای تابع داده‌شدهٔ f در U_ϵ، نشان دهیم $\delta > 0$ وجود دارد که گوی $B_\rho(f, \delta)$ کاملاً درون U_ϵ قرار می‌گیرد. قرار دهید

$$a = \sup\{d(x, y) : x, y \in X, f(x) = f(y)\}$$

[11] Baire Theorem، در مورد قضیهٔ بئر و کاربردهای آن، مقاله‌ای جامع در مجله ریاضی شریف، شماره پنجم به چاپ رسیده است.

از آنجا که X فشرده است، sup در عبارت بالا به max تبدیل می‌شود. و بنابر فرض ϵ−یک‌به‌یکی، داریم $a < \epsilon$. حال b را عددی دلخواه بین a و ϵ بگیرید و تعریف کنید

$$A = \{(x,y) : d(x,y) \geq b\}$$

بنابر نحوهٔ انتخاب b، تابع $|f(x)-f(y)|$، روی A مثبت است و چون f پیوسته و A فشرده است پس $|f(x)-f(y)|$ دارای مینیمم مثبتی روی A است. اکنون قرار دهید

$$\delta = \frac{۱}{۲}\min\{|f(x)-f(y)| : (x,y)\in A\}$$

ادعا می‌کنیم این δ خاصیت مورد نظر را دارد. فرض کنید $\rho(f,g) < \delta$. در این صورت هرگاه $g(x)=g(y)$، داریم

$$\begin{aligned}&|f(x)-f(y)| \leq \\ &|f(x)-g(x)| + |g(x)-g(y)| + |g(y)-f(y)| \\ &< ۲\delta\end{aligned}$$

پس بنابر نحوه انتخاب δ، داریم $(x,y)\notin A$ و در نتیجه $d(x,y) < b$. پس g نیز ϵ−یک‌به‌یک است و باز بودن U_ϵ ثابت می‌شود.

ب) U_ϵ در $\mathcal{C}(X,\mathbb{R}^N)$ چگال است. از فرض dبعدی بودن X در این بخش از اثبات استفاده می‌شود. فرض کنید نگاشت پیوستهٔ f در $\mathcal{C}(X,\mathbb{R}^N)$ و $\delta > ۰$ داده شده است. می‌خواهیم $g\in U_\epsilon$ بیابیم که $\rho(f,g) < \delta$.

به دلیل فشردگی X، ثابت $\alpha > ۰$ وجود دارد که هرگاه $d(x,y) < \alpha$ آنگاه $|f(x)-f(y)| < \frac{\delta}{۲}$، و همچنین می‌توانیم فرض کنیم $\alpha < \frac{\epsilon}{۲}$. اکنون پوشش بازی از X با گوی‌های به شعاع $\frac{\alpha}{۲}$ در نظر بگیرید. بنابر فرض d بعدی بودن X، تظریفی از این پوشش مانند $\{U_۱,\ldots,U_n\}$ وجود دارد به طوری که هر نقطه از X در حداکثر $d+۱$ تا از U_iها هست. بنابر این U_iها این خواص را دارند:

۱. $\mathrm{diam}U_i < \frac{\epsilon}{۲}$
۲. $\mathrm{diam}f(U_i) < \frac{\delta}{۲}$
۳. هر نقطهٔ X در حداکثر $d+۱$ تا از U_iها هست.

اکنون ϕ_iها را یک افراز واحد نسبت به U_iها بگیرید و برای هر i، x_i را نقطه‌ای دلخواه در U_i در نظر بگیرید. اکنون به تعریفی از جبر خطی نیاز داریم.

تعریف ۴. نقاط $z_۱$، ...، z_m در $\mathbb{R}^N$ را در *وضعیت عمومی*[۱۲] گوییم اگر برای هر $k \leq N+۱$، هیچ k نقطهٔ آن بر روی یک زیرفضای آفین $k-۲$ بعدی قرار نداشته باشند. (یعنی هیچ سه نقطه‌ای هم‌خط نباشند، هیچ چهار نقطه‌ای هم‌صفحه نباشند و ... و هیچ $N+۱$ نقطه‌ای روی یک زیر فضای آفین $N-۱$ بعدی قرار نداشته باشند.)

همانطور که از نام این تعریف برمی‌آید، در وضعیت عمومی بودنِ نقاط، یک ویژگی عمومی است و مگر در حالات خاص همواره رخ می‌دهد. به راحتی می‌توان ثابت کرد که اگر تعدادی نقطه در وضعیت عمومی نباشند، می‌توان با جابه‌جا کردن آن‌ها به مقدار به دلخواه کوچک، آن‌ها را در وضعیت عمومی قرار داد.

با توجه به این توضیح می‌توان برای نقاط $f(x_i)$در $\mathbb{R}^N$ نقاط z_i را یافت که در وضعیت عمومی باشند و علاوه بر آن $|z_i - f(x_i)| < \frac{\delta}{۲}$.

حال تعریف کنید

$$g(x) = \sum_{i=۱}^{n}\phi_i(x)z_i$$

ادعا می‌کنیم g همان تابع مطلوب است. اولاً با توجه به این‌که $\sum\phi_i(x) = ۱$،

$$\begin{aligned}g(x)-f(x) &= \sum_{i=۱}^{n}\phi_i(x)z_i - \sum_{i=۱}^{n}\phi_i(x)f(x) \\ &= \sum_{i=۱}^{n}\phi_i(x)(z_i - f(x_i)) + \sum_{i=۱}^{n}\phi_i(x)(f(x_i)-f(x))\end{aligned}$$

اکنون با توجه به نحوه انتخاب z_iها، $|z_i - f(x_i)| < \frac{\delta}{۲}$ و همچنین اگر برای یک i، $\phi_i(x)\neq ۰$، آن‌گاه $x\in U_i$ و از آن‌جا که $\mathrm{diam}f(U_i) < \frac{\delta}{۲}$ ، پس $|f(x_i)-f(x)| < \frac{\delta}{۲}$. پس هر دو جملهٔ عبارت بالا کمتر از $\frac{\delta}{۲}$ هستند و در نتیجه $|g(x)-f(x)| < \delta$. پس $\rho(g,f) < \delta$.

دوماً نشان می‌دهیم $g\in U_\epsilon$. ثابت می‌کنیم اگر $g(x)=g(y)$ آنگاه x و y متعلق به یکی از U_iها هستند و در نتیجه $d(x,y) < \frac{\epsilon}{۲}$. پس g، ϵ−یک‌به‌یک است. فرض کنید $g(x)=g(y)$. آن‌گاه

$$\sum_{i=۱}^{n}(\phi_i(x)-\phi_i(y))z_i = ۰$$

[۱۲] General position

از آن‌جا که هر x در حداکثر $d+1$ تا از U_i ها هست، پس حداکثر $d+1$ تا از $\phi_i(x)$ها و حداکثر $d+1$ تا از $\phi_i(y)$ها ناصفر هستند. پس اگر قرار دهیم $c_i = \phi_i(x) - \phi_i(y)$ آن‌گاه حداکثر $2d+2 = N+1$ تا از c_iها ناصفر هستند و از طرف دیگر

$$\sum c_i = \sum(\phi_i(x)-\phi_i(y)) = 1-1=0$$

بدون کم شدن از کلیت می‌توانیم فرض کنیم $c_1,\dots,c_k \neq 0$ و $c_{k+1},\dots,c_n = 0$ می‌نویسیم

$$0 = \sum_{i=1}^{n} c_i z_i = \sum_{i=2}^{k} c_i(z_i - z_1)$$

پس بردارهای $z_2 - z_1$، ...، $z_k - z_1$ وابسته خطی هستند مگر این‌که همهٔ c_iها صفر باشند. اما وابسته خطی بودن این بردارها به این معناست که این بردارها در یک زیرفضای $k-2$ بعدی قرار دارند و در نتیجه نقاط z_1، ...، z_k روی یک زیرفضای آفین $k-2$ بعدی قرار دارند که این با توجه به $k \le N+1$ فرض وضعیت عمومی داشتن نقاط z_i در $\mathbb{R}^N$ در تناقض است. پس همهٔ c_iها باید صفر باشند و در نتیجه برای هر i، $\phi_i(x) = \phi_i(y)$. حال یکی از iهایی را در نظر بگیرید که $\phi_i(x) \neq 0$، در نتیجه $\phi_i(y) \neq 0$. پس x و y هر دو عضو U_i هستند و حکم ثابت می‌شود.

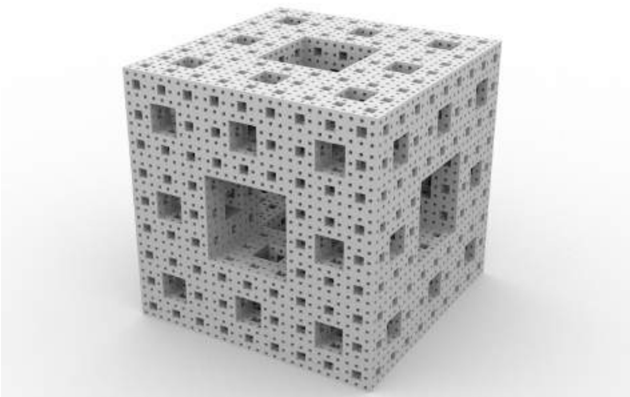
□

مسأله ۸. نشان دهید هر فضای متریک فشرده صفر بعدی را می‌توان در مجموعهٔ کانتور نشاند.

مسأله ۹. نشان دهید هر فضای متریک فشرده صفر بعدی را می‌توان در مجموعهٔ اعداد گنگ نشاند.

مسأله ۱۰. نشان دهید هر زیرمجموعهٔ فشرده یک‌بعدی از صفحه را می‌توان در فرش سرپینسکی نشاند.

مسأله ۱۱. نشان دهید هر زیرمجموعهٔ فشرده یک‌بعدی را می‌توان در اسفنج منگر[13] نشاند.



شکل ۶: اسفنج منگر

مراجع

[۱] جان و. میلنر، *توپولوژی از دیدگاه حساب دیفرانسیل*، ترجمهٔ سیاوش شهشهانی، انتشارات دانشگاه شریف (۱۳۵۸)

[2] *The Princeton companion to mathematics*. Princeton University Press, 2010.

[3] J. R. Munkres (2000) *Topology*. Second Edition.

[4] A. Hatcher, Algebraic topology, 2002.

[5] Lipscomb, S. L. *The Quest for Universal Spaces in Dimension Theory.* Notices of the AMS 56.11 (2009).

[13] Menger's sponge

نگاهی هندسی به مسأله‌ای حدی در احتمال

روزبه فرهودی

۱ چکیده

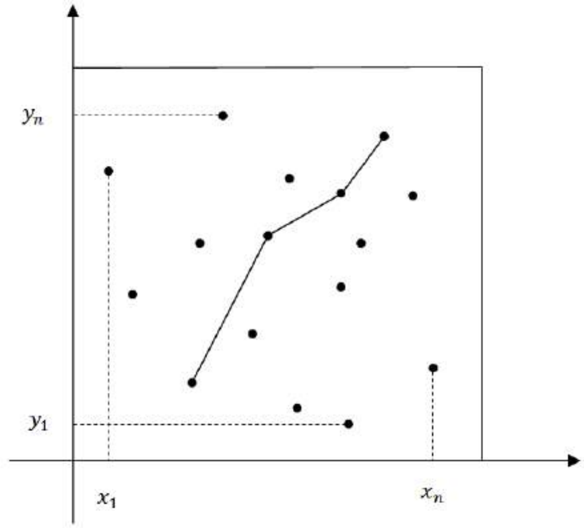
بررسی رفتارهای حدی پدیده‌های تصادفی، در احتمالات بسیار اهمیت دارد. زیرا معمولاً بررسی کمیت‌های منفرد سخت است ولی با بررسی همزمان تمام آن‌ها می‌توان امید داشت با مسأله‌ی ساده‌تری روبرو شویم. یکی از مثال‌های این موضوع بزرگترین زیردنباله‌ی صعودی یک جایگشت است. در این مقاله نشان می‌دهیم برای جایگشت‌های تصادفی می‌توان تخمین خوبی از طول بزرگترین زیردنباله‌ی صعودی بدست آورد. در بخش اول صورت مسأله و تعبیر هندسی آن را می‌گوییم و در بخش‌های بعدی محاسبات لازم انجام می‌دهیم.

۲ بزرگترین زیردنباله‌ی صعودی جایگشت تصادفی

می‌دانیم $n!$ جایگشت از اعداد ۱ تا n وجود دارد. با یک الگوریتم هندسی می‌توانیم جایگشتی را به طور یکنواخت از بین آن‌ها انتخاب کنیم: ابتدا n نقطه به تصادف و به طور یکنواخت از مربع $I = [۰,۱] \times [۰,۱]$ انتخاب می‌کنیم (شکل ۱). از آن جا که احتمال قرار گرفتن دو نقطه روی یک خط افقی یا عمودی صفر است؛ تصویر این نقاط روی هریک از محورهای x و y نیز n نقطه‌ی متمایز بدست می‌دهد. آن‌ها را به طور صعودی مرتب می‌کنیم تا به $x_۱ < x_۲ < \cdots < x_n$ و $y_۱ < y_۲ < \cdots < y_n$ برسیم. به ازای یک جایگشت، $\sigma \in S_n$ نقاط در I به شکل $\{(x_i, y_{\sigma(i)}) | ۱ \leq i \leq n\}$ هستند و به علت تقارن، جایگشت σ به طور یکنواخت از بین تمام جایگشت‌ها انتخاب شده است. به این طریق، انتخاب n نقطه در مربع به یک جایگشت یکنواخت منجر می‌شود. در این تعبیر هندسی یک زیر دنباله‌ی صعودی از این جایگشت متناظر در نظر گرفتن نقاطی است که پاره‌خط‌های واصل بین آن‌ها شیب مثبت داشته باشد. در نتیجه بزرگترین زیردنباله‌ی صعودی معادل طولانی‌ترین دنباله از خطوط شکسته و با شیب مثبت از این نقاط است.

از دید احتمالاتی کار کردن با تعبیر هندسی راحت تر است. در دهه ۶۰ میلادی اولام[۱] با شبیه سازی کامپیوتری حدس زد طول بزرگترین زیردنباله‌ی صعودی جایگشتی تصادفی از S_n، از مرتبه‌ی $۲\sqrt{n}$ است. چند سال بعد

[۱] Ulam



شکل ۱: انتخاب یک جایگشت تصادفی معادل انتخاب تعدادی نقطه در مربع است.

همرزلی[۲] با تعبیر هندسی که در بالا گفتیم نشان داد این طول از مرتبه‌ی $c\sqrt{n}$ است و کران‌هایی برای c بدست آورد[۱]. ایده‌ی او این بود که اگر در مربع به ضلع n ، $n^۲$ نقطه و در مربع به ضلع m ، $m^۲$ نقطه به تصادف قرار دهیم و آنها را از روی قطر به امتداد یکدیگر بچسبانیم طول بزرگترین خط شکسته با شیب مثبت در شکل حاصل تقریباً برابر جمع این طول برای دو شکل است. از آنجا شکل چسبانده شده قسمتی از مربع به ضلع $m+n$ است، یک نامساوی برای میانگین طول بزرگترین زیردنباله‌ی صعودی بدست می‌آوریم که مرتبه را نتیجه می‌دهد. در اینجا قصد داریم این استدلال را به طور دقیق ثابت کنیم.

۳ تقریب با فرایند پواسون

احتمالاً از درس احتمال با فرایند پواسون آشنا هستید. این فرایند برای مدل کردن زمان‌هایی که افرادی به صف وارد می‌شوند استفاده می‌شود. از آنجا که این زمان‌ها را می‌توان با نقاطی روی خط نشان داد به آن، فرایند نقطه‌ای پواسون هم می‌گویند. ویژگی اصلی آن تقارن نسبت به مکان است. یعنی به طور شهودی چگالی نقاط در طول خط ثابت است. مقدار این چگالی نرخ ورود افراد به صف را مشخص می‌کند. در این نگاه، مفهوم زمان کنار گذاشته می‌شود و به فرایند پواسون به شکل یک اندازه‌ی احتمال روی فضای تمام زیرمجموعه‌های گسسته‌ی اعداد حقیقی نگاه می‌شود. این نگاه مجرد را می‌توان به دو بعد تعمیم داد و بنابراین فرایند نقطه‌ای پواسون با نرخ یک در صفحه، فرایندی تصادفی با دو شرط است :

[۲] Hammersley

شکل ۲: جان همرزلی

۱. تعداد نقاط واقع در هر مستطیل از صفحه متغیری پواسون با نرخ مساحت آن است.

۲. تعداد نقاط دو زیر مجموعه‌ی مجزا از صفحه مستقل از هم است.

متغیر تصادفی $L_n(\sigma)$ را طول بزرگترین زیردنباله‌ی صعودی $\sigma \in S_n$ می‌گیریم و برای فرایند نقطه‌ای پواسون $\mathcal{N}$ روی کل صفحه، متغیر تصادفی $L^{\nearrow}(x,y)$ را برابر با تعداد نقاط روی بزرگترین خط شکسته‌ی با شیب مثبت در ناحیه‌ی $[0,x]\times[0,y]$ قرار می‌دهیم و برای $x \in \mathbb{R}$ تعریف می‌کنیم:

$$g(x)=\mathbb{E}(L^{\nearrow}(x,x))$$

از آن‌جا که برای فرایند $\mathcal{N}$ در مربع $[0,\sqrt{n}]\times[0,\sqrt{n}]$ تقریباً n نقطه وجود دارد، انتظار داریم $L^{\nearrow}(\sqrt{n},\sqrt{n})$ و L_n نزدیک به یکدیگر باشند. برای سادگی متغیر تصادفی $L^{\nearrow}(\sqrt{x},\sqrt{x})$ را با $L_x^{\nearrow}$ نشان می‌دهیم.

همان‌طور که در بالا گفتیم؛ همرزلی مشاهده کرد که با کنار هم قرار دادن دو خط شکسته‌ی با شیب مثبت، یکی از $(0,0)$ تا (t,t) و دیگری از (t,t) تا $(t+s,t+s)$ یک خط شکسته‌ی با شیب مثبت از $(0,0)$ تا $(t+s,t+s)$ بدست می‌آید. به علت ناوردا بودن فرایند پواسون تحت انتقال، متغیر تصادفیِ بزرگترین خط شکسته با شیب مثبت از نقطه‌ی (t,t) تا $(t+s,t+s)$ فرقی با متغیر تصادفی مشابهی که از $(0,0)$ به (s,s) می‌رسد، ندارد. در نتیجه متغیر تصادفی $\tilde{L}^{\nearrow}(s,s)$ هم‌توزیع با $L^{\nearrow}(s,s)$ وجود دارد که:

$$L^{\nearrow}(t,t)+\tilde{L}^{\nearrow}(s,s)\le L^{\nearrow}(t+s,t+s)$$

و با امید ریاضی گرفتن از طرفین:

$$g(t)+g(s)\le g(t+s) \tag{1}$$

به توابعی که برای هر s و t در رابطه ۱ صدق می‌کنند ابرجمعی می‌گوییم. لم مشهوری که در زیر می‌آوریم نشان می‌دهد رشد توابع ابرجمعی به صورت خطی است.

لم ۱ (فکت). [۳] *فرض کنید تابع $g:(0,+\infty)\longrightarrow\mathbb{R}^+$ ابرجمعی باشد و $c=\sup_t\frac{g(t)}{t}$. در این صورت:*

$$c=\lim_{t\longrightarrow+\infty}\frac{g(t)}{t}$$

اثبات. $t\in\mathbb{R}^+$ را ثابت بگیرید. برای هر $x\in\mathbb{R}^+$ عدد صحیح و نامنفی n و عدد حقیقی $0\le r<t$ یافت می‌شوند که $x=nt+r$. از ابرجمعی بودن نتیجه می‌شود:

$$ng(t)+g(r)\le g(x)$$

در نتیجه:

$$\frac{n}{nt+r}g(t)+\frac{g(r)}{x}\le\frac{g(x)}{x}$$

و بنابراین با میل دادن x به بینهایت:

$$\frac{g(t)}{t}=\liminf_{n\to+\infty}(\frac{n}{nt+r}g(t)+\frac{g(r)}{x})\le\liminf_{x\to+\infty}\frac{g(x)}{x}$$

از آن‌جا که نامساوی بالا برای هر $t\in\mathbb{R}^+$ برقرار است:

$$\limsup_{t\to+\infty}\frac{g(t)}{t}\le\liminf_{x\to+\infty}\frac{g(x)}{x}$$

بنابراین $\lim_{x\to+\infty}\frac{g(x)}{x}$ وجود دارد و برابر $c=\sup_t\frac{g(t)}{t}$ است □

از لم بالا نتیجه می‌شود $\lim_{x\to+\infty}\frac{\mathbb{E}(L_x^{\nearrow})}{\sqrt{x}}$ وجود دارد. علاوه بر این با استفاده از لم فکت برای متغیرهای تصادفی می‌توان نشان داد حد $\frac{L_x^{\nearrow}}{\sqrt{x}}$ تقریباً به ازای هر نمونه از فرایند نقطه‌ای پواسون وجود دارد[۲].

۴ وجود حد برای جایگشت‌ها

تاکنون دیدیم $\lim_{n\to+\infty}\frac{\mathbb{E}(L_n^{\nearrow})}{\sqrt{n}}$ وجود دارد. اما در مربع $[0,\sqrt{n}]\times[0,\sqrt{n}]$ الزاماً n نقطه قرار ندارد. در این قسمت نشان می‌دهیم بین دو متغیر تصادفی $L_n^{\nearrow}$ و L_n تفاوت زیادی نیست و می‌توان وجود $\lim_{n\to+\infty}\frac{\mathbb{E}(L_n)}{\sqrt{n}}$ را هم نشان داد. این کار با دو لم انجام می‌شود:

لم ۲. *برای هر k و n طبیعی:*

$$\mathbb{P}(L_n\le k)\le\mathbb{P}(L_{n-1}\le k)$$

[۳] Fekete lemma

اثبات. نگاشت $R: S_n \longrightarrow S_{n-1}$ را به این شکل تعریف می‌کنیم که ابتدا هر جایگشت $\sigma \in S_n$ را با دنباله‌ی $[\sigma(1),\sigma(2),\dots,\sigma(n)]$ نمایش می‌دهیم. حال با حذف جمله‌ای که برابر با n شده است به جایگشتی از S_{n-1} می‌رسیم.
از آنجا که R نگاشتی n به ۱ است، اندازه‌ی احتمال یکنواخت را ناوردا نگاه می‌دارد. از تعریف این نگاشت می‌توان دید

$$\forall\sigma\in S_n: \quad L_{n-1}(R(\sigma)) \le L_n(\sigma)$$

که به راحتی لم را نتیجه می‌دهد. □

لم ۳. *فرض کنید n عددی طبیعی و به اندازه‌ی کافی بزرگ باشد. k را عددی دل‌خواه در فاصله ۱ تا n بگیرید. تعریف می‌کنیم*

$$M_n = \lfloor n + 4\sqrt{n\log n}\rfloor$$

و

$$m_n = \lfloor n - 4\sqrt{n\log n}\rfloor$$

در این صورت ثابت c مستقل از n و k یافت می‌شود که:

$$\mathbb{P}(L_{M_n}\le k) - \frac{c}{n^3} \le \mathbb{P}(L_n^{\nearrow}\le k) \le \mathbb{P}(L_{m_n}\le k)+\frac{c}{n^3}$$

اثبات. با شرطی کردن روی تعداد نقاط در مربع $[0,\sqrt{n}]\times[0,\sqrt{n}]$ داریم:

$$\mathbb{P}(L_n^{\nearrow}\le k) = \sum_{t=1}^{+\infty}\frac{e^{-n}n^t}{t!}\mathbb{P}(L_t\le k) = \sum_{t=1}^{+\infty}\omega_n(t)\mathbb{P}(L_t\le k)$$

که $\omega_n(t) = \frac{e^{-n}n^t}{t!}$. سری بالا را به چهار دسته زیر تقسیم می‌کنیم:

$$\sum_{t=1}^{+\infty}\omega_n(t)\mathbb{P}(L_t\le k) = \sum_{1}^{n-4\sqrt{n\log n}} + \sum_{n-4\sqrt{n\log n}}^{n+4\sqrt{n\log n}} + \sum_{n+4\sqrt{n\log n}}^{2n} + \sum_{2n}^{\infty}$$

ثابت می‌کنیم سهم اصلی سری بالا روی جمله دوم (tهای نزدیک به n) است.

ابتدا از برای تقریب استرلینگ استفاده می‌کنیم:

$$\omega_n(t)\sim \frac{1}{\sqrt{2\pi t}}\exp(-(n+t\log\frac{t}{n}-t)) = \frac{1}{\sqrt{2\pi t}}\exp(-f(t))$$

که $f(t) = n + t\log\frac{t}{n} - t$. با محاسباتی سرراست دیده می‌شود که $f''(t) = \frac{1}{t}$ و در نتیجه f تابعی محدب است که علاوه بر آن چون تنها در نقطه $t=n$ صفر می‌شود، تابعی نامنفی نیز هست.

برای جمله‌ی چهارم سری بالا یعنی $t > 2n$ می‌توان نوشت:

$$f(t)\ge f(2n) + (t-2n)f'(2n) =$$
$$n(2\log 2 - 1) + (t-2n)\log 2 = t\log 2 - n$$

که نشان می‌دهد:

$$\sum_{2n\le t}\omega_n(t)\mathbb{P}(L_t\le k)\le \sum_{2n\le t}\omega_n(t)\le e^{-c_1 n}$$

که c_1 عددی مثبت و مستقل از n است. پس با افزایش n این سری به مقدار کافی کوچک خواهد شد.

برای جملات اول و سوم می‌دانیم $2n \le t$. چون در این فاصله مشتق دوم $f(t) - \frac{1}{4n}(t-n)^2$ بیشتر از صفر است، این تابع مثبت است و در نتیجه برای نقاطی که حداقل فاصله $4\sqrt{n\log n}$ با نقطه n دارند:

$$4\log n \le f(t)$$

پس:

$$\omega_n(t)\le\exp(-f(t))\le\frac{c_2}{n^4}$$

که c_2 نیز عددی مثبت و مستقل از n است و مانند مرحله‌ی قبل نشان می‌دهد مجموع جملات اول و سوم نیز حداکثر برابر $\frac{c_2}{n^3}$ است. در نهایت تنها جملات دوم باقی می‌مانند. تعداد آنها $4\sqrt{n\log n} = o(n)$ است می‌دانیم جمع $\omega_n(t)$های ظاهر شده در این سری نزدیک به یک است (زیرا در بالا دیدیم که جمع $\omega_n(t)$ ها برای nهای ظاهر شده در سه سری دیگر کوچک است.) و با توجه به لم ۲ $\mathbb{P}(L_t\le k)$ها نزولی هستند. حال برای رسیدن به کران پایین و بالا به ترتیب جمله‌های M_nام و m_nام را جای‌گذاری می‌کنیم و حکم ثابت می‌شود. □

حال وجود حد را برای جایگشت‌ها اثبات می‌کنیم. می‌دانیم پیشامدهای $L_t\le k$ و $L_t > k$ مکمل هم‌دیگرند. پس:

$$\mathbb{P}(L_t\le k) = 1-\mathbb{P}(L_t>k)$$

و درنتیجه لم قبل را می‌توان به صورت زیر بازنویسی کرد:

$$\mathbb{P}(L_{m_n}>k)-\frac{c}{n^3}\le\mathbb{P}(L_n^{\nearrow}>k)\le\mathbb{P}(L_{M_n}>k)+\frac{c}{n^3}$$

از آنجا که برای هر متغیر تصادفی مثبت:

$$\mathbb{E}(X)=\sum_{k\ge 0}\mathbb{P}(X>k)$$

با جمع زدن روی $k = 1,\dots,2n$ بدست می‌آوریم:

$$\mathbb{E}(L_{m_n}) - \frac{c}{n^2}\le\mathbb{E}(L_n^{\nearrow}) - \sum_{k\ge 2n}\mathbb{P}(L_n^{\nearrow}>k)\le\mathbb{E}(L_{M_n})+\frac{c}{n^2}$$

اما اگر L_n بیشتر از k باشد، باید داخل مربع $[0,\sqrt{n}]\times[0,\sqrt{n}]$ بیش از k نقطه وجود داشته باشد که این نشان می‌دهد:

$$\mathbb{P}(L_n^{\nearrow}>k)\le\frac{e^{-n}n^k}{k!}\le e^{-ck}$$

که نشان می‌دهد جمله $\sum_{k\ge 2n}\mathbb{P}(L_n^{\nearrow}>k)$ حداکثر به شکل e^{-cn} است و در نتیجه:

$$\lim_{n\to+\infty}\frac{\mathbb{E}(L_n)}{\sqrt{n}} = \lim_{n\to+\infty}\frac{\mathbb{E}(L_n^{\nearrow})}{\sqrt{n}}$$

۵ کران‌های بالا و پایین برای c

در بخش قبل دیدیم $c = \lim_{n\to+\infty} \frac{\mathbb{E}(L_n)}{\sqrt{n}}$ وجود دارد. در این بخش کران‌هایی برای c بدست خواهیم آورد و مشاهده می‌کنیم $c \neq 0, +\infty$. ابتدا کران پایینی برای c ارائه می‌دهیم. قضیه‌ی زیر نتیجه یکی از قضایای معروف درباره زنجیرها در یک مجموعه‌ی مجهز به طور جزیی مرتب است که اثبات آن در اکثر مرجع‌های ترکیبیاتی یافت می‌شود.

قضیه ۴ (اردوش-زاکر). [۴] *فرض کنید n و m دو عدد طبیعی هستند. در این صورت هر جایگشت در S_{nm+1} یا زیردنباله‌ای صعودی به طول $n+1$ و یا زیردنباله‌ای نزولی به طول $m+1$ دارد.*

اگر در این قضیه m را برابر n قرار دهیم نتیجه می‌شود هر جایگشت در S_{n^2+1} یا یک زیردنباله‌ی صعودی یا یک زیردنباله‌ی نزولی به طول n دارد. با استفاده از این موضوع ثابت می‌کنیم $c \geq \frac{1}{2}$. می‌دانیم c حد دنباله‌ی $\frac{\mathbb{E}(L_n)}{\sqrt{n}}$ است. در نتیجه با در نظر گرفتن زیردنباله‌ی $\{n^2+1\}_{n=1}^{+\infty}$ از اعداد طبیعی:

$$c = \lim_{n\to+\infty} \frac{\mathbb{E}(L_{n^2+1})}{n}$$

برای هر عدد طبیعی n، $l_n(\sigma)$ را متغیر تصادفی طول بلندترین زیردنباله‌ی نزولی σ بگیرید. از آن‌جا که با از چپ به راست نوشتن یک جایگشت، زیردنباله‌های نزولی به صعودی تبدیل می‌شوند؛ توزیع‌های دو متغیر تصادفی l_n و L_n یکی است و بنابراین $\mathbb{E}(L_n) = \mathbb{E}(l_n)$ و درنتیجه قضیه‌ی اردوش-زاکر معادل است با اینکه برای هر $\sigma \in S_{n^2+1}$:

$$n \le L_{n^2+1}(\sigma) + l_{n^2+1}(\sigma)$$

و اگر از طرفین امید ریاضی بگیریم:

$$n \le \mathbb{E}(L_{n^2+1}) + \mathbb{E}(l_{n^2+1})$$

که نتیجه می‌دهد:

$$\frac{1}{2} \le \lim_{n\to+\infty} \frac{\mathbb{E}(L_{n^2+1})}{n}$$

بنابراین $c \ge \frac{1}{2}$.
البته با استفاده از لم ۱ می‌توان کران پایین‌های دیگری یافت، زیرا هر جمله $\frac{\mathbb{E}(L_n^{\nearrow})}{\sqrt{n}}$ یک کران پایین برای c است. با قرار دادن $n = 1$ خواهیم داشت $\mathbb{E}(L_1^{\nearrow}) \le c$. ولی در مربع $[0,1] \times [0,1]$ به احتمال $\frac{1}{e}$ حداقل یک نقطه واقع است و این نشان می‌دهد که:

$$\frac{1}{e} \le c$$

با استفاده از لم زیر کران بالایی برای c بدست می‌آوریم:

لم ۵. *برای هر $k \le n$ طبیعی:*

$$\mathbb{P}(k \le L_n) \le \min\{1, \frac{1}{k!}\binom{n}{k}\}$$

اثبات. کافیست نشان دهیم $\mathbb{P}(k \le L_n) \le \frac{1}{k!}\binom{n}{k}$. اعداد

$$1 \le i_1 < i_2 < \cdots < i_k \le n$$

را ثابت فرض کنید. بنا به تقارن احتمال این‌که

$$\sigma(i_1), \sigma(i_2), \ldots, \sigma(i_k)$$

زیردنباله‌ای صعودی در σ باشد $\frac{1}{k!}$ است. از طرفی به $\binom{n}{k}$ طریق می‌توان اعداد $i_1, i_2, \ldots, i_k$ را انتخاب نمود و اگر $k \le L_n(\sigma)$ باید یکی از این زیردنباله‌ها صعودی باشد که این حکم نتیجه می‌دهد. □

می‌خواهیم از طرف راست لم بالا حد بگیریم و نشان دهیم: $c \le e$. داریم:

$$\mathbb{E}(L_n) = \sum_{1\le k\le n} \mathbb{P}(k \le L_n) \le \sum_{1\le k\le n} \min\{1, \frac{1}{k!}\binom{n}{k}\}$$
$$\le \sum_{1\le k\le n} \min\{1, \frac{n^k}{(k!)^2}\} = \sum_{1\le k\le e'\sqrt{n}} + \sum_{e''\sqrt{n}<k\le n}$$

که $e \le e' \le e''$ اعدادی نزدیک به e هستند. جملات مجموع اول را با ۱ تخمین می‌زنیم و در نتیجه مقدار آن از $e'\sqrt{n}$ کمتر می‌شود. برای جملات مجموع دوم از تقریب استرلینگ استفاده می‌کنیم:

$$\frac{n^k}{(k!)^2} \le (\frac{e'\sqrt{n}}{k})^{2k} \le (\frac{e'}{e''})^{2e''\sqrt{n}}$$

در نتیجه مجموع کل جملات سری دوم از $n(\frac{e'}{e''})^{2e'\sqrt{n}}$ کمتر است. از آنجا که این عدد از مرتبه‌ی $o(\sqrt{n})$ است، در حد، قابل چشم پوشی است. بنابراین تنها جملات سری اول اهمیت دارند که با میل دادن e'' به e نتیجه می‌شود $c \le e$.

۶ ارتباط با مسائل دیگر

همان‌طور که گفته شد، قدمت این مسأله به دهه‌ی شصت میلادی می‌رسد و بعد از اثبات همرزلی تلاش‌های بسیاری شد تا مقدار c بدست آید. در نهایت این مسأله به عنوان حالت خاصی از یک مسأله‌ی کلی‌تر توسط ورشیک[۵] و کروو[۶] اثبات شد[۳]. مسأله‌ی کلی، شکل حدی نمایش‌های گروه جایگشتی است که در حدود سال‌های ۱۹۷۰ میلادی یکی از مسائل داغ نظریه‌ی نمایش بود. ارتباط جالبی بین بزرگترین زیردنباله‌ی صعودی و نمودار یانگ که در نمایش‌های گروه جایگشتی ظاهر می‌شوند وجود دارد. به تازگی نیز پیشرفت‌های جالبی درباره‌ی ارتباط آن با مسائل دیگر شناخته شده است که مهمترین آن مسأله‌ی ماتریس‌های تصادفی است. به عنوان مثال می‌توان نشان

[۴] Erdos-Szkeres Theorem

[۵] Vershik

[۶] Kerov

داد توزیع L_n با توزیع بزرگترین مقدار ویژه‌ی یک ماتریس تصادفی گوسی یکی است. علاوه بر این کمیت $L_n - ۲\sqrt{n}$ نوساناتی دارد. نشان داده شده این نوسانات از مرتبه‌ی $O(n^{\frac{1}{6}})$ است.

مراجع

[۱] Hammersley, J.M. A Few Seedlings of Research, Proc. 6th Berkeley symp. math. statist. and probability, vol. 1, pp. 345–394, California University of California Press 1972

[۲] Michael J. Steele (2011), CBMS Lectures on Probability Theory and Combinatorial Optimization, University of Cambridge.

[۳] Quantum Probability and Spectral Analysis of graph, Springer-Verlag, New York, 2007.

الگوریتم جستجوی گراور [۱]

عباس محرابیان[۲]

۱ مقدمه

فرض کنید مسابقه‌ای به شکل زیر برگزار می‌شود: صد بسته‌ی دربسته روبروی شما قرار دارد که نود و نه‌تای آن‌ها پوچ هستند و در یکی از آن‌ها جایزه‌ای به ارزش ۲۵هزار تومان قرار دارد. شما می‌توانید هر بسته را با پرداخت هزار تومان باز کنید، و اگر جایزه در آن بود، آن را بردارید. آیا شما در این مسابقه شرکت می‌کنید؟ پس از اندکی تأمل می‌بینید که به نفع شما نیست که در این مسابقه شرکت کنید. گراور [۲] یک ریاضی‌دان هندی است که نشان داد در دنیای الگوریتم‌های کوانتوم، به نفع شماست که در این مسابقه شرکت کنید! او در حقیقت الگوریتمی کوانتومی ارائه داد که با خرج کردن حدود ۱۰هزار تومان می‌تواند بسته‌ای را که جایزه در آن است پیدا کند. در این مقاله الگوریتم جستجوی گراور را توضیح می‌دهیم، که یک الگوریتم کوانتومی[۳] است. در نیمه‌ی نخست مقاله توضیحات کلی درباره‌ی الگوریتم‌های کوانتومی می‌دهیم بدون این که وارد جزئیات مربوط به نحوه‌ی پیاده‌سازی آن‌ها و مباحث مکانیک کوانتومی شویم، و در نیمه‌ی دوم الگوریتم گراور را توضیح می‌دهیم.

۲ الگوریتم‌های کوانتومی

در ادامه منظور از الگوریتمی *متعارف*، الگوریتمی غیرکوانتومی است. همان‌طور که در الگوریتم‌های متعارف، بیت‌ها واحدهای اطلاعاتی هستند، در الگوریتم‌های کوانتومی، *کیوبیت‌ها*[۴] واحدهای اطلاعاتی هستند. محتوای هر بیت می‌تواند یکی از دو عضو مجموعه‌ی $\{۰,۱\}$ باشد. محتوای یک کیوبیت می‌تواند هر عضوی از مجموعه‌ی

$$Q_۱ = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{C}^۲ : |a|^۲ + |b|^۲ = ۱ \right\}$$

باشد. در فضاهای برداری که در این مقاله و به طور کلی در الگوریتم‌های کوانتوم با آن‌ها سروکار داریم، میدان همیشه مجموعه‌ی اعداد مختلط است. نمادهای مخصوصی برای نشان دادن اعضای پایه‌ی استاندارد $\mathbb{C}^۲$ وجود دارد: $\begin{pmatrix} ۱ \\ ۰ \end{pmatrix}$ را با $|۰\rangle$ و $\begin{pmatrix} ۰ \\ ۱ \end{pmatrix}$ را با $|۱\rangle$ نشان می‌دهیم. نمادهای $\rangle$ و $|$ در این‌جا فقط مشخص می‌کنند که با یک عضو $Q_۱$ سروکار داریم که به شکل یک بردار ستونی است. در ابتدای یک الگوریتم، می‌توان هر یک از کیوبیت‌ها را با یکی از دو مقدار $|۰\rangle$ یا $|۱\rangle$ مقداردهی اولیه کرد.

یک الگوریتم متعارف را می‌توان به شکل یک مدار منطقی نشان داد که از تعدادی گیت[۵] مانند not، and، و or تشکیل شده است. ورودی هر گیت یک یا دو بیت و خروجی آن هم یک بیت است. الگوریتم‌های کوانتومی را هم می‌توان به همین صورت نشان داد. دو نوع گیت کوانتومی وجود دارد: نوع اول در ساده‌ترین حالت یک کیوبیت ورودی و یک کیوبیت خروجی دارد. فرض کنید U یک عملگر یکه[۶] دلخواه روی فضای برداری $\mathbb{C}^۲$ باشد. در این صورت می‌توان یک گیت کوانتومی متناظر با U در نظر گرفت که برای هر ورودی $|q\rangle \in Q_۱$، خروجی آن $U|q\rangle$ می‌باشد. توجه کنید که بی‌نهایت گونه گیت کوانتومی از نوع اول داریم.

نوع دوم، گیت *اندازه‌گیری*[۷] نام دارد، که ورودی آن یک کیوبیت و خروجی آن یک بیت است. فرض کنیم یک کیوبیت داریم که در حالت

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\,|۰\rangle + b\,|۱\rangle$$

قرار دارد. اگر این کیوبیت به یک گیت اندازه‌گیری وارد شود، خروجی آن به احتمال $|a|^۲$ مقدار ۰ و به احتمال $|b|^۲$ مقدار ۱ خواهد داشت. وقتی می‌گوییم یک کیوبیت را اندازه‌گیری می‌کنیم، منظور این است که یک گیت اندازه‌گیری سر راه آن قرار می‌دهیم. توجه کنید که فقط یک گونه گیت کوانتومی از نوع دوم داریم.

ضرب تانسوری[۸] دو ماتریس را با نماد $\otimes$ نشان می‌دهیم؛ مثلاً داریم

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}.$$

یک n-کیوبیت واحد اطلاعاتی بزرگ‌تری است که محتوای آن

[۱] Grover's search algorithm

[۲] نگارنده در سال ۱۳۸۸ مدرک کارشناسی خود را در دو رشته مهندسی کامپیوتر و ریاضیات از دانشگاه صنعتی شریف اخذ نمود و در حال حاضر دانشجوی دکتری دانشگاه واترلوی کاناداست. نشانی ای‌میل: amehrabi@uwaterloo.ca.

[۳] quantum algorithm

[۴] qubits

[۵] gate

[۶] unitary operator

[۷] measurement gate

[۸] tensor product

عضوی از مجموعه‌ی

$$Q_n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_{2^n} \end{pmatrix} \in \mathbb{C}^{2^n} : |x_1|^2 + |x_2|^2 + \cdots + |x_{2^n}|^2 = 1 \right\}$$

است. دقت کنید که در حقیقت، یک کیوبیت یک ۱-کیوبیت است (و از نظر فیزیکی، یک n-کیوبیت چیزی نیست جز n کیوبیت خاص و شماره‌گذاری‌شده که به صورت قراردادی به آن‌ها به شکل یک گروه منسجم نگاه می‌شود). فرض کنید n کیوبیت داشته باشیم که دارای مقادیر $|q_1\rangle, |q_2\rangle, \ldots, |q_n\rangle$ باشند. در این صورت از کنار هم قرار دادن این کیوبیت‌ها یک n-کیوبیت به دست می‌آید که محتوای آن $|q_1\rangle \otimes |q_2\rangle \otimes \cdots \otimes |q_n\rangle$ است که به اختصار آن را با $|q_1 q_2 \ldots q_n\rangle$ نشان می‌دهیم. این نمادگذاری نتیجه‌ی جالبی دارد. هر عدد صحیح مانند m بین ۰ و $2^n - 1$ را می‌توان به صورت یک رشته به طول n از ۰ها و ۱ها نمایش داد (که نمایش دودویی خوانده می‌شود) که آن را با $b(m)$ نشان می‌دهیم. در این صورت می‌توان وارسی کرد که $\{|b(i)\rangle : 0 \le i \le 2^n - 1\}$ همان پایه‌ی استاندارد $\mathbb{C}^{2^n}$ است. مثلاً $\{|111\rangle, |110\rangle, |101\rangle, |100\rangle, |011\rangle, |010\rangle, |001\rangle, |000\rangle\}$ پایه‌ی استاندارد $\mathbb{C}^8$ است. دقت کنید که همواره از ضرب تانسوری n کیوبیت یک n-کیوبیت به دست می‌آید، ولی یک n-کیوبیت را لزوماً نمی‌توان به صورت ضرب تانسوری n کیوبیت نوشت. این یکی از ویژگی‌های عجیب مکانیک کوانتومی است و به پدیده‌ی entanglement مرتبط است، که بحث درباره‌ی آن از حوصله‌ی این مقاله خارج است.

دو نوع گیتی که برای کیوبیت‌ها تعریف کردیم، به صورت طبیعی برای n-کیوبیت‌ها نیز قابل تعریف هستند. فرض کنید U یک عملگر یکه روی فضای برداری $\mathbb{C}^{2^n}$ باشد. در این صورت متناظر با U یک گیت کوانتومی داریم که ورودی و خروجی‌اش هر دو n-کیوبیت هستند و همانند حالت $n = 1$، برای هر ورودی $|q\rangle \in Q_n$، خروجی آن $U|q\rangle$ می‌باشد. اندازه‌گیری یک n-کیوبیت به ما n بیت می‌دهد. فرض کنید $|q\rangle$ یک n-کیوبیت باشد که در پایه‌ی استاندارد به صورت

$$|q\rangle = \sum_{i=0}^{2^n-1} \alpha_i |b(i)\rangle$$

نمایش داده شود. در این صورت به احتمال $|\alpha_i|^2$ حاصل اندازه‌گیری $b(i)$ خواهد بود.

۳ الگوریتم جستجوی گراور

فرض کنید $f : X \to \{0,1\}$ تابعی دلخواه باشد به طوری که $f^{-1}(1)$ ناتهی است. هدف پیدا کردن عنصری در $f^{-1}(1)$ است، که مجموعه‌ی *جواب‌ها* نامیده می‌شود. برای سهولت در نمادگذاری فرض می‌کنیم $X = \{0,1\}^n$. طبیعت تابع f بر الگوریتم پوشیده است و در این‌جا فرض می‌کنیم تابع f به صورت یک *جعبه‌ی سیاه* به الگوریتم داده شده است و الگوریتم تنها می‌تواند مقدار تابع f را در نقاط دلخواهی از دامنه‌اش بپرسد. هدف این است که با کم‌ترین پرسش از این جعبه‌ی سیاه، جوابی را پیدا کند. واضح است که هر الگوریتم متعارف برای پیدا کردن جواب در حالت کلی به $\Omega(2^n)$ پرسش نیاز دارد، ولی گراور [۲] الگوریتمی کوانتومی ارائه داد که در صورتی که تعداد جواب‌ها معلوم باشد، با $O\left(2^{n/2}\right)$ پرسش از جعبه‌ی سیاه، جوابی برای مساله پیدا می‌کند.

حال توضیح می‌دهیم که جعبه‌ی سیاه متناظر با f چگونه کار می‌کند. از نماد $\oplus$ برای نشان دادن xor دو عنصر در $\{0,1\}$ استفاده می‌کنیم (که همان جمع در مبنای دو است). یک گیت $(n+1)$-کیوبیتی به نام F تعریف می‌کنیم که یک عملگر یکه روی $\mathbb{C}^{2^n}$ است و نقش جعبه‌ی سیاه را بازی می‌کند. برای تعریف F کافی است اثر آن را روی اعضای پایه مشخص کنیم. به ازای هر $x_1, x_2, \ldots, x_n, y \in \{0,1\}$،

$$F|x_1x_2\ldots x_n y\rangle = |x_1x_2\ldots x_n\rangle \otimes |y \oplus f(x_1,x_2,\ldots,x_n)\rangle \tag{1}$$

در نگاه اول ممکن است نحوه‌ی کار این گیت کمی عجیب به نظر برسد. در حقیقت این گیت تنها مقدار $f(x_1,x_2,\ldots,x_n)$ را با کیوبیت آخر xor می‌کند، و از این طریق مقدار f را در این نقطه به ما می‌دهد. به یاد آورید که گیت‌های کوانتومی (به جز گیت‌های اندازه‌گیری) عملگرهای یکه هستند و لذا استفاده از گیت‌هایی مثل F به عنوان جعبه‌ی سیاه در الگوریتم‌های کوانتوم اجتناب‌ناپذیر است. الگوریتم گراور از $O\left(2^{n/2}\right)$ کپی از گیت F استفاده می‌کند، و به احتمال بیش از یک‌سوم جوابی برای مساله پیدا می‌کند. (شایان ذکر است که تابع f یک تابع قطعی[۹] است و الگوریتم مورد استفاده هم عمدتاً قطعی است و تنها جایی که تصادف وارد الگوریتم می‌شود انتهای الگوریتم و موقع استفاده از گیت‌های اندازه‌گیری است.) با تکرار این الگوریتم می‌توان احتمال موفقیت را به دلخواه زیاد کرد.

اینک الگوریتم گراور را توصیف می‌کنیم. تعریف کنید $N = 2^n$ و

$$|\mathbf{h}\rangle = \sum_{\mathbf{x}\in\{0,1\}^n} |\mathbf{x}\rangle / \sqrt{N}\,.$$

دقت کنید که $|\mathbf{h}\rangle \in Q_n$. بردار صفر در فضای $\mathbb{C}^{2^n}$ را با $\mathbf{O}$ نشان می‌دهیم. گیت $(n+1)$-کیوبیتی R را چنین تعریف می‌کنیم: به

[۹] deterministic

ازای یک ورودی مثل $|\mathbf{x}\rangle \in Q_n, |y\rangle \in Q_1$، فرض کنیم $|\mathbf{z}\rangle \in Q_n$ حاصل بازتاب $|\mathbf{x}\rangle$ حول خط $\mathbf{O}\,|\mathbf{h}\rangle$ باشد (در این‌جا و در ادامه‌ی مقاله، منظور از $\mathbf{O}\,|\mathbf{h}\rangle$ خط گذرنده از نقاط $\mathbf{O}, |\mathbf{h}\rangle \in \mathbb{C}^{2^n}$ است). در این صورت،

$$R\left(|\mathbf{x}\rangle \otimes |y\rangle\right) = |\mathbf{z}\rangle \otimes |y\rangle \ . \tag{۲}$$

تعریف کنید $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2} \in Q_1$. الگوریتم گراور از یک $(n+1)$-کیوبیت که در حالت اولیه‌ی $|\mathbf{h}\rangle \otimes |-\rangle$ است شروع می‌کند و عملگرهای F و R را یکی‌درمیان k بار روی آن اعمال می‌کند تا به حالت $(RF)^k(|\mathbf{h}\rangle \otimes |-\rangle)$ برسد. مقدار دقیق k پایین‌تر تعیین خواهد شد. الگوریتم سپس $(n+1)$-کیوبیت حاصل را اندازه‌گیری می‌کند تا به $n+1$ بیت $x_1, x_2, \ldots, x_n, y$ برسد. نشان می‌دهیم می‌توان k را طوری انتخاب کرد که $k = O(\sqrt{N})$ و به احتمال بیش از یک‌سوم داشته باشیم $f(x_1, x_2, \ldots, x_n) = 1$، یعنی $(x_1, x_2, \ldots, x_n)$ جوابی از مساله باشد.

حالا به تحلیل این الگوریتم می‌پردازیم. تعریف می‌کنیم

$$A = f^{-1}(1),\ B = f^{-1}(0),\ \alpha = |A|,\ \beta = |B| \, .$$

هم‌چنین تعریف می‌کنیم

$$|\mathbf{a}\rangle = \sum_{\mathbf{x}\in A} |\mathbf{x}\rangle / \sqrt{\alpha},\ |\mathbf{b}\rangle = \sum_{\mathbf{x}\in B} |\mathbf{x}\rangle / \sqrt{\beta} \, .$$

ملاحظه کنید که $|\mathbf{a}\rangle$ و $|\mathbf{b}\rangle$ دو بردار عمود بر هم و یکه در $\mathbb{C}^{2^n}$ هستند. فرض کنید $\mathcal{S} \subseteq \mathbb{C}^{2^n}$ زیرفضای خطی تولید شده توسط این دو بردار در $\mathbb{C}^{2^n}$ باشد که در حقیقت یک صفحه است.

لم ۱. *فرض کنید $|\mathbf{s}\rangle, |\mathbf{t}\rangle \in \mathcal{S}$ به طوری که $|\mathbf{t}\rangle$ حاصل بازتاب $|\mathbf{s}\rangle$ حول خط $\mathbf{O}\,|\mathbf{b}\rangle$ باشد. در این صورت،*

$$F(|\mathbf{s}\rangle \otimes |-\rangle) = |\mathbf{t}\rangle \otimes |-\rangle \, .$$

اثبات. چون F عملگری خطی است، کافی است نشان دهیم تأثیر آن روی اعضای پایه‌ی $\mathcal{S}$ همان تأثیر مطلوب است، یعنی کافی است نشان دهیم دو تساوی زیر برقرارند.

$$F(|\mathbf{b}\rangle \otimes |-\rangle) = |\mathbf{b}\rangle \otimes |-\rangle \ , \tag{۳}$$

$$F(|\mathbf{a}\rangle \otimes |-\rangle) = (-\,|\mathbf{a}\rangle) \otimes |-\rangle \ . \tag{۴}$$

از تعریف $|-\rangle$ و $|\mathbf{b}\rangle$ و به دلیل خطی بودن عملگر F داریم

$$\begin{aligned}\sqrt{2\beta}F(|\mathbf{b}\rangle \otimes |-\rangle) &= \sqrt{2}F\left(\sum_{\mathbf{x}\in B} |\mathbf{x}\rangle \otimes |-\rangle\right) \\ &= \sum_{\mathbf{x}\in B}\left[F\left(|\mathbf{x}\rangle \otimes |0\rangle\right) - F\left(|\mathbf{x}\rangle \otimes |1\rangle\right)\right] \, .\end{aligned}$$

طبق تعریف F در (۱) به ازای هر $\mathbf{x} \in B$ و هر $y \in \{0,1\}$ داریم

$$F\left(|\mathbf{x}\rangle \otimes |y\rangle\right) = |\mathbf{x}\rangle \otimes |f(\mathbf{x}) \oplus y\rangle = |\mathbf{x}\rangle \otimes |y\rangle \ .$$

در نتیجه داریم

$$\begin{aligned}\sqrt{2\beta}F(|\mathbf{b}\rangle \otimes |-\rangle) &= \sum_{\mathbf{x}\in B}\left[|\mathbf{x}\rangle \otimes |0\rangle - |\mathbf{x}\rangle \otimes |1\rangle\right] \\ &= \sqrt{2\beta}(|\mathbf{b}\rangle \otimes |-\rangle) \, ,\end{aligned}$$

بنابراین (۳) درست است.

حال (۴) را اثبات می‌کنیم. از تعریف $|-\rangle$ و $|\mathbf{a}\rangle$ و به دلیل خطی بودن عملگر F داریم

$$\begin{aligned}\sqrt{2\alpha}F(|\mathbf{a}\rangle \otimes |-\rangle) &= \sqrt{2}F\left(\sum_{\mathbf{x}\in A} |\mathbf{x}\rangle \otimes |-\rangle\right) \\ &= \sum_{\mathbf{x}\in A}\left[F\left(|\mathbf{x}\rangle \otimes |0\rangle\right) - F\left(|\mathbf{x}\rangle \otimes |1\rangle\right)\right] \, .\end{aligned}$$

طبق تعریف F در (۱) به ازای هر $\mathbf{x} \in A$ و هر $y \in \{0,1\}$ داریم

$$F\left(|\mathbf{x}\rangle \otimes |y\rangle\right) = |\mathbf{x}\rangle \otimes |f(\mathbf{x}) \oplus y\rangle = |\mathbf{x}\rangle \otimes |1-y\rangle \ .$$

در نتیجه داریم

$$\begin{aligned}\sqrt{2\alpha}F(|\mathbf{a}\rangle \otimes |-\rangle) &= \sum_{\mathbf{x}\in A}\left[F(|\mathbf{x}\rangle \otimes |0\rangle) - F(|\mathbf{x}\rangle \otimes |1\rangle)\right] \\ &= \sum_{\mathbf{x}\in A}\left[|\mathbf{x}\rangle \otimes |1\rangle - |\mathbf{x}\rangle \otimes |0\rangle\right] \\ &= \sum_{\mathbf{x}\in A}\left[(-\,|\mathbf{x}\rangle) \otimes |0\rangle - (-\,|\mathbf{x}\rangle) \otimes |1\rangle\right] \\ &= \sqrt{2\alpha}\left(-\,|\mathbf{a}\rangle \otimes |-\rangle\right) \ ,\end{aligned}$$

بنابراین (۴) درست است. □

صفحه‌ی $\mathcal{S}$ را در نظر بگیرید. فرض کنیم θ زاویه‌ی بین خطوط $\mathbf{O}\,|\mathbf{h}\rangle$ و $\mathbf{O}\,|\mathbf{b}\rangle$ باشد (برای این که علامت زوایا را تعیین کنیم، فرض می‌کنیم که خط $\mathbf{O}\,|\mathbf{b}\rangle$ محور x و خط $\mathbf{O}\,|\mathbf{a}\rangle$ محور y باشد.) در این صورت هر بار اعمال عملگر RF شبیه اعمال یک دوران با زاویه‌ی 2θ حول $\mathbf{O}$ در این صفحه می‌باشد. در واقع، فرض کنید $|\mathbf{x}\rangle \in \mathcal{S}$. در این صورت طبق لم ۱ داریم

$$F(|\mathbf{x}\rangle \otimes |-\rangle) = |\mathbf{y}\rangle \otimes |-\rangle \ ,$$

که در آن $|\mathbf{y}\rangle$ حاصل بازتاب $|\mathbf{x}\rangle$ حول $\mathbf{O}\,|\mathbf{b}\rangle$ می‌باشد. به علاوه، طبق تعریف R در (۲) داریم

$$F(|\mathbf{y}\rangle \otimes |-\rangle) = |\mathbf{z}\rangle \otimes |-\rangle \ ,$$

که در آن $|\mathbf{z}\rangle$ حاصل بازتاب $|\mathbf{y}\rangle$ حول $\mathbf{O}\,|\mathbf{h}\rangle$ می‌باشد. ترکیب دو بازتاب با محورهای $\mathbf{O}\,|\mathbf{b}\rangle$ و $\mathbf{O}\,|\mathbf{h}\rangle$ معادل دورانی با زاویه‌ی 2θ حول

$\mathbf{O}$ می‌باشد، بنابراین $|\mathbf{z}\rangle$ حاصل دوران $|\mathbf{x}\rangle$ حول $\mathbf{O}$ و با زاویه‌ی 2θ در صفحه‌ی $\mathcal{S}$ است. توجه کنید که

$$|\mathbf{h}\rangle = \sqrt{\alpha/N}\,|\mathbf{a}\rangle + \sqrt{\beta/N}\,|\mathbf{b}\rangle \ ,$$

بنابراین $|\mathbf{h}\rangle \in \mathcal{S}$ و چون $\alpha > 0$ و $\beta \geq 0$ پس $0 < \theta \leq \pi/2$. نتیجه‌ی زیر حاصل می‌شود.

نتیجه ۲. *داریم $(RF)^k(|\mathbf{h}\rangle \otimes |-\rangle) = |\mathbf{h}'\rangle \otimes |-\rangle$ که در آن $|\mathbf{h}'\rangle$ حاصل دوران $|\mathbf{h}\rangle$ حول $\mathbf{O}$ با زاویه‌ی $2k\theta$ در صفحه‌ی $\mathcal{S}$ است.*

یک نکته‌ی جالب در این‌جا این است که کیوبیت آخر همیشه در حالت $|-\rangle$ می‌ماند و هیچ‌گاه در طول الگوریتم تغییری نمی‌کند، با این حال نمی‌توان این کیوبیت را از الگوریتم حذف کرد!

لم ۳. *فرض کنید $|\mathbf{q}\rangle \in Q_n \cap \mathcal{S}$ و ψ زاویه‌ی بین خطوط $\mathbf{O}\,|\mathbf{q}\rangle$ و $\mathbf{O}\,|\mathbf{b}\rangle$ باشد. فرض کنید پس از اندازه‌گیری $(n+1)$-کیوبیت $|\mathbf{q}\rangle \otimes |-\rangle$، بیت‌های $x_1, x_2, \ldots, x_n, y$ به دست آیند. در این صورت احتمال $f(x_1, \ldots, x_n) = 1$ حداقل $\sin^2(\psi)$ است.*

اثبات. فرض کنید

$$|\mathbf{q}\rangle = \sum_{\mathbf{x} \in \{0,1\}^n} \mathbf{q}_\mathbf{x}\,|\mathbf{x}\rangle$$

نمایش $|\mathbf{q}\rangle$ در پایه‌ی استاندارد باشد. چون $A = f^{-1}(1)$، احتمال مورد نظر برابر است با $\sum_{\mathbf{x}\in A} \mathbf{q}_\mathbf{x}^2$. فرض کنید ϕ زاویه‌ی بین خطوط $\mathbf{O}\,|\mathbf{q}\rangle$ و $\mathbf{O}\,|\mathbf{a}\rangle$ باشد. چون $\mathbf{O}\,|\mathbf{a}\rangle$ بر $\mathbf{O}\,|\mathbf{b}\rangle$ عمود است، داریم $\cos^2(\phi) = \sin^2(\psi)$. طبق نامساوی کوشی-شوارز،

$$\begin{aligned}
\sum_{\mathbf{x}\in A} \mathbf{q}_\mathbf{x}^2 &= \left(\sum_{\mathbf{x}\in A} \mathbf{q}_\mathbf{x}^2\right)\left(\sum_{\mathbf{x}\in A} \frac{1}{\alpha}\right) \\
&\geq \left(\sum_{\mathbf{x}\in A} \mathbf{q}_\mathbf{x}/\sqrt{\alpha}\right)^2 \\
&= \langle |\mathbf{q}\rangle \,, |\mathbf{a}\rangle\rangle^2 = \|\,|\mathbf{q}\rangle\,\|^2 \cdot \|\,|\mathbf{a}\rangle\,\|^2 \cdot \cos^2(\phi) \\
&= \cos^2(\phi) = \sin^2(\psi) \ .
\end{aligned}$$

□

اکنون ابزارهای لازم را برای انتخاب k مناسب و تحلیل الگوریتم گراور داریم. طبق تعریف $|\mathbf{h}\rangle$ و $|\mathbf{b}\rangle$ داریم

$$\cos\theta = \langle |\mathbf{h}\rangle \,, |\mathbf{b}\rangle\rangle = \sum_{\mathbf{x}\in B} 1/\sqrt{N\beta} = \sqrt{\beta/N} \ .$$

در نتیجه داریم $\sin\theta = \sqrt{\alpha/N} \geq 1/\sqrt{N}$. چون تعداد جواب‌های مساله برابر α است و بر الگوریتم معلوم است، θ نیز بر الگوریتم معلوم است. می‌دانیم $\mathbf{O}\,|\mathbf{h}\rangle$ با $\mathbf{O}\,|\mathbf{b}\rangle$ زاویه‌ی θ می‌سازد و k بار اعمال RF معادل دوران به اندازه‌ی $2k\theta$ حول $\mathbf{O}$ می‌باشد. بنابراین طبق نتیجه ۲ پس از k بار اعمال این عملگر به $(n+1)$-کیوبیتی مانند $|\mathbf{h}'\rangle \otimes |-\rangle$ می‌رسیم به طوری که $\mathbf{O}\,|\mathbf{h}'\rangle$ با $\mathbf{O}\,|\mathbf{b}\rangle$ زاویه‌ی $(2k+1)\theta$ را می‌سازد. طبق لم ۳ با اندازه‌گیری این $(n+1)$-کیوبیت به احتمال حداقل $\sin^2((2k+1)\theta)$ جوابی از مساله را پیدا می‌کنیم.

اگر $\pi/4 \leq \theta \leq \pi/2$، کافی است k را برابر صفر انتخاب کنیم، در این صورت احتمال موفقیت الگوریتم گراور حداقل یک‌دوم خواهد بود. اگر $0 < \theta < \pi/4$، فرض کنید k کوچک‌ترین عدد طبیعی باشد که $(2k+1)\theta \geq \pi/4$. در این صورت

$$(2k-1)/\sqrt{N} \leq (2k-1)\sin\theta \leq (2k-1)\theta < \pi/4$$

در نتیجه $k < (\pi\sqrt{N}+4)/8 < \sqrt{N}$. از طرف دیگر داریم $\pi/4 \leq (2k+1)\theta \leq 3\pi/4$؛ لذا با انتخاب این k، احتمال موفقیت الگوریتم گراور حداقل یک‌دوم خواهد بود.

۴ مؤخره

در بخش ۳ مساله‌ی جستجویی تعریف کردیم و الگوریتمی کوانتومی برای حل آن ارائه دادیم. البته این مساله دقیقاً با مسابقه‌ای که در مقدمه مطرح کردیم متناظر نیست؛ به‌هرحال الگوریتم ارائه‌شده بسیار جالب‌توجه است و تفاوت بارزی را بین قدرت الگوریتم‌های کوانتومی و الگوریتم‌های متعارف نشان می‌دهد.

گراور [۲] نشان داد هر الگوریتم کوانتومی که برای مساله‌ی جستجوی بالا ارائه شود که احتمال موفقیتش حداقل یک‌سوم باشد، از $\Omega(\sqrt{N})$ گیت جعبه‌سیاه استفاده می‌کند و بنابراین مرتبه‌ی[۱۰] تعداد پرسش‌های الگوریتم او بهینه است.

نویسندگان مقاله‌ی [۱] نشان دادند که تعداد گیت‌های جعبه‌سیاه الگوریتم گراور بهینه است. هم‌چنین آن‌ها الگوریتمی ارائه دادند که مساله‌ی جستجوی بالا را در حالتی که تعداد جواب‌ها مثبت ولی نامعلوم است حل می‌کند و از $O(\sqrt{N})$ گیت جعبه‌سیاه استفاده می‌کند.

مراجع

[۱] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight bounds on quantum searching. *Fortschritte der Physik*, 46(4-5):493-505,1998.

[۲] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the*

[۱۰] order

Twenty-eighth Annual ACM Symposium on the Theory of Computing (STOC 1996), pages 212-219, New York, 1996, ACM.

بهینه‌سازی ترکیباتی روی گراف‌های با عرض درختی[۱] محدود

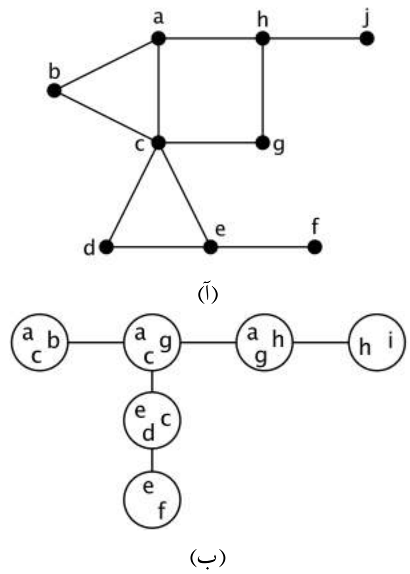
آرش حدادان

۱ مقدمه

بسیاری از مسائل بهینه‌سازی که به روی گراف تعریف می شوند NP-hard هستند. از جمله می‌توان به بزرگترین مجموعه‌ی مستقل، رنگ آمیزی رأسی و غیره اشاره کرد. اما تحت شرایط خاصی می‌توان بسیاری از این مسائل را در زمان چند جمله‌ای حل نمود. برای مثال برای بسیاری از مسائل بهینه‌سازی، در صورتی که گراف مسأله درخت باشد می‌توان الگوریتم چندجمله‌ای پیدا کرد. از جمله‌ی این مسائل می‌توان به کوچک‌ترین مجموعه‌ی مسلط و یا درخت استاینر اشاره نمود [۶]. این مشاهده‌ی کلیدی به تعریف پارامتری برای اندازه‌گیری میزان درخت بودن یک گراف انجامید. *عرض درختی* یک گراف به‌طور شهودی فاصله‌ی آن گراف از درخت بودن را نشان می‌دهد. (درخت‌ها عرض درختی یک دارند.) عرض درختی برای اولین بار توسط رودولف هالین[۲] در ۱۹۶۷ معرفی شد. اما در ۱۹۸۴ رابرتسون[۳] و سیمور[۴] این تعریف را در کنار چند پارامتر عرضی دیگر از جمله عرض مسیری[۵] دوباره معرفی کردند. از آن زمان تاکنون عرض درختی به مفهومی پرکاربرد در نظریه‌ی گراف و طراحی الگوریتم بدل شده است.

در این گزارش با یک الگوریتم عمومی برنامه‌ریزی پویا برای بسیاری مسائل بهینه‌سازی روی گراف‌های با عرض درختی محدود آشنا می‌شویم. این الگوریتم در زمان چندجمله‌ای (غالباً خطی) اجرا شده و به ما برای حل بسیاری مسائل NP-hard روی گراف‌های با عرض درختی محدود کمک می‌کند. در این گزارش ابتدا تجزیه‌ی درختی و سپس تعریفی دقیق از عرض درختی ارائه می‌دهیم. سپس تجزیه‌ی درختی خوب را معرفی می‌کنیم که به ما برای طراحی الگوریتم برنامه‌ریزی پویا کمک می‌کند. علاوه بر این در این بخش چند خاصیت کلیدی تجزیه‌ی درختی و عرض درختی را مرور

[۱] Treewidth.
[۲] Rudolf Halin
[۳] Neil Robertson
[۴] Paul D. Seymour
[۵] Pathwidth



شکل ۱: گرافی با عرض درختی ۲ (آ) و تجزیه‌ی درختی آن (ب)

می‌کنیم. در بخش ۳، الگوریتم چندجمله‌ای برنامه‌ریزی پویا برای مسأله‌ی بزرگترین مجموعه‌ی مستقل وزن‌دار را می‌بینیم. در ادامه‌ی این بخش به‌طور مختصر چارچوب کلی استفاده ازین روش را در حالت کلی خواهیم دید. در نهایت در بخش ۴ با یک نتیجه‌گیری مختصر این گزارش را به پایان می‌رسانیم.

۲ مقدمات

اساسی‌ترین مفهوم در رابطه با عرض درختی، تجزیه‌ی درختی است.

تعریف ۱. *گراف $G = (V, E)$ را در نظر بگیرید. درخت $T = (I, F)$ را تجزیه‌ی درختی G می‌گوییم اگر:*

۱. *به هر نقطه‌ی $i \in I$ یک کیسه*[۶] *$B_i \subseteq V$ نسبت دهیم.*

۲. *برای هر $v \in V$ حداقل یک نقطه $i \in I$ وجود داشته باشد به‌طوری‌که $v \in B_i$.*

۳. *برای هر یال $e = uv$ در G نقطه‌ای مانند $i \in I$ وجود داشته باشد که $\{u, v\} \subset B_i$.*

۴. *برای هر رأس $v \in V$ مجموعه $Y = \{i \in I |\ v \in B_i\}$ یک زیردرخت روی T القا کند.*

همچنین عرض تجزیه‌ی درختی T عبارت است از

$$\max_{i \in I} |B_i| - ۱$$

[۶] Bag

عرض درختی گراف G که با $tw(G)$ نشان می‌دهیم عبارت است از عرض کم عرض‌ترین تجزیه‌ی درختی G. دقت کنید که ۱− در تعریف عرض درختی تنها برای این است که برای درختها عرض درختی برابر با یک باشد. در شکل ۱ یک گراف و تجزیه‌ی درختی آن نمایش داده شده است. توجه کنید که عرض تجزیه‌ی درختی نشان داده شده ۲ است. از طرفی چون گراف داده شده دور دارد، ویژگی چهارم تجزیه‌ی درختی ما را مجبور می‌کند که در تجزیه‌ی درختی گراف داده شده، یک کیسه با حداقل سه رأس داشته باشیم. پس عرض درختی گراف داده شده در این شکل ۲ است. از این‌جا به بعد برای جلوگیری از شبهه، از کلمه **رأس** برای رئوس گراف اصلی و از کلمه **نقطه** برای اشاره به رئوس تجزیه‌ی درختی استفاده می‌کنیم. قضیه‌ی زیر یکی از اصلی‌ترین نتایج در مورد محاسبه‌ی عرض درختی گراف‌ها در حالت کلی است [۱] .

قضیه ۲. *برای گراف G و عدد صحیح k، تشخیص این‌که عرض درختی G حداکثر k است یا خیر، NP-complete است.*

ولی قضیه‌ی زیر نشان می‌دهد که اگر عرض درختی گراف محدود باشد می‌توان تجزیه‌ی درختی آن را در زمان چندجمله‌ای محاسبه نمود [۲] .

قضیه ۳. *برای هر $k \in \mathbb{N}$ یک الگوریتم وجود دارد که در زمان خطی، بررسی می‌کند که عرض درختی یک گراف حداکثر k است یا خیر و اگر بود یک تجزیه‌ی درختی با عرض حداکثر k به ما می‌دهد.*

پیش از رسیدن به تکنیک برنامه‌ریزی پویا که در بخش بعدی به آن خواهیم پرداخت؛ لازم است تعریف دیگری که ما را در طراحی الگوریتم برنامه‌ریزی پویا کمک می‌کند، مطرح کنیم.

تعریف ۴. *یک تجزیه‌ی درختی T را تجزیه‌ی درختی خوب می‌گوییم در صورتی که اگر یکی از نقطه‌های T مانند r را به عنوان ریشه درنظر بگیریم، نقطه‌های درخت ریشه‌دار T_r یکی از چهار شکل زیر باشند:*

- **برگ**

  یک نقطه i در T_r برگ است اگر فرزندی نداشته باشد و $|B_i| = 1$.

- **الحاقی**

  یک نقطه i در T_r الحاقی است اگر دقیقاً دو فرزند مانند j_1 و j_2 داشته باشد و $B_i = B_{j_1} = B_{j_2}$.

- **فراموشی**

  یک نقطه i در T_r فراموشی است اگر دقیقاً یک فرزند مانند j داشته باشد و رأسی مانند $v \in V$ داشته باشیم به‌طوری‌که $B_j = B_i \cup \{v\}$ و $|B_i| = |B_j| - 1$.

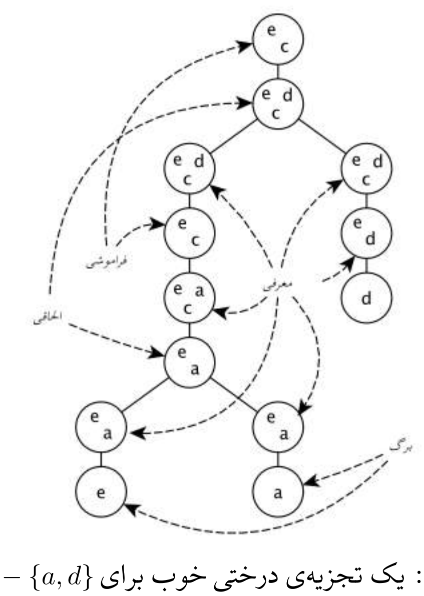
- **معرفی**

  یک نقطه i در T_r معرفی است اگر دقیقاً یک فرزند مانند j داشته باشد و رأسی مانند $v \in V$ داشته باشیم به‌طوری‌که $B_i = B_j \cup \{v\}$ و $|B_j| = |B_i| - 1$.

دقت کنید که هر نقطه در تجزیه‌ی درختی خوب باید دقیقاً یکی ازین چهار برچسب را داشته باشد. در شکل ۲ تجزیه‌ی درختی خوب گراف K_4 بدون یال $\{a, d\}$ را می‌بینید. توجه کنید که وجه تسمیه نقاط فراموشی و معرفی در صورتی واضح خواهد بود که به درخت از پایین به بالا نگاه کنیم. قضیه‌ی زیر نشان می‌دهد که پیدا کردن تجزیه‌ی درختی خوب کار مشکلی نیست.

قضیه ۵. *یک تجزیه‌ی درختی با عرض k از یک گراف با n رأس را می‌توان در زمان $O(k^2 n)$ به یک تجزیه‌ی درختی خوب با عرض k و حداکثر $O(kn)$ نقطه تبدیل کرد.*

اثبات این قضیه را به خواننده واگذار می‌کنیم (راهنمایی: با استقرا روی تعداد نقاط تجزیه‌ی درختی داده شده نشان دهید می‌توان هر نقطه تجزیه‌ی درختی را به عنوان ریشه در نظر گرفت و تجزیه‌ی درختی خوب را بر اساس آن ساخت.).



شکل ۲: یک تجزیه‌ی درختی خوب برای $K_4 - \{a, d\}$.

۳ الگوریتم

در مسأله‌ی بزرگترین مجموعه‌ی مستقل وزن‌دار برای یک گراف داده شده $G = (V, E)$ با وزن رأسی نامنفی $c(v)$، به دنبال پروزن‌ترین مجموعه‌ی رئوس دو به دو غیرمجاور از رئوس G هستیم. این مسأله در حالت کلی NP-hard است. در این بخش برای حالتی که عرض درختی G برابر k باشد، یک الگوریتم با زمان $O(۲^k n)$ ارائه می‌دهیم.

فرض کنید که گراف $G = (V, E)$ به همراه تجزیه‌ی درختی آن $T = (I, F)$ داده شده است. می‌توان فرض کرد که T یک تجزیه‌ی درختی خوب است. در غیر این صورت با کمک قضیه‌ی ۵ می‌توان از روی آن یک تجزیه‌ی درختی خوب ساخت. به هر نقطه $i \in I$ یک گراف G_iنسبت می‌دهیم که رئوس آن مجموعه‌ی $V_i = \bigcup_{j\in J} B_j$ هستند که J مجموعه‌ی تمام نوادگان i در T و خود i است. گراف G_i زیرگراف القا شده به وسیله‌ی V_i بروی G است. حال برای هر نقطه‌ی $i \in I$ می‌خواهیم یک جدول C_i را محاسبه کنیم. جدول C_i برای هر زیرمجموعه‌ی B_i یک خانه دارد. در نتیجه جدول هر نقطه حداکثر $۲^{k+۱}$ خانه دارد (چون عرض درختی k است). برای $S \subseteq B_i$ مقدار $C_i(S)$ عبارت‌است از وزن پروزن‌ترین مجموعه‌ی مستقل $W \subset V_i$ به طوری که $B_i \cap W = S$. به عبارت دیگر $C_i(S)$ وزن پروزن‌ترین مجموعه‌ی مستقل G_i است که همه‌ی رئوس S را دارد و هیچ‌یک از رئوس $B_i \setminus S$ را ندارد. اگر چنین مجموعه‌ی مستقلی وجود نداشته باشد (در زیر گراف القا شده توسط S به روی G_i دو رأس مجاور باشند) آن‌گاه قرار می‌دهیم $C_i(S) = -\infty$.

الگوریتم به این صورت عمل می‌کند که جدول را برای هر نقطه با استفاده از جدول فرزندهایش محاسبه می‌کند. به همین منظور بایستی جدول‌ها را از پایین به بالا محاسبه کرد. چون در یک تجزیه‌ی درختی خوب چهار نوع نقطه داریم تنها کافیست نحوه‌ی محاسبه‌ی جدول را برای هر نوع نقطه شرح دهیم.

۱.۳ نقاط برگ

برای یک نقطه‌ی برگ مانند i داریم $B_i = \{v\}$. در نتیجه برای کامل کردن جدول نقطه‌ی i کافیست که $C_i(\{v\})$ و $C_i(\emptyset)$ را محاسبه کنیم. به سادگی می‌توان دید که $C_i(\{v\}) = c(v)$ و $C_i(\emptyset) = ۰$. بدین ترتیب برای محاسبه‌ی جدول یک نقطه‌ی برگ به زمانی ثابت نیاز داریم.

۲.۳ نقاط معرفی

اگر i یک نقطه‌ی معرفی T باشد، آن‌گاه i دقیقاً یک فرزند مانند j دارد و رأسی مانند $v \in V$ وجود دارد که $B_i = B_j \cup \{v\}$. دقت کنید که رأس v مجاور هیچ یک از رئوس $V_j \setminus B_j$ نیست. زیرا در غیراین‌صورت از ویژگی چهارم تجزیه‌ی درختی می‌توان نتیجه گرفت که $v \in B_j$ است که تناقض است.

لم ۶. *اگر $S \subseteq B_j$ باشد،*

۱. $C_i(S) = C_j(S)$ ،

۲. *اگر یک رأس مانند $u \in S$ باشد که $\{u, v\} \in E$ آن‌گاه* $C_i(S \cup \{v\}) = -\infty$،

۳. *اگر به ازای هر $u \in S$ داشته باشیم $\{u, v\} \notin E$ آن‌گاه* $C_i(S \cup \{v\}) = C_j(S) + c(v)$.

اثبات. ۱. برای هر $S \subseteq B_j$ می‌دانیم که $S \subseteq B_i$. همچنین هر مجموعه‌ی مستقل W روی G_j به‌طوری‌که $W \cap B_j = S$ یک مجموعه‌ی مستقل روی G_i هم هست و $W \cap B_i = W \cap (B_j \cup \{v\}) = S$ چرا که $\{v\} \notin W$.

۲. اگر مجموعه‌ی مستقلی مانند W وجود داشته باشد، از آن‌جا $S \cup \{v\} \subseteq W$ یعنی W دو رأس مجاور دارد که تناقض است. پس $C_i(S \cup \{v\}) = -\infty$.

۳. از آن‌جا که v هیچ رأس مجاوری در $V_j \setminus B_j$ ندارد هر مجموعه‌ی مستقل مانند W روی G_j به‌طوری‌که $W \cap B_i = S$ و S رأسی مجاور v نداشته باشد را می‌توان با اضافه کردن v به مجموعه‌ی مستقلی برای G_i گسترش داد. دقت کنید که این مجموعه‌ی مستقل پروزن‌ترین مجموعه‌ی مستقل ممکن است. در غیر این صورت با بیشینه بودن W روی G_j به تناقض می‌رسیم. پس داریم $C_i(S \cup \{v\}) = C_j(S) + c(v)$.

□

برای محاسبه‌ی هر خانه از جدول یک نقطه‌ی معرفی مانند $C_i(S)$ بایستی مجاورت حداکثر $k + ۱$ رأسی که در S هستند را با v بررسی کنیم. خواننده به عنوان تمرینی ساده می‌تواند نشان دهد که از این بررسی مجاورت برای هر خانه به‌طور مجزا می‌توان پرهیز کرد و هر خانه‌ی جدول یک نقطه‌ی معرفی را در زمان ثابت محاسبه کرد. با این حساب می‌توان نشان داد که برای کامل کردن جدول یک نقطه‌ی معرفی، زمانی از مرتبه‌ی $O(۲^{k+۱})$ نیاز است.

۳.۳ نقاط فراموشی

اگر i یک نقطه‌ی فراموشی T باشد آن‌گاه i دقیقاً یک فرزند مانند j دارد و نقطه‌ای مانند $v \in V$ وجود دارد که $B_j = B_i \cup \{v\}$. به سادگی می‌توان دید که $G_i = G_j$.

لم ۷. *اگر $S \subseteq B_i$ باشد داریم:*

$$C_i(S) = \max\{C_j(S), C_j(S\cup\{v\})\}$$

اثبات. یک مجموعه‌ی مستقل بیشینه‌ی وزن مانند W روی G_i که $W\cap B_i = S$ را در نظر بگیرید. این مجموعه یا v را شامل می‌شود یا نه. اگر $\{v\} \in W$ آن‌گاه متناظر با مجموعه‌ی مستقل بیشینه‌ی وزن مانند W_1 روی G_j است که $W_1 \cap B_j = S\cup\{v\}$ که در این صورت وزن آن $C_j(S\cup\{v\})$ خواهد بود. اگر $\{v\}\notin W$ آن‌گاه متناظر با مجموعه‌ی مستقل بیشینه‌ی وزن W_2 روی G_j خواهد بود که $W_2\cap B_j = S$ و در این صورت وزن آن $C_j(S)$ خواهد بود. □

دقت کنید که می‌توان هریک از 2^{k+1} خانه‌ی جدول یک نقطه‌ی فراموشی را در $O(1)$ محاسبه نمود. بنابراین می‌توان C_i را وقتی i یک نقطه‌ی فراموشی است در $O(2^{k+1})$ محاسبه کرد.

۴.۳ نقاط الحاقی

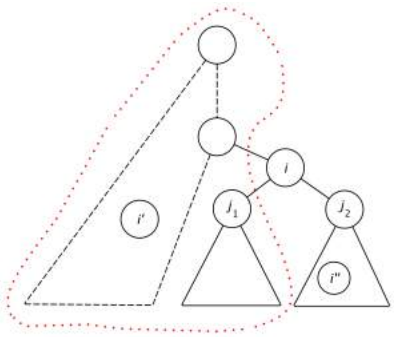
اگر i یک نقطه‌ی الحاقی T باشد آن‌گاه i دقیقاً دو فرزند مانند j_1 و j_2 دارد و $B_i = B_{j_1} = B_{j_2}$. به راحتی می‌توان دید که $V_i = V_{j_1}\cup V_{j_2}$ و در واقع $G_i = G_{j_1}\cup G_{j_2}$.

لم ۸. *اگر $u\in V_{j_1}$، $v\in V_{j_2}$ و $u,v\notin B_i$ داریم $\{u,v\}\notin E$.*

اثبات. برای رسیدن به تناقض فرض کنید $\{u,v\}\in E$. براساس ویژگی سوم تجزیه‌ی درختی نقطه‌ای مانند i' در T هست که $u,v\in B_{i'}$. بدون کاستن از کلیت فرض کنید که i' در زیردرخت با ریشه‌ی j_2 نیست (حالتی دیگر که i' در زیردرخت با ریشه‌ی j_1 نیست مشابه ادامه این اثبات خواهد بود.). از آن‌جا که $v\in V_{j_2}$، نقطه‌ای مانند i'' در زیردرخت با ریشه‌ی j_2 هست که $v\in B_{i''}$. با استناد به چهارمین ویژگی تجزیه‌ی درختی، چون v در کیسه‌ی هردو نقطه‌ی i' و i'' حضور دارد، v بایستی در کیسه تمامی نقاطی که در مسیر i' به i'' هستند نیز باشد. یکی از این نقاط i است. یعنی $v \in B_i$ که این یک تناقض است (شکل ۳). □

لم ۹. *اگر $S\subseteq B_i$ آن‌گاه*

$$C_i(S) = C_{j_1}(S)+C_{j_2}(S)-c(S),$$



شکل ۳: طرح کلی اثبات لم ۸

که $c(S)=\sum_{v\in S} c(v)$.

اثبات. ابتدا نشان می‌دهیم که $C_i(S)\le C_{j_1}(S)+C_{j_2}(S)-c(S)$. فرض کنید W مجموعه‌ی مستقل بیشینه‌ی وزن روی G_i باشد که $W\cap B_i = S$. چون $B_{j_1}=B_i$ داریم:

$$(W\cap V_{j_1})\cap B_{j_1} = S\cap V_{j_1} = S$$

همچنین $W\cap V_{j_1}$ یک مجموعه‌ی مستقل برای G_{j_1} نیز هست. مشابه این را برای $W\cap V_{j_2}$ روی G_{j_2} می‌توان گفت. در نتیجه:

$$\begin{aligned} C_i(S) &= c(W)\\ &= c(W\cap V_{j_1}) + c(W\cap V_{j_2}) - c(W\cap V_{j_1}\cap V_{j_2})\\ &\le C_{j_1}(S)+C_{j_2}(S)-c(S). \end{aligned}$$

حالا با نشان دادن این‌که $C_i(S)\ge C_{j_1}(S)+C_{j_2}(S)-c(S)$ اثبات را کامل می‌کنیم. فرض کنید W_1 مجموعه‌ی مستقل بیشینه‌ی وزن روی G_{j_1} باشد به‌طوری‌که $W_1\cap B_{j_1}=S$. مشابهاً W_2 را روی G_{j_2} در نظر بگیرید. ادعا می‌کنیم که $W_1\cup W_2$ یک مجموعه‌ی مستقل روی G_i است. در این صورت خواهیم داشت:

$$\begin{aligned} (W_1\cup W_2)\cap B_i &= (W_1\cap B_i)\cup(W_2\cap B_i)\\ &= (W_1\cap B_{j_1})\cup(W_2\cap B_{j_2})\\ &= S, \end{aligned}$$

و در نتیجه خواهیم داشت:

$$\begin{aligned} C_i(S) &\ge c(W_1\cup W_2)\\ &= c(W_1)+c(W_2)-c(W_1\cap W_2)\\ &= C_{j_1}(S)+C_{j_2}(S)-c(S). \end{aligned}$$

حال کافیست نشان دهیم که $W_1\cup W_2$ یک مجموعه‌ی مستقل روی G_i است. برای هر دو رأس $u,v\in W_1\cup W_2$ اگر هردو در W_1 یا W_2

باشند آن‌گاه $\{u,v\} \notin E$ زیرا W_1 و W_2 خود مجموعه‌های مستقل هستند. اگر $u \in W_1$ و $v \in W_2$ و $u,v \notin B_i$ یعنی $u \in V_{j_1}$ و $v \in V_{j_2}$ در نتیجه با استناد به لم ۸ داریم $\{u,v\} \notin E$. در نهایت اگر حداقل یکی از u یا v در B_i باشند، مثلاً u، آن‌گاه $u \in B_{j_1}$ در نتیجه $u \in V_{j_1}$. همچنین W_2 یک مجموعه‌ی مستقل روی G_{j_2} است. بنابراین $\{u,v\} \notin E$. □

توجه کنید که برای محاسبه هر $C_i(S)$ای بایستی $c(S)$ را محاسبه کنید و این کار به $O(k)$ زمان نیاز دارد. اما می‌توان با ذخیره این مقدار به راحتی برای محاسبه وزن زیرمجموعه‌های بزرگتر استفاده کرد. نشان دهید که با ذخیره این اطلاعات می‌توان $c(S)$ را در $O(1)$ محاسبه نمود. با این فرض حالا می‌توان دید که جدول نقاط الحاقی را می‌توان در زمان $O(2^{k+1})$ تکمیل کرد.

۵.۳ جمع‌بندی

برای مرتب کردن نقاط برای محاسبه‌ی جداول کافیست یک پس‌ترتیب[7] روی رئوس T انجام دهیم تا هنگامی که جدول یک نقطه را محاسبه می‌کنیم جدول فرزندان آن نقطه را پیشتر محاسبه کرده باشیم. پس‌ترتیب را می توان در زمان خطی روی تعداد نقاط تجزیه‌ی درختی خوب انجام داد. ریشه‌ی تجزیه‌ی درختی خوب، r، آخرین نقطه‌ای است که جدولش را محاسبه می‌کنیم. دقت کنید که $G = G_r$. بنابراین وزن مجموعه‌ی مستقل با بیشینه‌ی وزن روی G عبارت است از $\max_{S \in B_r} C_r(S)$. زمان اجرای این الگوریتم $O(2^k n)$ است زیرا برای محاسبه‌ی جدول هرنقطه به $O(2^k)$ زمان نیاز داریم و در مجموع $O(n)$ نقطه در تجزیه‌ی درختی داریم. به عنوان یک تمرین ساده نشان دهید می‌توان این الگوریتم را طوری تغییر داد که مجموعه‌ی مستقل با بیشینه‌ی وزن را بسازد.

این روش که در این بخش برای مسأله بزرگترین مجموعه‌ی مستقل وزن‌دار تشریح شد یک روش کلی است و می‌توان آن را برای پیدا کردن الگوریتم‌های چندجمله‌ای روی بسیاری از مسأله‌های بهینه‌سازی روی گراف‌های با عرض درختی محدود استفاده کرد. از جمله این مسائل می‌توان به مسأله‌ی پوشش رأسی، درخت استاینر، رنگ‌آمیزی گراف و غیره اشاره نمود. برای یک گراف داده شده مثل G با عرض درختی کم، این گام‌های اصلی در طراحی الگوریتمی مشابه به‌کار می‌روند:

۱. پیدا کردن تجزیه‌ی درختی گراف. زمان اجرا: $O(n)$.

۲. تبدیل کردن تجزیه‌ی درختی را به یک تجزیه‌ی درختی خوب. زمان اجرا: $O(n)$.

[7] Postorder

۳. پیدا کردن پس‌ترتیب روی نقاط تجزیه‌ی درختی خوب. زمان اجرا: $O(n)$.

۴. محاسبه کردن جدول برای هر نقطه با استفاده از ترتیبی که در مرحله‌ی قبل پیدا شد. زمان اجرا: $O(n)$.

۵. پیدا کردن خانه‌ای در جدول متناظر با ریشه‌ی تجزیه‌ی درختی خوب که بزرگترین (کوچکترین) مقدار را دارد. زمان اجرا $O(1)$.

بنابراین الگوریتم در زمان $O(n)$ اجرا می‌شود. [8]

۴ نتیجه‌گیری

همانطور که در این گزارش دیدیم بسیار از مسائل NP-hard روی درخت‌های با عرض درختی کم قابل ردیابی هستند. یک طرح نوین برای حل مسائل بهینه‌سازی این است که گرافی با عرض درختی محدود به عنوان تخمینی برای یک گراف داده شده با عرض درختی نامحدود پیدا کنیم و مسأله بهینه‌سازی داده شده را روی گرافی که تخمین زده‌ایم حل کنیم.

مراجع

[۱] Arnborg, Stefan, Corneil, Derek G., and Proskurowski, Andrzej. Complexity of finding embeddings in a k-tree. *SIAM J. Algebraic Discrete Methods*, 8(2):277–284, April 1987.

[۲] Bodlaender, Hans L. A linear time algorithm for finding tree-decompositions of small treewidth. In *Proceedings of the Twenty-fifth Annual ACM Symposium on Theory of Computing*, STOC '93, pp. 226–234, New York, NY, USA, 1993. ACM.

[۳] Bodlaender, Hans L., Drange, Pål Grønås, Dregi, Markus S., Fomin, Fedor V., Lokshtanov, Daniel, and Pilipczuk, Michal. A $o(c^k n)$ 5-approximation algorithm for treewidth. *CoRR*, abs/1304.6321, 2013.

[8] برای بعضی مسائل ممکن است زمان اجرای گام ۴ به پارامترهای دیگر مثل تعداد رنگ بستگی داشته باشد و متفاوت باشد.

[۴] Bodlaender, Hans L. and Koster, Arie M. C. A. Combinatorial optimization on graphs of bounded treewidth. *Comput. J.*, 51(3):255–269, May 2008.

[۵] Kloks, T. Treewidth. computations and approximations. 1994.

[۶] Markus Chimani and Petra Mutzel and Bernd Zey, Improved Steiner tree algorithms for bounded treewidt. *J. Discrete Algorithms*, 2012.

مسأله‌ها

مسأله ۱. فرض کنید G یک گروه و H_iها، $1 \le i \le n$ زیرگروه‌هایی از آن باشند. در این صورت نشان دهید اگر $G = \cup_{i=1}^{n} H_i$ و اندیس H_1 در G نامتناهی باشد؛ آن‌گاه می‌توان از این تساوی H_1 را حذف نمود: $G = \cup_{i=2}^{n} H_i$.

مسأله ۲. فرض کنید F یک میدان با مشخصه‌ی صفر و $x_1, \dots, x_{2n+1}$ اعضایی ناصفر از F باشند. ثابت کنید که می‌توان $n+1$ تا از آنها را انتخاب کرد که هیچ زیرمجموعه‌ای از این $n+1$ عضو، جمعی برابر صفر نداشته باشد.[1]

مسأله ۳. I ایده‌آلی از حلقه‌ی $\mathbb{Z}[x]$ است با این خواص که

- عناصر I هیچ مقسوم‌علیه مشترکی با درجه‌ی مثبت ندارند.
- I شامل چندجمله‌ای با جمله‌ی ثابت ۱ است.

ثابت کنید I عضوی به شکل $1 + x + x^2 + \cdots + x^{r-1}$ دارد که در آن $r \in \mathbb{N}$.

مسأله ۴. فرض کنید R یک حلقه‌ی جابجایی و متناهی باشد. ثابت کنید R یکدار است اگر و تنها اگر ۰ تنها پوچساز حلقه باشد. ($a \in R$ را پوچساز R می‌نامند هرگاه $aR = \{0\}$.)

مسأله ۵. $\alpha \in (0,1]$ یک عدد حقیقی است و $\{a_n\}_{n=1}^{\infty}$ دنباله‌ای است از اعداد حقیقی که در

$$a_{n+1} \le \alpha a_n + (1-\alpha) a_{n-1} \quad \forall n \ge 2$$

صدق می‌کند. اثبات کنید این دنباله در صورتی که از پایین کراندار باشد باید همگرا شود.

مسأله ۶. فرض کنید $g : [0,1] \to \mathbb{R}$ یک تابع مشتق‌پذیر با مشتق پیوسته باشد. ثابت کنید برای هر $x \in [0,1]$:

$$\int_0^1 \big(g(t) - |\, g'(t)\,|\big)\mathrm{d}t \le g(x) \le \int_0^1 \big(g(t) + |\, g'(t)\,|\big)\mathrm{d}t$$

مسأله ۷. $f_1, f_2, \dots, f_n : \mathbb{R}^n \to \mathbb{R}$ توابعی C^2 هستند. فرض کنید ثابت‌های c_{ij} چنان باشند که برای هر $1 \le i, j \le n$ داشته باشیم:

$$\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i} = c_{ij}$$

نشان دهید $g : \mathbb{R}^n \to \mathbb{R}$ موجود است به قسمی که برای هر $1 \le i \le n$، $f_i + \frac{\partial g}{\partial x_i}$ خطی است. (لازم به ذکر است که اینجا مختصات بر $\mathbb{R}^n$ را با $x = (x_1, \dots, x_n)$ نمایش داده‌ایم و منظور از تابع خطی بر $\mathbb{R}^n$ تابعی با ضابطه‌ی

$$a_0 + a_1 x_1 + \cdots + a_n x_n$$

است.)

مسأله ۸. فرض کنید $P_1, P_2, \dots$ دنباله‌ای چگال از نقاط متمایز در بازه‌ی $(0,1)$ باشد. نقاط $P_1, P_2, \dots, P_{n-1}$ این بازه را به $n-1$ بخش افراز می‌کنند و با اضافه شدن P_n یکی از این بخش‌ها به دو بازه‌ی کوچکتر افراز می‌شود که طول آنها را a_n و b_n می‌گیریم. نشان دهید:

$$\sum_{n=1}^{\infty} a_n b_n (a_n + b_n) = \frac{1}{3}$$

مسأله ۹. فرض کنید A و B دو ماتریس 2×2 متقارن نامنفی باشند. نشان دهید برای هر $m \in \mathbb{N}$، $g(t) = \mathrm{tr}(A + Bt)^m$ یک چندجمله‌ای با ضرایب نامنفی است (نامنفی بودن A و B یعنی $X^T A X \ge 0$ و $X^T B X \ge 0$ برای هر $X \in \mathbb{R}^2$ و tr نماد تریس است).

مسأله ۱۰. A و B دو ماتریس مربعی با درایه‌های در یک میدان و از اندازه‌ی یکسان هستند به قسمی که $ABAB = 0$. آیا این لزوماً نتیجه می‌دهد که $BABA = 0$؟

مسأله‌های پیشنهادی خود را به همراه راه‌حل برای ما بفرستید:

mathematicsjournal@gmail.com

[1] مطرح شده در شماره‌ی ۷ مجله‌ی ریاضی شریف، پاییز ۷۶

پاسخ مسأله‌ها

پاسخ ۱. نیاز به اثبات چند لم داریم:

لم ۱. فرض کنید G یک گروه باشد و $n,m\in\mathbb{N}$ و H_i $(i=1,\dots,n)$ و D زیرگروه‌هایی از G باشند. اگر $x_i,y_j\in G$ وجود داشته باشند به طوری که:

۱) $G=H_1x_1\cup\cdots\cup H_nx_n\cup Dy_1\cup\cdots\cup Dy_m$

۲) $G\neq Dy_1\cup\cdots\cup Dy_m$

آن‌گاه G اجتماع تعداد متناهی از هم‌دسته‌های زیرگروه‌های $H_1,\dots,H_n$ است.

اثبات) از شرط دوم نتیجه می‌شود یک عنصرِ $g\in G\backslash Dy_1\cup\cdots\cup Dy_m$ موجود است و در نتیجه:

$$Dg\cap(Dy_1\cup\cdots\cup Dy_m)=\emptyset$$

حال از شرط اول داریم:

$$Dg\subseteq H_1x_1\cup\cdots\cup H_nx_n$$
$$\Rightarrow Dy_i\subseteq(H_1x_1\cup\cdots\cup H_nx_n)g^{-1}y_i=\cup_{1\leq j\leq n}H_j(x_jg^{-1}y_i)$$
$$\Rightarrow Dy_1\cup\cdots\cup Dy_m\subseteq\cup_{1\leq i\leq m,1\leq j\leq n}H_j(x_jg^{-1}y_i)$$

و حکم نتیجه می‌شود.

لم ۲. اگر G توسط تعداد متناهی از هم‌دسته‌های راست زیرگروه‌هایش پوشش داده شود، آن‌گاه حداقل یکی از آن زیرگروه‌ها از اندیس متناهی در G است.

اثبات) فرض کنید $H_i<G$، $x_i\in G$ و $G=\cup_{i=1}^n H_ix_i$. می‌خواهیم ثابت کنیم وجود دارد $1\leq i\leq n$ به طوری که $[G:H_i]<\infty$. حکم را با استقرا ثابت می‌کنیم. برای $n=1$ که حکم بدیهی است.

حال فرض کنید برای n زیرگروه حکم برقرار باشد. حال $H_1,\dots,H_n$ را در نظر بگیرید و قرار دهید $D=H_{n+1}$ در این صورت شرط (۱) در لم قبلی برقرار است. اگر $[G:D]$ متناهی باشد که حکم نتیجه شده است در غیر این صورت $[G:D]$ نامتناهی است و بنابراین شرط (۲) در لم قبل نیز برقرار خواهد شد و در نتیجه می‌توان G را به صورت اجتماعی متناهی از هم‌دسته‌های $H_1,\dots,H_n$ نوشت که بنابر فرض استقرا نتیجه می‌شود یکی از آن‌ها از اندیس متناهی در G است.

و در نهایت آخرین لم:

لم ۳. فرض کنید $H_1,\dots,H_n$ زیرگروه‌هایی از G باشند و x_iها $1\leq i\leq n$ اعضای G به طوری که $G=H_1x_1\cup\cdots\cup H_nx_n$. اگر H_n از اندیس نامتناهی در G باشد آن‌گاه می‌توان نوشت $G=H_1x_1\cup\cdots\cup H_{n-1}x_{n-1}$.

اثبات. بنابر لم قبلی حداقل یکی از H_iها از اندیس متناهی در G است. پس با اندیس‌گذاریِ مجدد در صورت لزوم فرض کنید برای $i=1,\dots,k$، $[G:H_i]<\infty$ و برای $i=k+1,\dots,n$، $[G:H_i]=\infty$. قرار می‌دهیم:

$$D=H_1\cap\cdots\cap H_k$$

از آنجا که اشتراک متناهی زیرگروه از اندیس متناهی همواره از اندیس متناهی است[1]، داریم: $[G:D]<\infty$ و هر H_i، $1\leq i\leq k$ اجتماع تعدادی متناهی از هم‌دسته‌های D خواهد بود. پس:

$$\exists y_1,\dots,y_m\in G: H_1x_1\cup\cdots\cup H_kx_k=Dy_1\cup\cdots\cup Dy_m$$

حال اگر $G\neq Dy_1\cup\cdots\cup Dy_m$ آن‌گاه بنابر لم اول G به صورت اجتماع متناهی‌تا از هم‌دسته‌های راست $H_{k+1},\dots,H_n$ خواهد بود که بنابر لم قبلی این امکان‌پذیر نیست (داشتیم: $k+1\leq i\leq n,[G:H_i]=\infty$). لذا:

$$G=Dy_1\cup\cdots\cup Dy_m=H_1x_1\cup\cdots\cup H_kx_k$$
$$\subseteq H_1x_1\cup\cdots\cup H_{n-1}x_{n-1}$$

که اثباتِ این لم را تکمیل می‌کند.

اکنون در لم آخر قرار می‌دهیم $x_i=e$ برای هر i و حکمِ لم تبدیل می‌شود به اینکه هرگاه $G=\cup_{i=1}^n H_i$ و $H_i<G$، آن‌گاه برقراریِ $[G:H_1]=\infty$ مستلزم آن است $G=\cup_{i=2}^n H_i$. که همان حکم مسأله است.

پاسخ ۲. از آنجا که $char(F)=0$، F نسخه‌ای از $\mathbb{Q}$ را دربردارد که همان زیرمیدان اولِ[2] F است. پس می‌توان F را یک فضای برداری بر $\mathbb{Q}$ تلقی کرد و مسأله را برای زیرفضای برداری‌ای بر $\mathbb{Q}$ که این عناصر می‌سازند در نظر گرفت. اگر بُعدِ این زیرفضا بر $\mathbb{Q}$ را m بنامیم، آنگاه مسأله تبدیل می‌شود به: «$2n+1$ بردار ناصفرِ $x_1,\dots,x_{2n+1}$ در $\mathbb{Q}^m$ داده شده‌اند. یک زیرمجموعه‌ی $n+1$ عضوی از آنها موجود است که جمع هیچ زیرمجموعه‌ای از بردارهای متعلق به آن صفر نمی‌شود.» به منظور اثبات، $\mathbb{Q}^m$ را به ضرب داخلیِ معمولی مجهز کنید که از $\mathbb{R}^m$ به ارث می‌برد. با توجه به نامتناهی بودنِ میدان زمینه، یک ابرصفحه‌ی گذرا از مبدأ موجود است که شامل هیچ‌یک از این $2n+1$ بردارِ ناصفر نباشد. چنین ابرصفحه‌ای باید به ازای یک $v\in\mathbb{Q}^m-\{0\}$ با $\{x\in\mathbb{Q}^m\mid x.v=0\}$ داده شود. بنابراین $v.x_i\neq 0$ برای هر $1\leq i\leq 2n+1$. بنابر اصل لانه‌ی

[1] از این حکم استفاده کردیم که برای هر دو زیرگروهِ H و K از یک گروه دلخواهِ G داریم:
$$[G:H\cap K]\leq[G:H][G:K]$$

[2] prime subfield

کبوتری باید حداقل $n+1$ تا از این ضرب‌های داخلی هم‌علامت باشند. پس $1 \le i_1 < \cdots < i_{n+1} \le 2n+1$ موجودند به قسمی که یا همه‌ی $v.x_{i_j}$ها مثبتند یا همه‌ی آنها منفی. این به وضوح نشان می‌دهد که جمع هیچ تعدادی از بردارهای $x_{i_1},\ldots,x_{i_{n+1}}$ صفر نیست زیرا همواره ضرب داخلیِ چنین حاصل‌جمعی در v به ترتیب مثبت یا منفی خواهد بود.

پاسخ ۳. به منظور حل مسأله، ابتدا باید درستیِ چند حکم ساده‌تر را اثبات کنیم:

(الف) ایده‌آلِ I حتماً یک عددِ طبیعی را شامل می‌شود.

اثباتِ (الف): برای دیدنِ دلیل، ایده‌آلی که عناصرِ I در $\mathbb{Q}[x]$ -که برخلافِ $\mathbb{Z}[x]$ **PID** است- تولید می‌کنند را در نظر بگیرید. این ایده‌آل نمی‌تواند سره باشد. زیرا در غیر این‌صورت باید یک چندجمله‌ایِ $p(x) \in \mathbb{Q}[x]$ با درجه‌ی مثبت موجود باشد به قسمی که در $\mathbb{Q}[x]$ مقسوم‌علیه‌ای از هر $f(x) \in I$ باشد. با ضرب در عنصری از $\mathbb{Q}^{\times}$ بدون آنکه ایده‌آلی از $\mathbb{Q}[x]$ که $p(x)$ می‌سازد تغییر کند؛ می‌توان فرض کرد که $p(x)$ چندجمله‌ایِ با ضرایب صحیحی است که ب.م.م. ضرایبش یک است. ولی آنگاه به وضوح اینکه عضوی از $\mathbb{Z}[x]$ در $\mathbb{Q}[x]$ بر چنین چندجمله‌ای بخش‌پذیر باشد، معادل بخش‌پذیری در $\mathbb{Z}[x]$ برای این چندجمله‌ای‌ها خواهد بود. لذا $p(x)$ای که در بالا انتخاب شد، باید در $\mathbb{Z}[x]$ هم تمامی اعضای I را بشمارد. ولی چون $\deg p > 0$، این با شرط نخستی که مسأله بر ایده‌آلِ I گذاشته در تناقض است. پس ایده‌آلی که اعضای I در $\mathbb{Q}[x]$ می‌سازند کلِ این حلقه است. امری که نتیجه می‌دهد دو چندجمله‌ایِ $f(x), g(x) \in I$ وجود دارند به قسمی که ترکیب خطی‌ای از آنها با ضرایبِ در $\mathbb{Q}[x]$ یک می‌شود. به عبارت دیگر به ازای $p(x), q(x) \in \mathbb{Q}[x]$ مناسبی، داریم: $p(x)f(x) + q(x)g(x) = 1$. با ضرب طرفینِ این تساوی در عدد طبیعی مناسبی مانند N، $Np(x)$ و $Nq(x)$ با ضرایب صحیح خواهند بود و ترکیب خطیِ $f(x)$ و $g(x)$ با این ضرایب-که باید به ایده‌آلِ I از $\mathbb{Z}[x]$ تعلق داشته باشد- چندجمله‌ای ثابتِ N خواهد بود و این $(*)$ را بدست می‌دهد.
حکم دومی که به آن نیاز خواهیم داشت:

(ب) به ازای هر $d, s \in \mathbb{N}$، ایده‌آلِ تولید شده با چندجمله‌ای‌های $(x^s-1)^2$ و d در $\mathbb{Z}[x]$ عنصری به شکلِ $1+x+\cdots+x^{r-1}$ را دربرخواهد داشت.

اثباتِ (ب): کافی است چندجمله‌ای‌ای به صورتِ $1+x+\cdots+x^{r-1}$ بیابیم که به پیمانه‌ی $d \in \mathbb{N}$ با مضربی از $(x^s-1)^2$ همنهشت باشد. قرار دهید: $r = 2sd$. نشان خواهیم داد این عدد ویژگی مطلوب را برآورده می‌کند. می‌توان نوشت:

$$\begin{aligned}\sum_{i=0}^{2sd-1} x^i &= (\sum_{i=0}^{s-1} x^i)(\sum_{i=0}^{2d-1} x^{si})\\ &\equiv (\sum_{i=0}^{s-1} x^i)\big(\sum_{i=0}^{2d-1}(x^{si}-1)\big) \pmod{d}\\ &= (\sum_{i=0}^{s-1} x^i)(x^s-1)\big(\sum_{i=1}^{2d-1}\sum_{j=0}^{i-1} x^{sij}\big)\end{aligned}$$

لذا کافی است نشان دهیم $\sum_{i=1}^{2d-1}\sum_{j=0}^{i-1} x^{sij}$ به ایده‌آلی از $\mathbb{Z}[x]$ که d و x^s-1 می‌سازند تعلق دارد . این چندجمله‌ای به پیمانه‌ی x^s-1 برابر است با $\sum_{i=1}^{2d-1} i = d(2d-1)$ که مضربِ d است؛ امری که نشان می‌دهد چندجمله‌ای مذکور در ایده‌آلی از $\mathbb{Z}[x]$ که d و x^s-1 می‌سازند قرار دارد.
و در نهایت آخرین حکمی که به آن نیاز خواهیم داشت:

(ج) ایده‌آلی از $\mathbb{Z}[x]$ که یک عدد اولِ p و یک چندجمله‌ای مانند $f(x) \in \mathbb{Z}[x]$ با $f(0) = 1$ تولید می‌کنند؛ حتماً دارای عنصری به صورتِ x^t-1 خواهد بود.

اثباتِ (ج): یک همریختیِ $\begin{cases}\mathbb{Z}[x] \to \mathbb{Z}_p[x]\\ g(x) \mapsto \bar{g}(x)\end{cases}$ داریم که توسط آن چندجمله‌ای‌های با ضرایب صحیح به پیمانه‌ی p در نظر گرفته می‌شوند. پس چندجمله‌ایِ $f(x)$ در بالا به عضوی از $\mathbb{Z}_p[x]$ که آن را با $\bar{f}(x)$ نمایش می‌دهیم تصویر می‌شود و این‌که $f(0) = 1$ به آن معنی خواهد بود که در $\mathbb{Z}_p[x]$ -که یک **UFD** است- $\bar{f}(x)$ و x نسبت به هم اولند. از طرف دیگر حکمی استاندارد درباره‌ی میدان‌های متناهی بیان می‌کند که در $\mathbb{Z}_p[x]$:

$$x^{p^n} - x = \prod_{\substack{q(x)\in\mathbb{Z}_p[x]\ \text{تکین و تحویل‌ناپذیر از درجه‌ی}\ d|n}} q(x)$$

لذا اگر درجه‌ی $\bar{f}(x)$، k در نظر گرفته شود، این چندجمله‌ای باید $(x^{p^{k!}} - x)^{p^k} = x^{p^{k!+k}} - x^{p^k} = x^{p^k}(x^{p^{k!}-p^k} - 1)$ را در $\mathbb{Z}_p[x]$ بشمارد. با توجه به آنکه گفتیم در این حلقه‌ی چندجمله‌ایِ $\bar{f}(x)$ و x نسبت به هم اول هستند، این به آن معنی است که $\bar{f}(x) \mid x^t-1$ به ازای یک $t \in \mathbb{N}$. از این موارد می‌بینیم که x^t-1 به ایده‌آلی از $\mathbb{Z}_p[x]$ تعلق دارد که $\bar{f}(x)$ تولید می‌کند: $x^t-1 = \bar{f}(x)\bar{g}(x)$ که در آن $f(x), g(x) \in \mathbb{Z}[x]$. پس در $\mathbb{Z}[x]$ تمامی ضرایب چندجمله‌ایِ $(x^t-1) - f(x)g(x)$ مضربِ p هستند و این همان چیزی است که در پی‌اش بودیم.
در نهایت به حل مسأله می‌پردازیم: با استفاده از (الف) و با توجه به خاصیت دوم ایده‌آلِ I در صورت مسأله، در نظر گرفتنِ حالتی که I ایده‌آلِ تولید شده با یک عدد طبیعیِ $d \in \mathbb{N}$ و یک چندجمله‌ایِ $f(x) \in \mathbb{Z}[x]$ با $f(0) = 1$ باشد از کلیت نمی‌کاهد: $I =< f(x), d >$. با استقرا بر d اثبات خواهیم نمود که این ایده‌آل عضوی به صورت $1+x+\cdots+x^{r-1}$ را دربردارد. پایه‌ی استقرا یعنی وقتی که $d = 1$ بدیهی است چرا که در این حالت ایده‌آلِ

مذکور بر کل حلقه منطبق است. به منظور رسیدن از فرض استقرا به حکم استقرا، p را یک عامل اول دلخواه از d بگیرید. $I =< f(x), d >$ ضرب ایده‌آل‌های $< f(x), p >$ و $< f(x), \frac{d}{p} >$ را شامل می‌شود و از میان این دو ایده‌آل، اولی بنابر (ج) و دومی از فرض استقرا شامل چندجمله‌ای هایی همچون به ترتیب $x^t - ۱$ و $x^{r-۱} + \cdots + x + ۱$ هستند. ولی در $\mathbb{Z}[x]$:

$$(x^t-۱)(۱+x+\cdots+x^{r-۱}) \mid (x^t-۱)(x^r-۱) \mid (x^{rt}-۱)^۲$$

بنابراین چندجمله‌ای ثابت با مقدارِ $d \in \mathbb{N}$ و همچنین $(x^{rt}-۱)^۲$ عضو I هستند و حال (ب) حل را تکمیل می‌کند.

پاسخ ۴. یک سمتِ حکم واضح است: اگر در حلقه‌ای یکدار ضربِ عنصرِ a در تمامی اعضای حلقه صفر شود، بایستی $a = ۰$. عکس آن را با استقرا بر تعداد اعضای R اثبات خواهیم نمود: R را حلقه‌ای جابجایی می‌گیریم با این خاصیت که $aR = \{۰\}$ مستلزم آن است که $a = ۰$. می‌خواهیم نشان دهیم R یکدار است با این فرض که حکم مشابه برای حلقه‌های جابجایی‌ای که تعداد اعضایشان از تعداد اعضای R کمتر باشد معتبر است (پایه‌ی استقرا زمانی است که R تک‌عضوی باشد و بدیهی است). یک $a \in R - \{۰\}$ به دلخواه انتخاب کنید. ادعا می‌کنیم $b \in R$ موجود است به قسمی که $ab = a$. بدین‌منظور ایده‌آلِ

$$I = \{x \in R \mid ax = ۰\}$$

از R را در نظر می‌گیریم (با توجه به جابجایی بودنِ R به وضوح ایده‌آل است). اگر $I = \{۰\}$ که باید نگاشتِ $\begin{cases} R \to R \\ x \mapsto ax \end{cases}$ یک‌به‌یک باشد؛ امری که با توجه به متناهی بودنِ R پوشاییِ آن را هم بدست می‌دهد که در این صورت چون $a \in R$ باید در بُردِ این نگاشت باشد؛ به ازای $b \in R$ای خواهیم داشت $ab = a$ که همان چیزی است که در پی‌اش بودیم. پس فرض کنید ایده‌آلِ I ناصفر نباشد. پس تعداد اعضای حلقه‌ی خارج‌قسمتیِ $\frac{R}{I}$ -که در نظر گرفتنِ آن با توجه به جابجایی بودنِ R خالی از اشکال است- از تعداد اعضای R کمتر است و می‌توان فرضِ استقرا را به آن اعمال کرد: اگر این حلقه یکدار نباشد باید پوچسازی ناصفر همچون $c + I$ داشته باشد. پوچساز بودنِ این عنصر به آن معنی است که $cx \in I$ به ازای هر $x \in R$ و این هم بنابر تعریفِ I یعنی:

$$\forall x \in R : acx = ۰$$

که با توجه به شرط قرار داده شده بر R تنها وقتی می‌تواند رخ دهد که $ac = ۰$ یعنی $c \in I$ و این با ناصفر بودن عنصرِ $c + I$ از حلقه‌ی خارج‌قسمتی در تناقض است. بنابراین به ناچار $\frac{R}{I}$ یکدار است. اگر یکِ حلقه را هم‌دسته‌ی $b + I$ بگیریم، باید $bx - x \in I$ برای هر $x \in R$. مجدداً با استفاده از اینکه R پوچسازی به جز صفر ندارد و تعریفِ I خواهیم داشت:

$$\forall x \in R : a(bx-x) = ۰ \Rightarrow \forall x \in R : (ab-a)x = ۰ \Rightarrow ab = a$$

لذا حکم مطلوب حاصل شد: اگر $a \in R - \{۰\}$ می‌توان $b \in R$ را چنان یافت که $ab = a$. پس با تثبیتِ این a و b، ایده‌آلِ

$$J = \{x \in R \mid bx = x\}$$

از حلقه‌ی جابجاییِ R باید ناصفر باشد چرا که a به آن تعلق دارد. با تکرارِ همان ایده‌ی قبلی، فرض استقرا را به حلقه‌ی خارج‌قسمتیِ $\frac{R}{J}$ اعمال می‌کنیم: این حلقه باید یکدار باشد چرا که اگر $d + J$ را پوچسازی از آن بگیریم، باید $dx \in J$ برای هر x که از تعریفِ J به معنای

$$\forall x \in R : bdx = dx \Rightarrow \forall x \in R : (bd-d)x = ۰$$

است که برقراریِ آن مستلزم $bd = d \Leftrightarrow d \in J \Leftrightarrow d + J = ۰_{\frac{R}{J}}$ است. با توجه به این موارد، باید $\frac{R}{J}$ یکدار باشد. حال اگر $۱_{\frac{R}{J}}$ هم‌دسته‌ی $e + J$ فرض شود، باید $ex - x \in J$ یا معادلاً

$$b(ex-x) = ex - x$$

برای هر $x \in R$. تساویِ فوق را می‌توان به صورتِ $x = (b+e-be)x$ نوشت و لذا $b + e - be$ یکِ حلقه‌ی R خواهد بود و حل به انتها می‌رسد.

پاسخ ۵. با اضافه کردنِ $(۱-\alpha)a_n$ به طرفینِ نامساویِ داده شده در مسأله، می‌بینیم که دنباله‌ی

$$\{b_n := a_n + (۱-\alpha)a_{n-۱}\}_{n=۲}^{\infty}$$

نزولی است. همچنین به وضوح چون $\{a_n\}_{n=۱}^{\infty}$ از پایین کراندار بود، دنباله‌ی مذکور هم چنین است. پس باید به حدی همچون $L \in \mathbb{R}$ همگرا باشد. از این نتیجه خواهیم گرفت که $\{a_n\}_{n=۱}^{\infty}$ هم همگراست. اگر $\alpha = ۱$ که به وضوح $\lim_{n\to\infty} a_n = L$ زیرا $b_n = a_n$. بنابراین فرض می‌کنیم $\alpha \in (۰,۱)$ و اثبات خواهیم کرد که در این حالت $\lim_{n\to\infty} a_n = \frac{L}{۲-\alpha}$. توجه کنید که چون $b_n = a_n + (۱-\alpha)a_{n-۱}$ برای هر $n \geq ۲$، داریم:

$$\forall n \geq ۲ : a_n = \sum_{k=۲}^{n} (\alpha-۱)^{n-k} b_k + (\alpha-۱)^{n-۱} a_۱$$

$\epsilon > ۰$ را دلخواه بگیرید. چون $۰ < \alpha < ۱$ و $\lim_{n\to\infty} b_n = L$ بود، $N \in \mathbb{N}$ای موجود است با این ویژگی که اگر $n > N$ تمامی اعدادِ $|b_n - L|$، $|(\alpha-۱)^n a_۱|$ و $\sum_{k=N}^{\infty} (۱-\alpha)^k$ از ϵ کمترند. اکنون به کمک

تساوی فوق‌الذکر، برای هر $n > 2N$ تخمین‌های زیر را خواهیم داشت:

$$\begin{aligned}
&|a_n - \sum_{k=2}^{n}(\alpha-1)^{n-k}L| \\
&= \Big|\Big(\sum_{k=2}^{n}(\alpha-1)^{n-k}b_k + (\alpha-1)^{n-1}a_1\Big) - \sum_{k=2}^{n}(\alpha-1)^{n-k}L\Big| \\
&\le \Big|\sum_{k=N+1}^{n}(\alpha-1)^{n-k}(b_k-L)\Big| \\
&+ \Big|\sum_{k=2}^{N}(\alpha-1)^{n-k}(b_k-L)\Big| + |(\alpha-1)^{n-1}a_1| \\
&\le \epsilon\sum_{k=0}^{\infty}(1-\alpha)^k + \sup_{k\ge 2}|b_k-L|\overbrace{\sum_{k=N}^{\infty}(1-\alpha)^k}^{\le\epsilon} + \epsilon
\end{aligned}$$

پس اگر M را کران بالایی برای قدرمطلق جملاتِ دنباله‌ی همگرای $\{b_n\}_{n=2}^{\infty}$ بگیریم، نامساوی‌های فوق نشان می‌دهند که برای هر $n > 2N$:

$$\begin{aligned}
|a_n - \sum_{k=0}^{n-2}(\alpha-1)^k L| &= |a_n - \sum_{k=2}^{n}(\alpha-1)^{n-k}L| \\
&\le \frac{\epsilon}{\alpha} + (M+|L|)\epsilon + \epsilon
\end{aligned}$$

چون $\epsilon > 0$ دلخواه بود، این نشان می‌دهد که:

$$\lim_{n\to\infty} a_n = \sum_{k=0}^{\infty}(\alpha-1)^k L = \frac{L}{2-\alpha}$$

که حل را تکمیل می‌کند.

پاسخ ۶. ابتدا توجه کنید که تنها کافی است سمت راستِ نامساویِ داده شده را اثبات نماییم. زیرا با فرض آنکه

$$\forall x\in[0,1] : g(x) \le \int_0^1 \big(g(t) + |\, g'(t)\,|\big)\mathrm{d}t$$

به ازای هر تابع C^1ی $g : [0,1] \to \mathbb{R}$؛ اعمالِ این نامساوی به $-g$ با کمی ساده‌سازی نتیجه خواهد داد:

$$\forall x\in[0,1] : g(x) \ge \int_0^1 \big(g(t) - |\, g'(t)\,|\big)\mathrm{d}t$$

در ادامه سمت راستِ نامساوی را با برهان‌خلف ثابت خواهیم کرد. M و m را به ترتیب ماکسیمم و مینیمم مطلقِ g بگیرید و فرض کنید در نقاطِ به ترتیب x_2 و x_1 از $[0,1]$ رخ دهند. فرض‌خلف بیان می‌کند که

$$\int_0^1 \big(g(t) + |\, g'(t)\,|\big)\mathrm{d}t < M$$

از قضیه‌ی مقدار میانگین برای انتگرال بایستی $\int_0^1 g(t)\mathrm{d}t = g(x_3)$ به ازای یک نقطه‌ی مناسبِ $x_3 \in [0,1]$. پس:

$$g(x_3) < M - \int_0^1 |\, g'(t)\,|\,\mathrm{d}t$$

ادعا می‌کنیم $\int_0^1 |\, g'(t)\,|\,\mathrm{d}t \ge M - m$ و آنگاه ترکیب دو نامساویِ اخیر نتیجه خواهد داد $g(x_3) \le m$ با شیوه‌ی انتخابِ m مباینت دارد. لذا برای اتمام کار کافی است صحت آخرین نامساویِ ظاهر شده در بالا را نشان دهیم. داریم:

$$\begin{aligned}
\int_0^1 |\, g'(t)\,|\,\mathrm{d}t &\ge \int_0^{x_1}\big(-g'(t)\big)\mathrm{d}t + \int_{x_1}^1 g'(t)\mathrm{d}t \\
&= \big(g(0) - g(x_1)\big) + \big(g(1) - g(x_1)\big) = g(0) + g(1) - 2m
\end{aligned}$$

و مشابهاً

$$\begin{aligned}
\int_0^1 |\, g'(t)\,|\,\mathrm{d}t &\ge \int_0^{x_2} g'(t)\mathrm{d}t + \int_{x_2}^1 \big(-g'(t)\big)\mathrm{d}t \\
&= \big(g(x_2) - g(0)\big) + \big(g(x_2) - g(1)\big) = 2M - g(0) - g(1)
\end{aligned}$$

اکنون میانگین‌گیری از این دو نامساوی، حکم مطلوب مبنی بر

$$\int_0^1 |\, g'(t)\,|\,\mathrm{d}t \ge M - m$$

را بدست می‌دهد و حل تمام است.

پاسخ ۷. برای هر $1 \le i \le n$، $h_i : \mathbb{R}^n \to \mathbb{R}$ را تابع C^2ی

$$h_i(x) = f_i(x) - \frac{1}{2}\sum_{j=1}^{n} c_{ij}x_j$$

می‌گیریم. با تعویضِ جای i و j در تساوی داده شده در صورت مسأله، می‌بینیم که همواره باید $c_{ij} = -c_{ji}$. پس

$$\begin{aligned}
\forall 1 \le i,j \le n : \frac{\partial h_i}{\partial x_j} - \frac{\partial h_j}{\partial x_i} &= \Big(\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i}\Big) - \frac{1}{2}(c_{ij} - c_{ji}) \\
&= c_{ij} - \frac{1}{2}(c_{ij} - c_{ji}) = 0
\end{aligned}$$

حکمی کلاسیک بیان می‌کند که وجودِ چنین ویژگی‌ای برای nتابع C^1 مانندِ $h_1, \ldots, h_n$ بر $\mathbb{R}^n$، به معنای آن است که میدان برداریِ $x \mapsto \big(h_1(x), \ldots, h_n(x)\big)$ بر $\mathbb{R}^n$ یک میدانِ گرادیان است؛ یعنی یک تابعِ $u : \mathbb{R}^n \to \mathbb{R}$ موجود است چنان‌که $\frac{\partial u}{\partial x_i} = h_i$ برای هر i [3]. حال تابعِ $g := -u$ ویژگیِ مطلوب در صورتِ مسأله را برآورده می‌کند. چرا که از تساویِ پیشین و شیوه‌ی تعریفِ h_iها، به ازای هر iتابعِ

$$f_i(x) + \frac{\partial g}{\partial x_i} = \Big(h_i(x) + \frac{1}{2}\sum_{j=1}^{n} c_{ij}x_j\Big) - \frac{\partial u}{\partial x_i} = \frac{1}{2}\sum_{j=1}^{n} c_{ij}x_j$$

خطی است.

[3] به عبارت دقیق‌تر، از تساوی‌های

$$\forall 1 \le i,j \le n : \frac{\partial h_i}{\partial x_j} = \frac{\partial h_j}{\partial x_i}$$

نتیجه می‌شود که ۱-فرمِ

$$h_1 dx_1 + h_2 dx_2 + \cdots + h_n dx_n$$

بسته است. ولی طبق لم پوآنکاره هر ۱-فرمِ بسته بر $\mathbb{R}^n$ باید دقیق باشد. این دقیقاً به معنای وجودِ uای با خاصیت فوق است.

پاسخ ۸. $P_1, \dots, P_n$ بازه‌ی $(0,1)$ را به $n+1$ زیربازه افراز می‌کنند. جمع مکعبات طولِ این زیربازه‌ها را به x_n نمایش داده و قرار می‌دهیم $x_0 = 1$. وقتی افراز حاصل از $n-1$ نقطه‌ی $P_1, \dots, P_{n-1}$ در نظر گرفته شده باشد؛ با اضافه کردنِ نقطه‌ی P_n و رسیدن به یک افرازِ جدید، تغییری که در طول زیربازه‌های ظاهر شده در افراز رخ می‌دهد آن است که بازه‌ای به طولِ $a_n + b_n$ با دو زیربازه به طول‌های a_n و b_n جایگزین می‌شود. پس:

$$x_n - x_{n+1} = (a_n + b_n)^3 - a_n^3 - b_n^3 = 3a_n b_n (a_n + b_n)$$

لذا مجموع داده شده در صورت مسأله به طور تلسکوپی ساده شده و با توجه به $x_0 = 1$ برابر خواهد بود با $\frac{1}{3} - \lim_{n\to\infty} x_n$. تنها اثبات صفر بودنِ این حد باقی می‌ماند: باید نشان دهیم در افرازی از $(0,1)$ که نقاطِ $P_1, \dots, P_n$ معین می‌کنند، جمع مکعباتِ طول‌های زیربازه‌های حاصل از افراز، وقتی $n \to \infty$ به صفر می‌رود. به منظور اثبات، $0 < \epsilon < 1$ را دلخواه فرض کنید. چون P_iها در $(0,1)$ چگالند، اگر N به اندازه‌ی کافی بزرگ انتخاب شود در هریک از زیربازه‌های

$$(*) \quad (0, \epsilon), (\epsilon, 2\epsilon), \dots \big((k-1)\epsilon, k\epsilon\big), (k\epsilon, 1)$$

که در آن $k := \lfloor \frac{1}{\epsilon} \rfloor$ یکی از نقاطِ $P_1, \dots, P_N$ قرار خواهد داشت. حال فرض کنید $n > N$ و نقاط متمایزِ $P_1, \dots, P_n$ از $(0,1)$ در ترتیب صعودی به شکلِ $y_n > \dots > y_1$ مرتب شده باشند. می‌خواهیم $x_n = \sum_{i=1}^{n-1} |y_{i+1} - y_i|^3$ را تخمین بزنیم. زیربازه‌های ظاهر شده در $(*)$ را به ترتیب $I_1, \dots, I_k, I_{k+1}$ می‌نامیم: $I_i = \big((i-1)\epsilon, i\epsilon\big)$ برای هر $1 \le i \le k$ و $I_{k+1} = (k\epsilon, 1)$. طول هیچ‌یک از I_iها از ϵ تجاوز نمی‌کند و طبق شیوه‌ی انتخابِ n، باید در هریک از این زیربازه‌ها یکی از $y_1, \dots, y_n$ واقع باشد. برای هر $1 \le j \le k+1$، i_j را بزرگترین $1 \le s \le n$ای بگیرید که به ازای آن $y_s \in I_j$ و قرار دهید $i_0 = 1$. اکنون

$$1 = i_0 < i_1 < \dots < i_k < i_{k+1} = n$$

و داریم:

$$(**) \quad \begin{aligned} x_n &= \sum_{i=1}^{n-1} |y_{i+1} - y_i|^3 \\ &= \sum_{j=1}^{k+1} \sum_{i=i_{j-1}+1}^{i_j - 1} |y_{i+1} - y_i|^3 + \sum_{j=1}^{k} |y_{i_j+1} - y_{i_j}|^3 \end{aligned}$$

در اینجا به ازای هر $1 \le j \le k+1$ نقاطِ $y_{i_{j-1}+1} < \dots < y_{i_j}$ در بازه‌ی $\overline{I_j}$ واقعند که طولِ آن حداکثر ϵ است و لذا

$$\begin{aligned} \sum_{i=i_{j-1}+1}^{i_j - 1} |y_{i+1} - y_i|^3 &\le \Big(\sum_{i=i_{j-1}+1}^{i_j - 1} (y_{i+1} - y_i) \Big)^3 \\ &= (y_{i_j} - y_{i_{j-1}+1})^3 \le \epsilon^3 \end{aligned}$$

در موردِ $\sum_{j=1}^{k} |y_{i_j+1} - y_{i_j}|^3$ که آن هم در $(**)$ ظاهر شده؛ توجه کنید که این مجموع حداکثر $8k\epsilon^3$ است چرا که چون هریک از زیربازه‌های ظاهر شده در $(*)$ یکی از $y_n > \dots > y_1$ را شامل می‌شد، اختلاف هر دوتای متوالی از این اعداد حداکثر باید 2ϵ باشد. با قرار دادن این موارد در $(**)$، برای هر $n > N$ خواهیم داشت:

$$x_n \le 9k\epsilon^3 < 9(1 + \frac{1}{\epsilon})\epsilon^3$$

چون $0 < \epsilon$ دلخواه بود و هنگامی که به صفر میل کند سمت راست نامساوی بالا هم به صفر می‌رود؛ این $\lim_{n\to\infty} x_n = 0$ را اثبات می‌کند و سپس از آنچه که گفته بودیم، حکم مطلوب از حد فوق حاصل می‌شود.

پاسخ ۹. می‌دانیم که با داشتنِ یک عملگر خودالحاق بر یک فضای ضرب داخلیِ حقیقی و متناهی‌البعد، می‌توان پایه‌ای یکه‌متعامد برای فضا متشکل از بردار ویژه‌های آن عملگر ارائه کرد. اگر فضای ضرب داخلیِ مذکور را $\mathbb{R}^n$ مجهز به ضرب داخلیِ معمولش بگیریم، این حکم تبدیل می‌شود به اینکه هر ماتریس متقارن با درایه‌های حقیقی را می‌توان به وسیله‌ی یک ماتریس متعامد از همان مرتبه با درایه‌های حقیقی قطری کرد. با اعمالِ این حکم به B، وجود یک ماتریسِ 2×2 و متعامدِ P (یعنی $PP^T = \mathrm{I}_2$) با درایه‌های حقیقی نتیجه می‌شود که B را قطری می‌کند:

$$P^{-1}BP = P^T BP = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$$

با جایگزین کردنِ A و B با به ترتیب $P^T AP$ و $P^T BP$، شرایط مسأله تغییر نمی‌کند: این ماتریس‌ها با توجه به متعامد بودنِ P با به ترتیب A و B مزدوج هستند و لذا همواره

$$\mathrm{tr}(A + Bt)^m = \mathrm{tr}\big(P^T AP + (P^T BP)t\big)^m$$

زیرا $P^{-1}(A + Bt)^m P = (A + Bt)^m$؛ ماتریس‌ها با این تغییر کماکان متقارن می‌ماند و در نهایت همچنان نامنفی خواهند بود:

$$\forall X \in \mathbb{R}^2 : \begin{cases} X^T (P^T AP) X = (PX)^T A (PX) \ge 0 \\ X^T (P^T BP) X = (PX)^T B (PX) \ge 0 \end{cases}$$

این موارد نشان می‌دهند که بدون کاسته شدن از کلیت می‌توانیم توجه خود را به حالتی معطوف کنیم که $B = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$ قطری باشد. ماتریس متقارنِ A را به صورتِ $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ بگیرید. هرگاه در شرط نامنفی بودن: $X^T AX \ge 0$ و $X^T BX \ge 0$، به جای X اعضای پایه‌ی استاندارد قرار گیرند، نتیجه می‌شود که درایه‌های روی قطر در این دو ماتریس باید نامنفی باشند: $(*) : a, c, \lambda, \mu \ge 0$. پس کافی است نشان دهیم با فرضِ این نامساوی‌ها، ضرایب چندجمله‌ایِ

$$(**) \quad t \mapsto \mathrm{tr}\Big(\begin{bmatrix} a + \lambda t & b \\ b & c + \mu t \end{bmatrix}^m \Big)$$

نامنفی‌اند. بدین منظور ادعا می‌کنیم:

(***) برای هر $m \in \mathbb{N}$ چندجمله‌ای‌های

$$P_m(t), Q_m(t), R_m(t) \in \mathbb{R}[t]$$

با ضرایب نامنفی موجودند به قسمی که

$$\begin{bmatrix} a+\lambda t & b \\ b & c+\mu t \end{bmatrix}^m = \begin{bmatrix} P_m(t) & bQ_m(t) \\ bQ_m(t) & R_m(t) \end{bmatrix}$$

این، حکمِ مطلوب را بدست خواهد داد. زیرا در این صورت چندجمله‌ایِ ظاهر شده در $(**)$ با $P_m(t)+R_m(t)$ داده خواهد شد که با ضرایب نامنفی است. $(***)$ را با استقرا بر m ثابت خواهیم کرد: حالتِ $m=۱$ با توجه به اینکه $P_m(t)=a+\lambda t$، $R_m(t)=c+\mu t$ و $Q_m(t)=۱$ از $(*)$ حاصل می‌شود. با فرض درستیِ $(***)$ برای m، آن را برای $m+۱$ بدست می‌آوریم:

$$\begin{aligned} \begin{bmatrix} P_{m+۱}(t) & bQ_{m+۱}(t) \\ bQ_{m+۱}(t) & R_{m+۱}(t) \end{bmatrix} &= \begin{bmatrix} a+\lambda t & b \\ b & c+\mu t \end{bmatrix}^{m+۱} \\ &= \begin{bmatrix} a+\lambda t & b \\ b & c+\mu t \end{bmatrix} \begin{bmatrix} P_m(t) & bQ_m(t) \\ bQ_m(t) & R_m(t) \end{bmatrix} \\ &= \begin{bmatrix} (a+\lambda t)P_m(t)+b^۲Q_m(t) & b\big((a+\lambda t)Q_m(t)+R_m(t)\big) \\ b\big(P_m(t)+(c+\mu t)Q_m(t)\big) & b^۲Q_m(t)+(c+\mu t)R_m(t) \end{bmatrix} \end{aligned}$$

$$\Rightarrow \begin{cases} P_{m+۱}(t)=(a+\lambda t)P_m(t)+b^۲Q_m(t) \\ Q_{m+۱}(t)=(a+\lambda t)Q_m(t)+R_m(t) \\ R_{m+۱}(t)=(c+\mu t)R_m(t)+b^۲Q_m(t) \end{cases}$$

با توجه به فرض استقرا مبنی بر نامنفی بودن ضرایبِ $P_m(t)$، $Q_m(t)$ و $R_m(t)$ و حکمِ $(*)$ مبنی بر $a,c,\lambda,\mu \ge ۰$، تساوی‌های فوق نامنفی بودن ضرایبِ $P_{m+۱}(t)$، $Q_{m+۱}(t)$ و $R_{m+۱}(t)$ را نتیجه می‌دهند و حکم استقرا اثبات می‌شود.

پاسخ ۱۰. پاسخ منفی است. یک مثال نقض خواهیم ساخت. ابتدا توجه کنید که می‌توان مسأله را تقلیل داد به حالتی که در آن A شکل ساده‌ای داشته باشد. r را رتبه‌ی A بگیرید. پس ماتریس‌های $n\times n$ و وارون‌پذیرِ P,Q با درایه‌های در F موجودند به قسمی که:

$$PAQ = \begin{bmatrix} \mathrm{I}_r & \mathrm{O}_{r\times(n-r)} \\ \mathrm{O}_{(n-r)\times r} & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix}$$

که در آن منظور از $\mathrm{O}_{k\times l}$ و I_s به ترتیب ماتریسِ صفرِ مرتبه‌ی $k\times l$ ماتریسِ همانیِ مرتبه‌ی $s\times s$ است. با جایگزین کردنِ A و B در صورت مسأله با به ترتیب PAQ و $Q^{-۱}BP^{-۱}$، مسأله تقلیل می‌یابد به حالتی که در آن A این شکل بلوکی را دارد. چرا که:

$$\begin{cases} (PAQ)(Q^{-۱}BP^{-۱})(PAQ)(Q^{-۱}BP^{-۱}) = P(ABAB)P^{-۱} \\ \quad = \mathrm{O}_{n\times n} \Leftrightarrow ABAB = \mathrm{O}_{n\times n} \\ (Q^{-۱}BP^{-۱})(PAQ)(Q^{-۱}BP^{-۱})(PAQ) = Q^{-۱}(BABA)Q \\ \quad = \mathrm{O}_{n\times n} \Leftrightarrow BABA = \mathrm{O}_{n\times n} \end{cases}$$

لذا در ادامه A را با شکل بلوکیِ فوق در نظر گرفته و تجزیه‌ی بلوکیِ متناظر برای B را چنین می‌گیریم:

$$B = \begin{bmatrix} X_{r\times r} & Y_{r\times(n-r)} \\ Z_{(n-r)\times r} & W_{(n-r)\times(n-r)} \end{bmatrix}$$

که در آن X، Y، Z و W ماتریس‌هایی با درایه‌های در F از مرتبه‌های به ترتیب $r\times r$، $r\times(n-r)$، $(n-r)\times r$ و $(n-r)\times(n-r)$ هستند. حال با ضرب بلوکی بررسی خواهیم کرد که تساوی‌های $ABAB=\mathrm{O}_{n\times n}$ و $BABA=\mathrm{O}_{n\times n}$ به چه معنی هستند:

$$\begin{cases} AB = \begin{bmatrix} \mathrm{I}_r & \mathrm{O}_{r\times(n-r)} \\ \mathrm{O}_{(n-r)\times r} & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} \\ = \begin{bmatrix} X & Y \\ \mathrm{O}_{(n-r)\times r} & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \\ \Rightarrow ABAB = \begin{bmatrix} X^۲ & XY \\ \mathrm{O}_{(n-r)\times r} & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \\ BA = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} \begin{bmatrix} \mathrm{I}_r & \mathrm{O}_{r\times(n-r)} \\ \mathrm{O}_{(n-r)\times r} & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \\ = \begin{bmatrix} X & \mathrm{O}_{r\times(n-r)} \\ Z & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \\ \Rightarrow BABA = \begin{bmatrix} X^۲ & \mathrm{O}_{r\times(n-r)} \\ ZX & \mathrm{O}_{(n-r)\times(n-r)} \end{bmatrix} \end{cases}$$

لذا پاسخ در صورتی مثبت است که بتوان از $X^۲ = \mathrm{O}_{r\times r}$ و $XY = \mathrm{O}_{r\times(n-r)}$ نتیجه گرفت $ZX = \mathrm{O}_{(n-r)\times r}$. ولی به سادگی می‌توان مثالی زد که این برقرار نباشد. به عنوان نمونه در حالتِ $n=۳, r=۲$ قرار دهید

$$X = \begin{bmatrix} ۰ & ۱ \\ ۰ & ۰ \end{bmatrix} \quad Y = \begin{bmatrix} ۱ \\ ۰ \end{bmatrix} \quad Z = \begin{bmatrix} ۱ & ۰ \end{bmatrix}$$

بنابراین اگر داشته باشیم:

$$A = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۰ \end{bmatrix} \quad B = \begin{bmatrix} ۰ & ۱ & ۱ \\ ۰ & ۰ & ۰ \\ ۱ & ۰ & ۰ \end{bmatrix}$$

آنگاه $(AB)^۲$ صفر است و $(BA)^۲$ ناصفر. امری که به سادگی با محاسبه‌ی مستقیم هم قابل تحقیق است:

$$ABAB = \begin{bmatrix} ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ \end{bmatrix} \quad BABA = \begin{bmatrix} ۰ & ۰ & ۰ \\ ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ \end{bmatrix}$$

معرفی آزمایشگاه الگوریتم‌ها و حل مساله

گسترش فرهنگ علمی و استفاده از علوم پایه همچون ریاضی و علوم کامپیوتر نیازمند جذب و پرورش افراد مستعد در این زمینه و آشناسازی ایشان با زمینه‌های کاری و کمک به جهت‌یابی‌شان در مسیرهایی است که نیاز جامعه‌ی ما را در عرصه‌های مرتبط برطرف کند.

در این راستا، در دو دهه ۷۰ و ۸۰ هر ساله در دانشکده‌ی علوم ریاضی دانشگاه صنعتی شریف، فعالیت‌های ترویجی و به خصوص برنامه‌هایی با مخاطب دانش‌آموز برگزار می‌شده است. این فعالیت‌ها در قالب برگزاری کارگاه‌های آموزشی، توسعه‌ی باشگاه‌های اینترنتی(شبکه‌ی مدرسه)، طراحی خانه ریاضیات، و همکاری با المپیادهای ریاضی و کامپیوتر و برگزاری مسابقات ریاضی و کامپیوتر در دو دهه گذشته به انباشت تجربیات ارزنده‌ای انجامیده است.

آزمایشگاه الگوریتم‌ها و حل مساله مبتنی بر این تجربیات، با احساس نیاز وجود نهادی برای سازمان‌دهی جدی‌تر و برای انجام فعالیت‌های پژوهش، توسعه و ترویج در زمینه‌ی آموزش ریاضی و کامپیوتر در سال ۱۳۹۰ تاسیس شده است. انجام پژوهش‌های بنیادی و کاربردی در زمینه آموزش ریاضی و کامپیوتر و ترویج و توسعه فرهنگ علوم ریاضی از اهداف و برنامه‌های اصلی این آزمایشگاه است.

چشم‌انداز

ایجاد پل ارتباطی بین آموزش دانشگاهی و دبیرستانی-عمومی در زمینه‌ی علوم ریاضی و کامپیوتر و فراهم کردن بستری مناسب برای ارتباط موثر نسل‌های مختلف دانش‌آموختگان علوم ریاضی در جهت انتقال اطلاعات و تجربه‌ها در زمینه‌های مختلف فعالیت تخصصی ایشان در راستای ترویج فرهنگ علمی.

مأموریت

توسعه پژوهش‌های بنیادی و کاربردی و ترویج علوم ریاضی با انجام فعالیت‌های آموزشی متناسب با سطوح مختلف مخاطبان در سه سطح عمومی، دانش‌آموزی و دانشجویی.

اهداف

- تولید و توسعه‌ی محتوای آموزشی موثر در جهت آشنایی بهتر مخاطبان با زمینه‌های فعالیت علمی در حوزه‌ی علوم ریاضی و کامپیوتر.
- ایجاد شبکه‌ی اجتماعی دانشجویان، فارغ‌التحصیلان و اساتید علاقمند و درگیر کردن ایشان در پروژه‌های جمعی مرتبط.
- ایجاد ارتباط بین نهادهای مختلف مرتبط با موضوع.

برنامه‌ها

برنامه‌های آزمایشگاه الگوریتم و حل مساله در راستای چشم‌انداز و مأموریت، در محورهای زیر در نظر گرفته می‌شود:

- پژوهش‌های بنیادی
- پژوهش‌های کاربردی
- برگزاری کارگاه‌های آموزشی، جشنواره‌ها و مسابقات
- توسعه‌ی باشگاه‌های اینترنتی و آموزش الکترونیکی (با تکیه بر فضای اینترنت و تلفن همراه)
- برگزاری مسابقات
- تولید محتوای علمی و انتشار گزارش‌های فنی، نشریات و کتب

اساتید و همکاران

آزمایشگاه تحت سرپرستی دکتر یحیی تابش عضو هیأت علمی دانشکده علوم ریاضی دانشگاه صنعتی شریف اداره می‌شود. برنامه‌های آزمایشگاه توسط گروه‌های کاری زیر نظر کمیته علمی-اجرایی آزمایشگاه اجرا می‌شود. آزمایشگاه علاوه بر بهره‌گیری از فعالیت دانش‌جویان و فارغ‌التحصیلان دانشکده علوم ریاضی، تاکنون از کمک و همراهی اساتیدی چون دکتر سیاوش شهشهانی، دکتر دانشگر، دکتر فتوحی، دکتر رزوان، دکتر اصغری، دکتر علیشاهی، دکتر عباسیان، دکتر رمضانیان، دکتر امین‌زاده و دکتر خزایی بهره برده و از همکاری و نظرات همگی اساتید و دانشجویان و فارغ‌التحصیلان گرامی استقبال می‌کند.

در صورت علاقه به زمینه‌های کاری آزمایشگاه و برای اطلاع از برنامه‌های گذشته و آینده می‌توانید به وب‌سایت http://ApLab.sharif.ir مراجعه کنید و به وسیله پست الکترونیکی ApLab@sharif.ir با ما در تماس باشید.

گزارش مدرسه‌ی زمستانی موبایل، اسفند ۹۲

مدرسه‌ی زمستانی موبایل عنوان برنامه‌ای است که در روزهای ۷، ۸ و ۹ اسفندماه ۱۳۹۲ در دانشگاه صنعتی شریف برگزار شد. این مدرسه با هدف آشنایی دانش‌آموزان با ریاضیات دانشگاهی – که پایه‌ی علم موبایل و مخابرات را تشکیل می‌دهند– با همکاری اساتید، دانشجویان و فارغ‌التحصیلان دانشکده‌ی ریاضی، جمعی از دانشجویان دانشکده‌ی مهندسی برق و دانشکده‌ی مهندسی کامپیوتر دانشگاه صنعتی شریف برگزار شد.

در این مدرسه، ۱۰۰ دانش‌آموز از مدارس تهران، کرج، قزوین، اصفهان و زابل شرکت داشتند. این دانش‌آموزان در مقاطع اول، دوم و سوم دبیرستان مشغول به تحصیل بودند.

برنامه شامل چندین بخش بود که در ادامه هر بخش را به صورت مجزا توضیح خواهیم داد.

سخنرانی‌های علمی

روزهای اول و دوم این برنامه با یک سخنرانی علمی آغاز شد. این سخنرانی‌ها به صورت عمومی و برای تمامی دانش‌آموزان برگزار گردید. سخنرانی روز اول توسط دکتر شهرام خزایی از اساتید دانشکده‌ی علوم ریاضی با موضوع "سیستم‌های مخابراتی دیجیتال و امنیت آن‌ها" انجام شد. چکیده‌ی این سخنرانی به شرح زیر است:

سیستم‌های مخابراتی دیجیتال که نسل جدیدی از ارتباطات را برای بشر فراهم کرده است، روز به‌روز در حال رشد و گسترش هستند. در این سیستم‌ها یک منبع آنالوگ تولید اطلاعات (مانند خروجی یک میکروفن) یک سیگنال آنالوگ تولید می‌کند. این سیگنال آنالوگ به نحو مناسبی کد و تا حد ممکن فشرده شده که به این امر کدینگ منبع گفته می‌شود. اطلاعات کد شده را که معمولا رشته‌ای از بیت‌های صفر و یک هستند، نمی‌توان به همین صورت از یک کانال ارتباطی ارسال کرد. زیرا به علت وجود عواملی مانند نویز اطلاعات با خطا مواجه شده و به درستی به مقصد نمی‌رسند. برای ارسال صحیح اطلاعات به دست گیرنده، با افزودن اطلاعاتی اضافی به پیام فشرده شده (که افزونگی نامیده می‌شود.) شرایط را برای دریافت صحیح اطلاعات فراهم می‌سازند. اما ارسال صحیح اطلاعات تازه آغاز راه است؛ زیر افراد سود جو همواره بدنبال سوء استفاده از اطلاعات ارسالی در کانال‌های مخابراتی هستند. برای مبارزه با چنین خطراتی باید برای امن کردن اطلاعات ارسالی از ابزاری مانند رمزنگاری استفاده کرد. در این سخنرانی بخش‌های مختلف سیستم‌های مخابراتی دیجیتال معرفی شدند و مخاطبان با مفاهیم امنیتی شبکه‌های مخابراتی آشنا شدند.

اما در سخنرانی روز دوم به موضوعی که جزو مباحث داغ روز محسوب می‌شود، پرداخته شد. این سخنرانی با عنوان "آزادی و پیشرفت" توسط کوهیار حاجی‌سلیمی از دانشجویان دانشکده‌ی علوم ریاضی انجام شد. چکیده‌ی این سخنرانی به این شکل است:

در این بخش قرار است با شروع از مفهوم نرم‌افزارهای «متن‌باز»، تعریفی از آزادی در دنیای تکنولوژی ارایه دهیم. مقداری درباره‌ی قوانین حق انحصاری ثبت یک نوآوری (patenting) صحبت کنیم و ببینیم این قوانین، به خصوص در زمینه‌ی موبایل، از کجا موجب نوآوری‌های بیش‌تر می‌شوند و تا کجا جلوی نوآوری را می‌گیرند.

سپس به بحث داغ آزاد بودن سیستم‌عامل اندروید می‌پردازیم و مشکلات و فایده‌هایی که این آزادی برای اندروید ایجاد کرده است را بررسی می‌کنیم. صحبت‌ها بیش‌تر «مباحثه‌محور» بوده و شما باید در شکل‌گیری نتیجه‌گیری‌ها شرکت کنید؛ نتیجه‌ی نهایی بحث هنوز مشخص نیست!

کارگاه‌های آموزشی

در واقع می‌توان گفت که کارگاه‌های آموزشی هسته‌ی اصلی این برنامه بوده‌است. هدف این کارگاه‌ها آموزش مباحث دانشگاهی به دانش‌آموزان است. اما با کارسوق‌هایی که به صورت معمول در ایران برگزار می‌شود، تفاوت ویژه‌ای دارد. در واقع در این کارگاه‌ها، مربیان تنها مسیر را برای فراگیری دانش‌آموزان فراهم می‌آورند و یادگیری به عهده‌ی خود دانش‌آموزان است. سعی بر این بوده است که مفاهیم پیچیده‌ی دانشگاهی در قالب بازی‌های ساده و قابل فهم گنجانده شده و دانش‌آموزان در حین انجام این فعالیت‌ها بتوانند این مفاهیم را درک کنند. در انتخاب موضوعات کارگاه‌ها سعی شده است ریاضیات و علوم مرتبط با تلفن‌های همراه و مخابرات پوشش داده شود.

این کارگاه‌ها در دو روز اول یعنی روزهای هفتم و هشتم اسفندماه در دو نوبت صبح و بعدازظهر برگزار گردید. هر کارگاه با ظرفیت ۲۵ نفر به مدت ۳ ساعت اجرا شد. برای این برنامه شش کارگاه در نظر گرفته شده بود و در دو روز اول برنامه، هر گروه در ۴ کارگاه شرکت داشتند.

خلاصه‌ای از مباحث مطرح شده در این کارگاه‌ها در ادامه می‌آید:

تعامل انسان و رایانه

اگر موشواره نبود و تنها باید از صفحه‌کلید استفاده می‌کردید، آیا باز هم بسیاری از کارهایی که با رایانه انجام می‌دهید به همین سادگی بود؟! آیا به چیدمان صفحه‌کلید رایانه‌تان فکر کرده‌اید؟! اگر کلیدها به گونه‌ای دیگر چیده شده بودند کار با آن سخت‌تر می‌شد یا ساده‌تر؟! «تعامل انسان و رایانه» با هدف بررسی و مطالعه‌ی چنین مواردی شکل گرفت. در واقع تعامل انسان و رایانه به بررسی، برنامه‌ریزی و طراحی اثری که رایانه و کاربرانش بر یکدیگر می‌گذارند، می‌پردازد. علمی که در آن رشته‌های متعددی از قبیل علوم رایانه، علوم رفتاری و طراحی گرد هم می‌آیند تا کاربر بتواند به بهترین نحو ممکن با «ماشینی» که سروکار دارد رابطه برقرار کند و از تمامی قابلیت‌هایش به خوبی بهره ببرد و در مقابل «ماشین» نیاز کاربر را به درستی دریابد و آن‌ها را تا حد امکان برآورده سازد. آشنایی با این مبحث و پروژه‌های کوچک و بزرگی از این دست، تشخیص نیازهای پیرامونمان، و در نهایت طراحی بهترین «ماشین» برای رفع این نیازها با امکاناتی که در اختیار داریم، هدفی است که این کارگاه دنبال می‌کند.

دسترسی چندگانه

به تلویزیون که به یکی از پراستفاده‌ترین وسایل بشر تبدیل شده است دقت کرده‌اید؟ تنها با فشردن یک دکمه می‌توان شبکه‌ی خبر را رها کرد و به تماشای برنامه‌ی ۹۰ نشست! تعداد شبکه‌های تلویزیونی و ماهواره‌ای روز به روز رو به افزایش است و ما هیچ‌گاه انتظار نداریم که تصاویر شبکه‌ی ورزش از شبکه‌ای دیگر پخش شود! چگونه ممکن است که تمامی شبکه‌ها به کار خود ادامه می‌دهند بدون اینکه اختلالی در کار یکدیگر ایجاد کنند؟!

آیا در یک مکان بسیار شلوغ قرار گرفته‌اید که بخواهید از تلفن همراه خود استفاده کنید و با مشکل روبه‌رو شوید؟! یا حتی با گرفتن شماره‌ی دوستتان از طریق تلفن، با مشکل اشغال بودن شبکه مواجه شوید، حتی اگر تلفن دوستتان اشغال نباشد؟!

اینکه ممکن است ما همزمان به شبکه‌های مختلف دسترسی داشته باشیم، یا در یک منطقه تعداد زیادی از افراد بتوانند با تلفن همراه خود صحبت کنند، امکانی است که تکنیک‌های دسترسی چندگانه برای ما به ارمغان آورده است. هدف ما در این بخش این است که همه‌ی افراد شرکت‌کننده بتوانند همزمان به اطلاعات مورد نیاز خود دست یابند و اطلاعاتی که نیاز است را منتقل کنند!!

مدار منطقی

تا به حال شده است که به کنجکاوی ذهن خود پاسخ گفته باشید و یک ماشین‌حساب ساده (یا وسیله‌ی از این دست) را از هم بگشایید تا به اعداد درون آن دست یابید؟!! می‌توان با زدن چند دکمه دو عدد را با هم جمع کرد، از هم کم کرد، ضرب کرد و یا عمل تقسیم را روی آن دو انجام داد! و تنها چیزی که صرف می‌شود کمی از انرژی باطری ماشین‌حسابتان است! تنها کافی است تا دوشاخه‌ی رایانه‌ی خود را

به برق وصل کنید تا با آن بتوانید هزاران کار سخت‌تر و پیچیده‌تر از جمع و ضرب را انجام دهید! جالب این‌جاست که این «ماشین»ها تنها دو عدد را آموخته‌اند: صفر و یک! عدم وجود جریان الکتریکی و وجود جریان الکتریکی!

ریاضیات پشت‌پرده‌ی این صفر و یک‌ها زیبایی خاصی دارد و دنیای جدیدی را به روی آدم می‌گشاید! این ریاضیات آغاز بحث ما در این‌جاست! ریاضیاتی که پایه و اساس ساده‌ترین تا پیچیده‌ترین این «ماشین»هاست! در ادامه‌ی این ریاضیات به کنجکاوی ذهن شما پاسخ خواهیم گفت!

رمزنگاری

به کمک خدمات جدید ارائه شده از سوی اپراتورهای تلفن همراه، انجام پرداخت قبوض، کارهای بانکی و بسیاری از این قبیل کارها به سادگی فشردن چند دکمه شده است. اما آیا می‌توان به این سیستم‌ها اعتماد کرد و اطلاعات کارت‌های بانکیِ خود را در اختیار آنان قرار داد؟! به سیستم خرید الکترونیکی چطور؟

مساله‌ی فوق شاید در ظاهر بسیار مدرن به نظر برسد، اما امنیت اطلاعات از دیرباز برای جوامع بشری مطرح بوده است که یکی از مهم‌ترین آن‌ها تامین امنیت پیام‌های جنگی در حین ارسال بوده، سوالی که هم‌چنان از دغدغه‌های بشری است!!

هدفی که در این‌جا در پی آن هستیم، یافتن پاسخ برای این سوالات است. پاسخی که بسیار ساده است: رمزنگاری! احتمالاً پیش آمده که با دوستان خود قرادادهایی گذاشته باشید که تنها خودتان قادر به فهم‌شان بوده‌اید. در این صورت شما یک سیستم رمز بین خودتان و دوستتان برقرار کرده‌اید! اما آیا رمزها قابل اعتمادند و هیچ‌راهی برای دست یافتن به آن‌ها نیست؟!

هندسه محاسباتی

بسیاری از اوقات،ما مسائل هندسی رابا کمک توانایی‌های ویژه بینایی درذهن حل می‌کنیم. اما چطور می‌توانیم از الگوریتم‌ها برای حل این مسائل استفاده کنیم؟ به عنوان مثال ما معمولا به راحتی می‌توانیم تعیین کنیم یک نقطه در کدام ناحیه از یک نقشه است یا تقاطعی بین دو شکل وجود دارد یا خیر. اما چطور می‌توانیم این توانایی را برای یک رایانه طراحی کنیم؟ به علاوه بعضی از این مسائل با استفاده از توانایی‌های ذهنی ما نیز به سادگی قابل حل نیستند. نمونه این‌گونه مسائل می‌تواند قرار دادن کمترین تعداد دکل مخابراتی برای آنتن‌دهی یک فضا یا پیدا کردن کوتاه‌ترین مسیر میان دو نقطه باشد. هندسه محاسباتی علم چگونگی حل مسائل هندسی توسط رایانه با بیشترین بازدهی است و محققین و متخصصین در این حوزه به تلاش برای حل مسائل هندسی با تولید ابزارهای محاسباتی برای حل این‌گونه مسائل می‌پردازند. از آنجایی که قابل جابه‌جایی بودن در شبکه‌های تلفن همراه جزء ذات این تکنولوژی است، مسائلی در رابطه با مکان و چگونگی قرار گرفتن در محیط باعث می‌شود هندسه محاسباتی بخش جدا نشدنی از تکنولوژی مذکور باشد. به این ترتیب در این کارگاه سعی می‌کنیم با معرفی مسائلی جالب در این زمینه، مقدمات درک بهتر این شاخه از علم رایانه و کاربردهای آن در شبکه‌های تلفن همراه را فراهم کنیم.

شبکه

چند ده میلیون تلفن همراه در ایران در حال استفاده است و شما تنها با فشردن ۱۱ عدد می‌توانید دوست خود را در بین این چند ده میلیون پیدا کنید! و پیامک شما بدون این‌که در بین آن‌ها گم شود به مقصد درست خود می‌رسد!

اینترنت! که هر روز بیشتر در زندگی ما نفوذ می‌کند! به طرز کار آن فکر کرده‌اید؟ چگونه رایانامه‌های شما در تمام دنیا می‌چرخند و مقصدشان را می‌یابند؟! چگونه این امکان برای شما فراهم می‌شود که با یک تلفن همراه به اینترنت متصل شده و با دوست‌تان که پشت رایانه‌اش قرار گرفته ارتباط برقرار کنید؟!

جهان مدرن جهان ارتباطات است. مبنای زندگی ما، از کارهای روزمره مانند فرستادن پیامک و یا چک کردن رایانامه گرفته تا مدیریت کردن بحران در شهرهای زلزله زده و یا خبررسانی صحیح در شرایط لازم، همه و همه بر شبکه‌های ارتباطاتی استوارند. این که این شبکه‌ها چگونه چنین محیطی را برای ما فراهم می‌آورند، هسته‌ی اصلی این کارگاه را تشکیل می‌دهد. در ادامه نیز سعی بر این است که مشکلات اصلی این شبکه‌ها و راه‌کارهایی که برای این مشکلات مطرح است بررسی شود.

مسابقه

در نهایت در روز آخر مطالبی که به دانش‌آموزان در طی کارگاه‌ها آموزش داده شده بود، به بوته‌ی آزمایش گذاشته شد. این کار به این صورت انجام شد که در روز آخر مسابقه‌ای برگزار گردید که طی آن سعی بر این بوده است علاوه بر سنجش آموخته‌های دانش‌آموزان، با ایجاد فضای رقابتی و دخیل کردن مفاهیم نظریه‌ی بازی‌ها، مدلی از دنیای واقعی به وجود آید تا دانش‌آموزان به تصمیم‌گیری‌های جمعی و همکاری‌های گروهی روی آورند.

مسابقه با شور و هیجان زیادی برگزار شد و گروه اول با رقابت

بسیار نزدیک مشخص شد. در نهایت از برندگان این مسابقه با اهدای جوایزی تقدیر گردید.

میزگرد

در پایان کار و در جریان اختتامیه، با حضور کیهان اصغری و پوریا کاویانی از اعضای "کافه‌بازار" میزگردی پیرامون نرم‌افزارهای موبایل و بازار آن در ایران برگزار شد. در این برنامه دانش‌آموزان به طور مستقیم می‌توانستند سؤالات خود را در این‌باره با افرادی که تجربه‌ی بالایی در این زمینه داشتند مطرح کنند.

بازی گروهی

در این برنامه یک فعالیت جانبی نیز در نظر گرفته شده بود و آن انجام یک بازی گروهی بود که تمامی شرکت‌کنندگان و برگزارکنندگان در سه روز اجرای برنامه در آن شرکت داشتند. در این بازی چندین هدف دنبال می‌شد:

- برقراری تعامل دانش‌آموزان با یکدیگر و مجریان برنامه
- تقویت حس رقابت بین دانش‌آموزان
- ایجاد روحیه‌ی فعالیت‌های گروهی
- آشنایی با شبکه‌های اجتماعی و اهمیت آن‌ها
- مطرح شدن مفاهیمی از علم رمزنگاری و انتقال داده

که در نهایت هیچ‌یک از شرکت‌کنندگان نتوانستند این بازی را با موفقیت به پایان برسانند!

دوره چهارم مسابقه‌ی طرح و حل

علی خزلی

آیا تا کنون تجربه حل مسئله به شکل گروهی داشته‌اید؟ بعید می‌دانم تا کنون در قالب یک مسابقه چنین کاری انجام داده باشید. مسابقه طرح و حل مسابقه‌ای به همین شکل است که تا کنون چهار دوره آن برگزار شده است. تا قبل از دوره چهارم که خودم در این مسابقه شرکت کردم به زیبایی و فواید آن پی نبرده بودم. به همین خاطر این مطلب را می‌نویسم تا شمایی که شاید علاقه‌مند باشید هم با آن آشنا شوید.

هدف اصلی مسابقه طرح و حل، برگزاری مسابقه‌ای است که همه‌ی کارهای آن توسط شرکت کنندگان انجام شود؛ بدون نیاز به یک کمیته‌ی علمی برای طرح سؤال، تصحیح برگه‌ها و ... . ایده‌ی اولیه این بوده است که هر تیم یک یا چند سؤال طرح کند، سؤالات طرح شده توسط تیم‌های دیگر را حل کند و پاسخ‌های تیم‌های دیگر به سؤال خودش را تصحیح کند. اما باید قوانین هوشمندانه‌ای وضع شود تا مسابقه منصفانه باشد؛ مثلاً جلوی این گرفته شود که همه‌ی تیم‌ها سؤالات بسیار سختی بدهند تا هیچ تیم دیگری نتواند آن‌ها را حل کند! برای حل این معضل، برای هر تیم نمره‌ای برای حل سؤالات و نمره‌ای برای طراحی سؤالات داده می‌شود که در نهایت با هم جمع می‌شوند. قبل از پرداختن به قوانین مسابقه، ترجیح می‌دهم تجربه‌ی شرکت در مسابقه را بنویسم تا حوصله‌تان سر نرود.

تا کنون مسابقات طرح و حل توسط انجمن علمی دانشکده‌ی ریاضی و علوم کامپیوتر دانشگاه صنعتی امیرکبیر برگزار شده‌اند که دوره‌ی چهارم آن در روز پنج‌شنبه ۱۸ اردیبهشت ۱۳۹۳ با حضور شش تیم برگزار شد. مسابقه در اصل اینترنتی است و هر تیم یک دانشکده است؛ تمام اعضای دانشکده! اما این دوره، چون دانشگاه صنعتی امیرکبیر حاضر بود چند تیم چند نفره را میزبانی کند، با یک تیم پنج نفره از دانشجویان کارشناسی (به همراه خودم به عنوان سرپرست تیم) در مسابقه شرکت کردیم: کامیار امینی، عرفان خانیکی، حامد دشتی، مینا دلیروی فرد و سهند فرهودی.

پس از انتخاب تیم و چند روز قبل از مسابقه جلسه‌ای تشکیل دادیم تا برای طرح یک سؤال هم‌فکری کنیم. چند ایده‌ی خام مطرح کردیم و با هم فکر کردیم که چه طور می‌شود از روی آن‌ها یک مسئله ساخت. ما اغلب طراحی سؤال را تجربه نمی‌کنیم، به خصوص با زمانی که قرار است با یکدیگر در این باره هم فکری کنیم. به نظر من این می‌تواند یک تمرین برای تحقیق علمی باشد، زیرا کمک می‌کند تا ایده‌هایمان را پرورش دهیم و به تعمیم مطالبی که دیده‌ایم فکر کنیم؛ چیزی که جایش در دوره‌ی کارشناسی خالی است. در پایانِ جلسه ایده‌ی اولیه‌ی یکی از دانشجویان به یک سؤال خوب منجر شد که در انتها می‌بینید.

باید مشخص می‌کردیم که درجه‌ی سادگی سؤال چقدر است. به بیان دیگر، انتظار داریم چند تیم آن را حل کنند. این عدد باید لااقل ۲ و حداکثر دوتا کمتر از تعداد تیم‌ها می‌بود. نمره‌ی طرح سؤال، تابعی از درجه‌ی سادگی و تعداد تیم‌هایی که سؤال را حل می‌کنند است. درجه‌ی سادگی سؤالمان را ۲ تعیین کردیم، اما متأسفانه در نهایت تنها یک تیم آن را حل کرد و قسمتی از نمره‌ی طرح سوال را از دست دادیم.

حدود ساعت ۹ صبح پنج‌شنبه بود که تیم‌ها سؤالاتشان را برای یکدیگر فرستادند. مسابقه کتاب‌باز و در واقع اینترنت‌باز بود! این قسمت هم چند نکته‌ی جالب داشت، از جمله تجربه‌ی فکر کردن گروهی روی سؤالات و تقسیم کار برای نوشتن راه‌حل‌ها. جالب‌ترینش این بود که راه‌حل‌های نوشته‌شده را می‌خواندیم و نحوه‌ی نوشتن هم‌دیگر را اصلاح می‌کردیم. نوشتن ریاضی نیاز به آموزش و تمرین دارد که در این فرصت آن را تجربه کردیم.

پس از پایان وقت راه‌حل‌ها را فرستادیم و تصحیح برگه شروع شد. نمره‌ی هر سؤال صفر یا کامل داده می‌شد و این تصحیح را آسان می‌کرد (این قانون برای جلوگیری از اثر اختلاف سلیقه در نمره‌دهی و سهولت رفع مناقشات وضع شده است). پس مصحح باید تصمیم بگیرد که پاسخ نوشته شده یک راه حل کامل، یا با کمی ایراد جزئی است یا این که یک حل ناقص است. چون تنها پنج برگه را باید تصحیح می‌کردیم، کار سختی نداشتیم و این مرحله به سرعت انجام شد. سپس تقاضاهای تجدید نظر رسیدگی شد. هم می‌توانستیم برای پاسخ‌های خودمان تقاضای تجدید نظر بدهیم هم برای پاسخ‌های دیگران! چون، حداقل روی کاغذ، یک تیم ممکن است برای افزایش نمره‌ی طرحش تصحیح غیرواقعی انجام دهد. پس از آن مراسم اختتامیه برگزار شد و جوایز نقدی که دانشگاه امیرکبیر پیش بینی کرده بود به سه تیم برتر داده شد، که تصادفاً نمره‌ی مساوی داشتند!

با توجه به منافعی که ذکر شد، به نظرم خیلی خوب است که این نوع مسابقه در سطح کوچک‌تر در دانشکده برگزار شود. حتی از این روش به عنوان امتحان میان‌ترم یک درس نیز استفاده شده است[1]. مثلاً می‌توان در یک کلاس در تیم‌های سه یا چهار نفره مسابقه را برگزار کرد.

قوانین مسابقه

بد نیست به قوانین نمره‌دهی هم اشاره کنم، چون وضع هوشمندانه‌ی آن‌ها خودش مسئلهٔ جالبی است! اگر درجه‌ی سادگی یک سؤال k پیش بینی شده باشد، بارم آن سؤال باید تابعی از k باشد که آن را با $b(k)$ نشان می‌دهیم. هم‌چنین اگر n تیم آن را حل کنند، نمره‌ی طرح آن تابعی از n و k است که آن را با $T(k,n)$ نشان می‌دهیم. نمره‌ی حل هر تیم نیز جمع بارم سؤالاتی است که حل کرده است (نمره‌ی کامل یا صفر برای هر سؤال). قوانین مسابقه طرح و حل باید به شکلی باشد که رقابت تیم‌ها برای رسیدن به به‌ترین رتبه، هم‌راستا با برگزاری یک مسابقه مفید و منصفانه باشد. با توجه به این‌که سؤال‌هایی که دو تیم مختلف باید حل کنند در یک مورد با هم تفاوت دارد و سادگی سؤال‌ها نیز قرار است یکی نباشد، طراحی چنین قوانینی توجه به نکاتی را ایجاب می‌کند:

۱. تیم‌ها برای رسیدن به رتبه به‌تر باید هر چه می‌توانند سؤال‌های دیگران را حل کنند.

۲. طراح هر سؤال، اگر تصمیمش را در مورد درجهی سادگی سؤال گرفته، باید به نفعش باشد که سؤالی طرح کند که درجه سادگی‌اش به آن عدد نزدیک‌تر باشد.

۳. طراح هر سؤال باید به نفعش باشد که در مورد درجه سادگی، به‌ترین برآورد خود را بیان کند.

۴. اگر دو تیم از نمره‌ی طرح و نمره‌ی حل بیش‌ترین مقدار ممکن را بگیرند، می‌بایست مستقل از درجه‌ی سادگی سؤالشان نمره‌ی نهایی برابر بگیرند.

نکته جالب این است که همین چهار اصل، فرمول نمره‌دهی را به شکل یک‌تا مشخص می‌کند!

$$b(k)=\frac{c}{k-۱}$$

$$T(k,n)=\begin{cases} nb(k) & n\le k \\ b(k)+\frac{c(n-۱)}{۲n-k-۱} & n>k \end{cases}$$

همان طور که می‌بینید، درجه‌ی سادگی سؤال باید حداقل ۲ باشد. هم‌چنین برای جلوگیری از سادگی بیش از حد می‌توان کران بالایی روی درجه‌ی سادگی سؤالات گذاشت که در آخرین دوره ۴ بود. هم‌چنین توجه کنید که این فرمول‌ها مستقل از تعداد شرکت کنندگان است. به نظر شما آیا با وجود این قوانین که به طور یکتا مشخص شده‌اند، هم‌چنان انتخاب درجه‌های سادگی مختلف برای طرح سؤال تأثیری در رتبه خواهد گذاشت؟

سؤالات دوره‌ی چهارم مسابقه‌ی طرح و حل

دانشگاه تهران

تمام اعداد صحیح مثبت x,m,n را که در رابطه‌ی $x^m+۱|(x+۱)^n$ صدق می‌کنند بیابید.

[1] http://tarhohal.persianblog.ir/page/aazmaayeshi02

نام دانشگاه	۱	۲	۳	۴	۵	۶	k	$b(k)$	n	T(k،n)	نمره‌ی حل	مجموع
دانشگاه تهران		۱	۱	۱	۰	۰	۲	۴۲۰	۲	۸۴۰	۸۴۰	۱۶۸۰
دانشگاه خواجه‌نصیر	۰		۱	۱	۰	۰	۳	۲۱۰	۵	۴۹۰	۶۳۰	۱۱۲۰
دانشگاه شهید باهنر	۰	۱		۰	۰	۰	۲	۴۲۰	۲	۸۴۰	۲۱۰	۱۰۵۰
دانشگاه شهید بهشتی	۰	۱	۰		۰	۰	۳	۲۱۰	۴	۵۲۵	۲۱۰	۷۳۵
دانشگاه صنعتی امیرکبیر	۱	۱	۰	۱		۱	۲	۴۲۰	۱	۴۲۰	۱۲۶۰	۱۶۸۰
دانشگاه صنعتی شریف	۱	۱	۰	۱	۱		۲	۴۲۰	۱	۴۲۰	۱۲۶۰	۱۶۸۰

جدول ۱: نتایج چهارمین دوره‌ی مسابقه‌ی طرح و حل

دانشگاه خواجه نصیرالدین طوسی

فرض کنید (X,d) یک فضای متریک غیر فشرده باشد که فقط یک نقطه‌ی حدی داشته باشد. آیا X لزوماً شمارا است؟ چرا؟

دانشگاه شهید باهنر کرمان

اگر V یک فضای برداری روی میدان اعداد حقیقی باشد، آنگاه به ازای هر $x,y \in V$ تعریف می‌کنیم $[x,y)=\{\lambda y+(1-\lambda)x | 0 \leq \lambda < 1\}$ و به ازای هر $A \subseteq V$ تعریف می‌کنیم

$$A^b=\{x\in A | \forall v \in V \exists \delta > 0 : [x, x+\delta v) \subseteq A$$

$$A^m = A \cup \{x \in V | \exists c \in A : [c,x) \subseteq A\}$$

فرض کنید V یک فضای برداری روی میدان اعداد حقیقی با پایه‌ی شمارای $B=\{x_1,x_2,x_3,\dots\}$ باشد. اگر

$$F=\{x\in V | x=\sum_{i=1}^{k} \lambda_i x_i, k \in N, \lambda_k > 0\}$$

آنگاه نشان دهید $F^b=\emptyset$ و $F^m=V$.

دانشگاه شهید بهشتی

اگر $1 < p \leq q$ دو عدد طبیعی و $f:[0,1]\to[0,1]$ و

$$f(x)=\sum_{n=1}^{\infty}\frac{x_n}{q^n}$$

که $x=\sum_{n=1}^{\infty}\frac{x_n}{p_n}$ بسط نرمال x در مبنای p است، مقدار $\int_0^1 f(x)\mathrm{d}x$ را محاسبه کنید.

دانشگاه صنعتی امیرکبیر

می‌دانیم که به ازای هر عدد طبیعی k، حاصل عبارت $1^k+2^k+\cdots+n^k$ یک چندجمله‌ای درجه‌ی $k+1$ بر حسب n است. این چندجمله‌ای را با $S_k(n)$ نمایش می‌دهیم. به عنوان مثال $S_1=\frac{1}{2}x(x+1)$ و $S_2=\frac{1}{6}x(x+1)(2x+1)$. تعریف کنید $R_k(x)=S_k(x-\frac{1}{2})$. نشان دهید اگر k زوج باشد، R_k فرد است و اگر k فرد باشد R_k زوج است.

دانشگاه صنعتی شریف

فرض کنید $n \geq 4$ و A ماتریسی با n سطر و $2n+1$ ستون باشد که درایه‌های آن در یک میدان دو عضوی هستند. ثابت کنید $n+1$ ستون A پیدا می‌شود که رتبه‌ی ماتریسی که از کنار هم گذاشتن آن‌ها به دست می‌آید کمتر از n باشد. [2]

وبلاگ مسابقه: http://tarhohal.persianblog.ir

[2] با راه حل دانشگاه صنعتی امیرکبیر برای این سؤال، می‌توان به جای $2n+1$ قرار داد $n+5$. این نتیجه را می‌توان برای $n \geq 5$ تا $n+3$ بهبود داد.

مجله شفاهی کارشناسی

مجله‌ی شفاهی کارشناسی ارائه‌ای دانشجویی است که در آخرین چهارشنبه‌ی هر ماه یک ساعت قبل از مجله‌ی شفاهی یعنی به طور معمول ساعت ۱۳ آغاز می‌شود. ارائه‌ها تقریباً یک ساعت طول می‌کشد که در آن یکی از دانشجویان مطلبی قابل فهم برای دانشجویان کارشناسی را ارائه می‌دهد.

تقریباً یک ماه قبل از ارائه فرد ارائه دهنده اعلام آمادگی می‌کند و پس از مشخص شدن موضوع و آمادگی کامل با تعدادی از اساتید جلسه‌ای گذاشته می‌شود که در آن نکات علمی مربوط و حتی نحوه‌ی ارائه دادن را به فرد تذکر می‌دهند و با این کار فرد به جمع‌بندی و آمادگی کامل می‌رسد.

تبلیغات مجله تقریباً یک هفته قبل در سراسر دانشکده نصب می‌گردد. در هر شماره از مجله شفاهی نیز از تعدادی از اساتید برای شرکت در جلسه دعوت می‌شود. همچنین از جلسه فیلم‌برداری می‌شود تا مورد استفاده عموم قرار گیرد. از جمله اهداف برگزاری مجله شفاهی کارشناسی:

- **پر کردن اوقات خالی دانشجویان ریاضی با فعالیت ریاضی:** بسیاری از دانشجویان هستند که اوقات خالی خود را با فعالیتی غیر ریاضی پر می‌کنند و با ایجاد چنین بستری زمینه برای بارور کردن استعداد دانشجویان و جلوگیری از اتلاف وقت آنها فراهم می‌شود.

- **ایجاد توانایی در ارائه دادن دانشجویان کارشناسی:** ایجاد مجله شفاهی برای دانشجویان لیسانس باعث تقویت مهارت آنان در ارائه دادن مطالب عمیق‌ترِ علمی می‌شود.

- **مهیا ساختن بستری مناسب برای فراگیری مطالب علاوه بر واحدهای درسی:** با ایجاد چنین مجله‌ای انگیزه‌ای مضاعف برای دانشجویان لیسانس برای فراگیری مطالب عمیق‌تر و زیباتر فراهم می‌شود و توانایی علمی آنان را ارتقا می‌دهد.

سه شماره از مجله شفاهی کارشناسی را گذرانده‌ایم. موضوع شماره‌ی اول تقابل درجه دوم توسط کامیار امینی، شماره‌ی دوم از ترمودینامیک آماری تا نظریه اعداد توسط روح ا... مهکام و شماره‌ی سوم الگوریتم چند جمله‌ای برای بررسی اول بودن یک عدد توسط محمد قیاسی بود.

به امید برگزاری بهتر شمارهای بعدی ان شا ا...

برنامه‌ی پخش ویدئوی هفتگی

محمد قیاسی

برنامه‌ی پخش ویدئوی هفتگی دانشکده ریاضی از ابتدای سال جدید (سال ۱۳۹۳) در دانشکده آغاز شده است و قصد داریم که در طول سال تحصیلی هر هفته یک ویدئوی ریاضی (مرتبط با ریاضیات) در دانشکده پخش کنیم. ویدئوهایی که برای پخش در هر هفته انتخاب می شوند ممکن است از نوع یک سخنرانی ریاضی (lecture) یا یک فیلم (movie) مرتبط با ریاضیات یا ... باشند.

در زیر سایت‌هایی را لیست کرده‌ایم که ویدئوهایی با موضوع ریاضی برای دریافت قرار داده اند و شما نیز می‌توانید با مراجعه به آنها ویدئوهای موردعلاقه‌تان را دریافت کنید یا اینکه لینک آن را به ما هم پیشنهاد کنید تا از آنها در این برنامه استفاده کنیم.

https://video.ias.edu

http://www.msri.org/web/msri/online-videos

http://www.claymath.org/publications/video-catalogue

https://www.edx.org

http://video.impa.br

http://www.fields.utoronto.ca/video-archive

همچنین لیست ویدئوهایی که تاکنون پخش شده‌اند به شرح زیر است که می‌توانید آنها را از طریق آدرس https://videos.math.sharif.ir/weekly دریافت کنید.

۱) موسیقی اعداد اول، مارکوس دو سوتوی

۲) Dynamics on Moduli Space of Curves، مریم میرزاخانی

۳) Structure and Randomness in Prime Numbers، ترنس تآ او

۴) مستند قضیه‌ی آخر فرما

۵) The Importance of Mathematics، تیموتی گاورز

۶) The Mystry of Three-Manifolds، ویلیام ترستن

در پایان تشکر می‌کنیم از دکتر جعفری که در انتخاب ویدئوها و برگزاری برنامه ما را یاری کردند و همچنین مطلوب ماست که پیشنهادات و انتقادات خودتان را از طریق ارسال ایمیل به mathvidoesite@yahoo.com به اطلاع ما برسانید.

انجمن علمی و فوق برنامه‌ی دانشکده‌ی علوم ریاضی
بهار ۹۳

ساعت چای دانشکده ریاضی

از آغاز به کار شورای صنفی در دانشکده یکی از اهداف اصلی آن، ایجاد فضایی دوستانه بین دانشجویان و اساتید بوده است. از این رو با مشورت دانشجویان و اساتید محترم دانشکده، تصمیم به برگزاری سلسله نشست هایی با عنوان "ساعت چای" گرفته شد. با مشورت اعضای شورا و بررسی زمان های مناسب برای اجرای برنامه، تصمیم گرفتیم به مناسبت روز ملی ریاضیات، اولین جلسه از این سری جلسات را در تاریخ ۲۹ اردیبهشت برگزار کنیم. خوشبختانه این جلسه با استقبال پرشور تعداد کثیری از دانشجویان و اساتید برگزار شد. هر چند جای خالی چند نفر از اساتید محترم در جمع احساس می‌شد. برنامه ساعت ۱۴ در محوطه باز طبقه اول دانشکده آغاز شد و به مدت حدودی دو ساعت به طول انجامید. در یک ساعت نخست از دانشجویان و اساتید پذیرایی صورت گرفت و افراد به گپ و گفت پرداختند، در ادامه بحثی میان چند نفر از دانشجویان و دکتر علیشاهی شکل گرفت که منجر به تشکیل یک حلقه مباحثه شد و کم‌کم دانشجویان و چند تن دیگر از اساتید نیز به این حلقه پیوستند. در پایان ادامه بحث به جلسات بعدی موکول شد و جلسه در ساعت ۴ به پایان رسید. همانند جلسه‌ی نخست، زمان جلسه‌ی بعدی نیز در دانشکده اعلام می‌گردد.

[illegible]

مجله‌ی ریاضی شریف از هر گونه همکاری در تمامی زمینه‌ها از جمله تهیه یا معرفی مطالب علمی و توصیفی از جانب دانشجویان و اساتید استقبال به عمل می‌آورد. لازم به ذکر است که اکثر همکاران فعلی این مجله به‌صورت کاملا داوطلبانه با مجله همکاری می‌کنند و اساس کار این نشریه بر مبنای همکاری داوطلبانه‌ی اهالی دانشکده‌ی ریاضی قرار گرفته‌است.

تماس با ما:

mathematicsjournal@gmail.com
www.sharifmathjournal.ir

قیمت: ۲۰۰۰ تومان