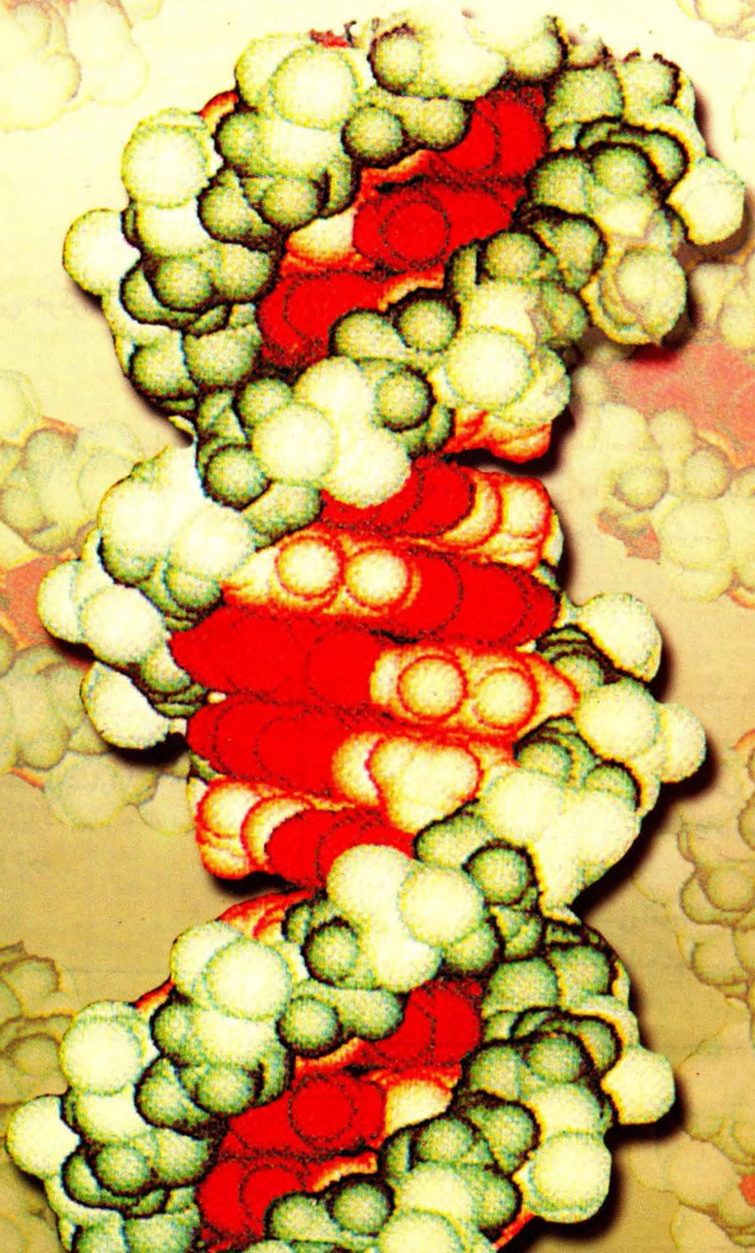
# شیمی

سال ۱۲، شمارهٔ ۱ و ۲

شمارهٔ پیاپی: ۲۳



## بسم‌الله الرحمن الرحیم

## مجلهٔ ریاضی مرکز نشر دانشگاهی

**شمارهٔ ۸۵ خیابان پارک خیابان دکتر بهشتی، تهران ۱۵۱۳۴**





بهای این شماره ۵۵۰۰ ریال؛ حق اشتراک سالانه برای داخل کشور ۱۱۰۰۰ ریال. (برای دانشجویان با ٪۳۰ تخفیف)

وجه اشتراک به حساب شمارهٔ ۹۰۰۰۹ بانک ملی شعبهٔ خیابان پارک تهران به نام مرکز نشر دانشگاهی واریز شود.


***نشر ریاضی*** از انتشارات ادواری مرکز نشر دانشگاهی است که هر شش ماه یک بار منتشر می‌شود. هدفهای اصلی انتشار مجله عبارت است از

- معرّفی پیشرفتهای جدید ریاضیات؛
- معرّفی شاخه‌های جدید علوم ریاضی و همچنین مباحثی که مورد توجّه پژوهشگران است؛
- معرّفی جنبه‌های کاربردی، فرهنگی، فلسفی، و تاریخی ریاضیات؛
- معرّفی کارهای ریاضی پژوهان فارسی زبان و ایجاد ارتباط بین آنان؛
- طرح مسائل آموزشی ریاضیات، به ویژه مسائل مربوط به آموزش ریاضیات دانشگاهی در ایران.

***نشر ریاضی*** از همکاری تمام علاقه‌مندان استقبال می‌کند. مقاله‌های ارسالی باید در چهارچوب هدفهای فوق و با سبکی مشابه با سبک مقاله‌های چاپ‌شده در ***نشر ریاضی*** باشد. به همکارانی که مایل‌اند مقاله‌ای را به فارسی برگردانند و برای درج به مجله بفرستند توصیه می‌شود ابتدا اصل مقاله را با ذکر منبع برای بررسی و تصویب ارسال دارند. ترجمهٔ آزاد پذیرفته نمی‌شود، و فرستادن اصل مقاله‌های ترجمه شده الزامی است. مقاله‌های ارسالی پس فرستاده نمی‌شود. هر مقاله‌ای مطابق ضوابط رایج داوری می‌شود. هیأت ویراستاران در رد، قبول، و حک و اصلاح مقالات آزاد است. ویرایش، نگارش، انتخاب واژه‌ها، و ضبط اسامی و اعلام مطابق ضوابط گروه ریاضی در مرکز نشر دانشگاهی انجام خواهد گرفت.

### یادآوری

- متن مقاله روی یک طرف کاغذ، یک خط در میان، و با حاشیهٔ کافی ماشین، یا با خط خوانا نوشته شود.
- نحوهٔ نگارش، بخشبندی، فرمول‌نویسی، و شیوهٔ ارجاع به منابع حتی‌المقدور مطابق با مقاله‌های چاپ شدهٔ در ***نشر ریاضی*** باشد.
- فهرست معادلهای انگلیسی اصطلاحاتی که در مقاله به‌کار می‌رود همراه با مقاله فرستاده شود.

# نشر ریاضی

**سال ۱۲، شمارهٔ ۱ و ۲**

تاریخ انتشار: شهریور ۱۳۸۰

شمارهٔ پیاپی: ۲۳

nashriaz@rose.ipm.ac.ir

**صاحب امتیاز:** مرکز نشر دانشگاهی

**مدیر مسؤول:** سیاوش شهشهانی





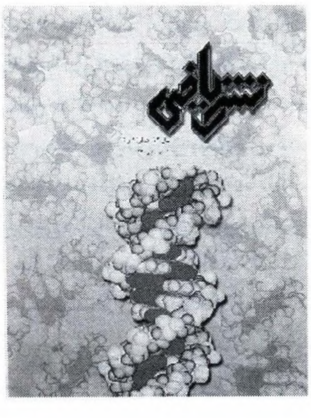
- **دستیارفنی:** زهرا دلاوری
- **طراح:** شکوه پیله‌فروشها
- **حروفچین و صفحه‌آرا:** سیده مریم طاهریان
- **ناظر چاپ:** علی صادقی

- **لیتوگرافی:** مردمک
- **چاپ و صحافی:** منفرد (میدان شهدا، خ شهید کفایی، شمارهٔ ۳۶)




## فهرست





**روی جلد**

به مناسبت درج مقالهٔ «کارآیی نامعقول ریاضیات در زیست‌شناسی مولکولی»

# گزارش

در ربع قرن اخیر شاهد تأثیرات متقابل و ارتباط کم‌سابقه‌ای میان بعضی از پیشرفته‌ترین بخشهای توپولوژی و هندسه از یک سو و پیشرفته‌ترین مباحث فیزیک میدانها و ذرات بنیادی از سوی دیگر بوده‌ایم. اولین اثر چشمگیر و درخشان این ارتباط در ریاضیات، موفقیت دانلدسن در بررسی خمینه‌های چهاربعدی بود که اعطای مدال فیلدز را به او در سال ۱۹۸۶ به دنبال داشت. در سال ۱۹۹۴، ادوارد ویتن که بعضیها او را علمدار فیزیک نظری انرژیهای بالا می‌دانند نیز به پاس تحقیقات میان‌رشته‌ای‌اش مدال فیلدز را دریافت کرد. مقالهٔ تألیفی این شماره، «نظریهٔ زایبرگ‌ـویتن چیست؟»، به نظریهٔ پربار دیگری در این زمینه می‌پردازد که در سال ۱۹۹۴ مطرح شد و تاکنون منشأ کشفیات هندسی قابل‌توجهی بوده است.

در بخش دیگری از علوم تجربی، زیست‌شناسی مولکولی و ژنتیک شاید پرتحرک‌ترین رشته‌های ربع قرن اخیر بوده باشند. مقالهٔ «کارایی نامعقول ریاضیات در زیست‌شناسی مولکولی» که ترجمهٔ آن در این شماره می‌آید، داستانی شنیدنی دارد. این مقاله مبتنی است بر سخنرانی نویسنده در جلسهٔ پایانی سمپوزیومی با عنوان «عملکرد زیست‌ـمولکولی و تکامل در چارچوب پروژهٔ ژنوم» که در دسامبر سال ۱۹۹۸ در مؤسسهٔ علوم ریاضی آیزک نیوتنِ کیمبریج برگزار شد. نویسندهٔ مقاله در نامه‌ای که پس از نشر مقاله در مجلهٔ متمتیکال اینتلیجنسر به آن مجله فرستاد، می‌نویسد که نخست به یکی از مسؤولان سمپوزیوم که زیست‌شناس بود، پیشنهاد کرده بود سخنرانیی با عنوان «ناکارایی نامعقول ریاضیات در زیست‌شناسی مولکولی» ایراد کند. آن مسؤول اعتراض می‌کند که چنین عنوانی اثر منفی بر حال‌وهوای سمینار خواهد داشت که هدف آن نزدیک کردن ریاضیدانان و زیست‌شناسان به یکدیگر است، و پیشنهاد می‌کند در عنوان سخنرانی، «کارایی» را به جای «ناکارایی» بگذارد؛ و عنوان مقالهٔ حاضر از آن پیشنهاد ناشی شده است. ولی ویراستاران مجلهٔ اینتلیجنسر همراه با چاپ مقاله، از قول گلفاند ریاضیدان معروف که تحقیقات گسترده‌ای در فیزیولوژی نیز دارد ولی ریاضیات و زیست‌شناسی را با هم مخلوط نمی‌کند، به اصل «ناکارایی نامعقول ریاضیات در علوم زیستی» اشاره کردند. نویسنده در نامهٔ فوق‌الذکر می‌نویسد که اگر از وجههٔ علمی گلفاند برخوردار بود احتمالاً بر پیشنهاد اولیهٔ خودش در مورد عنوان سخنرانی (و مقاله) پافشاری کرده بود. وی می‌افزاید: «ریاضیات بی‌شک، برای انسجام‌بخشیدن به مشاهدات، در زیست‌شناسی کارایی دارد ولی زیست‌شناسی فاقد آن ایجاز چشمگیر علوم فیزیکی است که در آنها تعدادی معدود اصل بنیادی، پیش‌بینی کمّی مشاهدات را با دقت بسیار ممکن می‌سازند. زیست‌شناسی که با تودهٔ بزرگی از مشاهدات تبیین نشده روبه‌روست، مشکل بتواند ایمان داشته باشد که کشف ساختار ریاضی مناسب، با برملاساختن نظام نهفته در مشاهدات، همهٔ معضلات را حل خواهد کرد. مسأله این است که نقش تصادف تاریخی در اینجا بسیار پراهمیت است. فیزیکدان معروفی یک بار می‌خواست کار مرا با این جمله که 'کار تو علم نیست، باستان‌شناسی است' بی‌اهمیت جلوه دهد. احساس من این است که نظر او نسبت به من (و قطعاً نسبت به باستان‌شناسان) بی‌انصافانه بود، ولی در هر صورت به مانع جدی و اصیلی که در برابر کاربرد ریاضی در زیست‌شناسی وجود دارد، اشاره داشت.» نویسنده نامه‌اش را این‌گونه به پایان می‌رساند که می‌گوید سؤالِ «آیا ریاضیات در زیست‌شناسی کارایی دارد؟» او را به یاد این گفتگو می‌اندازد که از کسی حال همسرش را پرسیدند و او جواب داد: «در مقایسه با چه؟»

بیشتر مقاله‌های این شماره را مطالب میان‌رشته‌ای تشکیل می‌دهند. جوزف دوب از بزرگترین متخصصان احتمال ریاضی در قرن بیستم، داستان پیدایش دقت ریاضی در علم احتمال را بازگو می‌کند. مقالهٔ کلاسیک این شماره نوشتهٔ توصیفی کلاسیکی از سول کریپکی است که بعضیها آثار او را سرآغاز مطالعات نوین در منطق موجهات می‌دانند.

در سالهای اخیر، کتابهای ریاضی مفید متعددی به زبان فارسی منتشر شده است که تعدادی از آنها شایستهٔ نقد و معرفی‌اند. در این شماره به بررسی دو تا از آنها که ترجمهٔ آثار شایان توجه خارجی‌اند می‌پردازیم و این کار را در آینده نیز ادامه خواهیم داد.

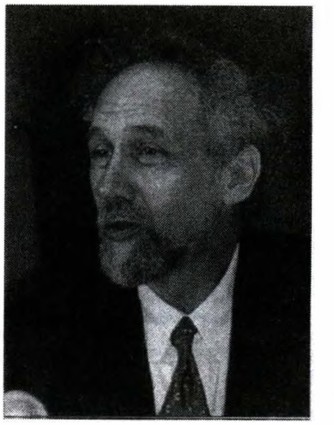
در دو شمارهٔ پیشین مجله، بخش مسأله را، به ترتیب، آقایان ایمان افتخاری و کیوان ملاحی، از اعضای تیمهای موفق المپیاد ریاضی کشور در سالهای گذشته، نوشته بودند. این روند در این شماره نیز ادامه دارد و خانم میرزاخانی از المپیادیهای سابق نوشتن این بخش را به عهده داشته است. امیدواریم تهیهٔ این بخش به قلم دوستان المپیادی به‌صورت یک سنت ادامه یابد.

******

## جایزهٔ کرافورد

جایزهٔ سال ۲۰۰۱ کرافورد را آکادمی سلطنتی سوئد در مهر ماه سال جاری به ریاضیدان فرانسوی **الن کُن** اعطا خواهد کرد. الن کُن[1] که از بزرگترین ریاضیدانان معاصر محسوب می‌شود در سال ۱۹۸۳ به مناسبت آثار مهمش در جبر عملگرها برندهٔ مدال فیلدز شد. پس از آن کُن نقشی اساسی در ایجاد و پیشبرد هندسهٔ ناجابه‌جایی داشت و آثار دیگری که در

1. Alain Connes



الن کُن

نظریهٔ بازبهنجارش و مدلهای استاندارد فیزیک ذرات داشته است به‌زعم بعضی ریاضیدانان راه را برای درک فرضیهٔ ریمان گشوده است. او متولد سال ۱۹۴۷ و استاد مؤسسهٔ مطالعات عالی علمی (IHES) وکُلِژ دوفرانس در فرانسه است.

جایزهٔ کرافورد را که شامل یک مدال طلا و پانصد هزار دلار آمریکاست بنیادی به این نام از سال ۱۹۸۰ به نوبت در پنج رشتهٔ ریاضیات، اخترشناسی، علوم‌زیستی (به‌خصوص محیط زیست)، علوم‌زمین، و مطالعات آرتروز_رماتیسمی اعطا می‌کند. برندگان قبلی این جایزه در ریاضیات عبارت‌اند از ولادیمیر آرنولد، لوئیس نیرنبرگ، پیردُلین، آلکساندر گروتندیک (که جایزه را رد کرد)، سایمن دانلدسن، و شینگ_تونگ‌یاو. (خلاصه شده از متن رسمی آکادمی سلطنتی سوئد در این زمینه.)

## جایزهٔ وُلف

جایزهٔ ۱۰۰۰۰۰ دلاری وُلف در سال ۲۰۰۱ به **ولادیمیر آرنولد و ساهارون شلاه** تعلق گرفت.

ولادیمیر آرنولد در رشته‌های گوناگون ریاضی از جمله دستگاههای دینامیکی، نظریهٔ تکینه‌ها، هندسهٔ همتافته و بخشهای گوناگون فیزیک ریاضی آثار مهمی دارد. رسالهٔ دکتری او شامل حل مسألهٔ سیزده هیلبرت بود. نظریهٔ KAM (کولموگوروف_آرنولد_موزر) که اثبات حالت تحلیلی آن از اوست، به‌صورت یکی از ابزارهای مهم معادلات دیفرانسیل، دستگاههای دینامیکی، و بعضی از بخشهای فیزیک ریاضی درآمده است. او یکی از پایه‌گذاران توپولوژی همتافته است.

ساهارون شلاه[1] یکی از برجسته‌ترین منطقدانان ریاضی محسوب می‌شود. او که بیش از ۷۰۰ مقالهٔ تحقیقاتی نوشته است، ایجادکنندهٔ چند جریان جدید در منطق ریاضی و نظریهٔ مجموعه‌هاست. در نظریهٔ مدلها، کار او با عنوان «نظریهٔ پایداری» به جریان غالب در این نظریه تبدیل شده است. آثار او منجر به حل مسائلی دیرینه در کاربرد نظریهٔ مدلها و نظریهٔ مجموعه‌ها در شاخه‌های گوناگون ریاضی مانند نظریهٔ گروهها، توپولوژی، نظریهٔ اندازه، فضاهای باناخ، و ترکیبیات شده است.

1. Saharon Shelah

## جایزهٔ وبلن

جایزهٔ وبلن در هندسه و توپولوژی هر پنج سال یک بار توسط انجمن ریاضی آمریکا اعطا می‌شود. این جایزه در سال ۲۰۰۱ به جف چیگر[1]، یاکوف الیاشبرگ[2]، و مایکل هاپکینز[3] تعلق گرفت. جف چیگر آثار متعددی در هندسهٔ ریمانی دارد، کار الیاشبرگ در هندسهٔ همتافته نقش اساسی در پیشرفتهای عمدهٔ این رشته داشته است، وکار هاپکینز عمدتاً در نظریهٔ هوموتوپی در توپولوژی جبری است.

س. ش

## درگذشت کواین

هفته‌ای پیش از پایانِ قرنِ بیستم، روزِ بیست و پنجمِ دسامبرِ ۲۰۰۰ (چهارمِ دیِ ۱۳۷۹) ویلارد کواین[4] درگذشت، و او از مهمترین و پرنفوذترین فیلسوفانِ قرن بود. حتی در میانِ غیرِ اهلِ فلسفه هم آراء او بر ضدِ رده‌بندیِ تحلیلی/ترکیبیِ احکام و نظرِ او در موردِ عدمِ تعیّنِ ترجمه و (احتمالاً) نقدهای او بر مفاهیمِ وجهیِ ضرورت و امکان معروف است. او فلسفه را بخشی از علم (به مفهوم تجربیِ آن) می‌دانست. از کواین ۲۲ کتاب منتشر شده که مهمترینشان لفظ و شیء (۱۹۶۰) است. او در ۱۹۸۰ در اکرون در ایالتِ اوهایوی امریکا متولد شد، در ۱۹۳۲ تحتِ نظرِ الفرد نورث وایتهد دکتریِ فلسفه گرفت، در ۱۹۷۸ از دانشگاه هاروارِدِ امریکا بازنشسته شد، و تا اواخرِ عمر ذهنِ فعالِ نقّادی داشت.

کواین در ابتدا به واسطهٔ آثارش در منطقِ ریاضی مشهور شد، گرچه مسلماً او در شمارِ منطقدانانِ بسیار برجستهٔ قرنِ گذشته نیست؛ در هر صورت:

معروفترین اثرِ کواین در منطقِ ریاضی مقالهٔ «مبانیِ جدید برای منطقِ ریاضی»[5] (۱۹۳۷) است که در آن یک نظامِ نگرهٔ مجموعه‌ها («NF»، برگرفته از عنوانِ همین مقاله) پیش نهاده شده است. می‌دانیم که در پیدایشِ پارادوکسِ راسل «تقصیر» با اصلِ تصریح است: به ازای هر فرمولِ $\varphi$، $\{x : \varphi(x)\}$ یک مجموعه است. در نظامِ تسرملو_فرنکل (ZF) این اصل با ملاحظاتی در موردِ بزرگیِ مجموعه‌ها مقیّد شده است: به ازای هر مجموعهٔ $A$ و هر فرمولِ $\varphi$، $\{x \in A : \varphi(x)\}$ یک مجموعه است. در NF اصلِ تصریح با توجه به خصوصیاتِ $\varphi$ تحدید شده است: به‌ازای هر $\varphi$ی لایه‌لایه[6]، $\{x : \varphi(x)\}$ یک مجموعه است ــ «لایه‌لایه»بودنِ فرمولها از طریقِ نگرهٔ نوعها تعریف می‌شود؛ مثلاً "$x = x$" لایه‌لایه هست و (نقیضِ) "$x \in x$" لایه‌لایه نیست. به خلافِ ZF، در NF مجموعهٔ جهانی وجود دارد و نقیضِ اصل موضوعِ انتخاب هم اثبات‌پذیر است. از جمله به دلیلِ این مطلبِ اخیر است که NF را مبنای مناسبی برای ریاضیات نمی‌دانند. هنوز دانسته نیست (حتی با فرضِ سازگاریِ ZF) که NF سازگار هست یا نه.

ک. ل.

1. Jeff Cheeger 2. Yacov Eliashberg 3. Micheal Hopkins
4. Willard Van Orman Quine
5. "New Foundations for Mathematical Logic" 6. stratified

نشر *ریاضی*، سال ۱۲، شمارهٔ ۱ و ۲، صص ۴ـ۱۱

# نظریهٔ زایبرگ‌ـویتن چیست؟

پدرام صفری*

**اشاره.** این مقاله، گزارش‌گونه‌ای است بسیار مختصر از تحولات چند سال اخیر در نظریهٔ خمینه‌های چهاربعدی که به دنبال کشف معادلاتی توسط زایبرگ و ویتن صورت یافت. برای ملاحظهٔ شرح مفصلتر، خواننده می‌تواند به منابع انتهای مقاله رجوع کند. جنبهٔ هندسی این نتایج است که منظور نظر این نوشتار خواهد بود، گرچه فیزیکدانان و ریاضی‌ـفیزیکدانان هم دیدگاههای خاص خود را در بارهٔ این مبحث دارند که نگارنده آن را به اهلش وامی‌نهد.

## نگاه کلی

در پاییز سال ۱۹۹۴ میلادی، ادوارد ویتن[1] و ناتان زایبرگ[2]، ضمن کار بر روی نظریه‌های فرامتقارن پیمانه‌ای[3]، به معادلات دیفرانسیلی پی بردند که طبق روال فیزیکدانان باید همان نتایج نظریهٔ دانلدسن[4] را در مورد خمینه‌های چهاربعدی به‌دست می‌داد [W]. در عین حال این معادلات مزیت ویژه‌ای داشت: گروه پیمانه‌ای در این معادلات $\mathrm{U}(1)$ است، که گروهی است آبلی، و این امر محاسبات را بسیار ساده می‌کند. در ضمن، برخلاف نظریهٔ دانلدسن، فضای پیمانهٔ[5] وابسته به معادلات زایبرگ‌ـویتن فضایی است فشرده، و همین باعث می‌شود که از برخی مشکلات فنی اجتناب شود.

در فاصلهٔ چند ماه پس از کشف این نظریه، علاوه بر ارائه اثباتهای کوتاهتر و ساده‌تر از نتایج نظریهٔ دانلدسن، پیشرفتهایی نیز در نظریهٔ خمینه‌های چهاربعدی حاصل شد:

**الف) اثبات حدس توم[6] توسط کرونهایمر[7] و مُروکا[8]**

فرض کنید $c$ یک ردهٔ هومولوژی دوبعدی در $\mathbb{CP}^2$ باشد. رویه‌های دوبعدی متفاوتی در $\mathbb{CP}^2$ می‌توانند نمایندهٔ این ردهٔ هومولوژی باشند ــ این امر نتیجه‌ای از نظریهٔ کوبوردیسم توم است. توجه کنید که اگر $\Sigma$ چنین نماینده‌ای باشد، می‌توان با چسباندن دستوارهٔ[1] مناسب کوچکی، نمایندهٔ دیگری به‌دست آورد که گونهٔ[2] بزرگتری دارد. بنابراین برای گونهٔ این نماینده‌ها کران بالایی نیست. اما می‌توان از کران پایین گونه‌ها پرسید. حدس توم این بود که این کران پایین توسط یک خم جبری که ردهٔ $c$ را نمایش می‌دهد محقّق می‌شود (این خم جبری با تقریب جهت همواره وجود دارد). حدس توم بعداً صورت تعمیم‌یافته‌ای هم یافت. مورگان[3] این پرسش را مطرح کرد که آیا این حدس در حالتی که $\mathbb{CP}^2$ با یک رویهٔ $M$ با بعد مختلط ۲ جایگزین شود هم برقرار است یا نه. کرونهایمر و مروکا با استفاده از روشهای پیمانه‌ای دانلدسن، تعمیم حدس توم را در حالتی که $b_2^+(M) > 1$ ثابت کردند، ولی روش آنان در حالتی که $b_2^+ = 1$ (و به‌ویژه حالت $M = \mathbb{CP}^2$، که حدس اصلی توم بود) به اشکالات عمده‌ای برمی‌خورد [KM1–KM3]. نظریهٔ بدیع زایبرگ و ویتن راه را برای اثبات این حالت باز کرد: کرونهایمر و مروکا مسأله را در حالت $\mathbb{CP}^2$ حل کردند [KM] و از سوی دیگر، مورگان، سابو[4]، و تاوبز[5] با روشهای مشابهی و با به‌دست آوردن یک فرمول ضربی[6]، تعمیم حدس توم را در حالتی که ردهٔ هومولوژی $c$ عدد خودْتقاطعی[7] نامنفی داشته باشد به اثبات رساندند [MST]. نهایتاً سابو کلیترین صورت حدس توم را در حالتی که خمینه همتافته[8] باشد، و بدون هیچ فرضی روی عدد خودتقاطعی، ثابت کرد [OSz].

**ب) نتایج تاوبز در هندسهٔ همتافته: $\mathrm{SW} = \mathrm{Gr}$**

در عین حال، تاوبز معادل‌بودن ناورداهای زایبرگ‌ـویتن و ناورداهای گروموف[9] برای خمینه‌های چهاربعدی همتافته را ثابت کرد [T1–T8]. ناورداهای گروموف قبلاً براساس تعداد خمهای شبه‌هولومورف[10] در یک خمینهٔ همتافته تعریف شده بودند [G]. علاوه بر این، تاوبز مقدار ناوردا را برای خمینه‌های چهاربعدی همتافته با $b_2^+ > 1$ صریحاً محاسبه کرد. صورت این نتیجه را بعداً بیان خواهیم کرد.

1. Edward Witten 2. Nathan Seiberg
3. super-symmetric gauge theories 4. Simon Donaldson
5. moduli space 6. René Thom 7. Peter Kronheimer
8. Tomasz Mrowka

1. handle 2. genus 3. John W. Morgan
4. Zoltán Szabó 5. Clifford H. Taubes 6. product formula
7. self-intersection 8. symplectic 9. Michael Gromov
10. pseudo-holomorphic

### ج) پیامدها در هندسهٔ جبری

علاوه بر اثباتهای کوتاهتر از نتایج قبلی نظریهٔ دانلدسن در هندسهٔ جبری، قضایای جدیدی هم به اثبات رسید، ازجمله ناوردایی کثیرگونه[1] تحت نگاشتهای هموار [FM1]، که درواقع تعمیمی از حدس وان دِون[2] است که قبلاً به کمک چندجمله‌ای‌های دانلدسن به اثبات رسیده بود، یعنی اینکه بُعدِ کُدایرا[3] ناوردایی هموار است [FQ]. به دلیل وجود ساختار مختلط، جوابهای معادلات زایبرگ‌ـ‌ویتن در بسیاری از حالات (مثلاً برای رویه‌های کِی‌لِر[4]) به‌طور صریح به‌دست می‌آیند [M]، [W].

### د) بازنگری در نظریهٔ دانلدسن

با استفاده از نظریهٔ زایبرگ‌ـ‌ویتن، اثبات کوتاهتر و ساده‌تری از قابل‌توجه‌ترین قضیهٔ دانلدسن به‌دست آمد.

• قضیهٔ دانلدسن، فرمهای تقاطع منفی معین یک خمینهٔ چهاربعدی هموارِ بستهٔ ساده‌ـ‌همبند، قطری‌شدنی هستند.

برای توضیحات بیشتر در بارهٔ فرمهای تقاطع و این قضیه، رجوع کنید به [ک]. اثباتی از قضیهٔ فوق با استفاده از نظریهٔ زایبرگ‌ـ‌ویتن در [N] آمده است.

• حدس $\frac{11}{8}$. در یک خمینهٔ چهاربعدی هموارِ بستهٔ ساده‌ـ‌همبند، رتبهٔ[5] فرم تقاطع حداقل $\frac{11}{8}$ نشان[6] آن است.

برای رویه‌های $K3$، این حداقل محقق می‌شود، بنابراین تخمین فوق قابل بهبود نیست. اثبات $\frac{10}{8}$ ادعا شده است [F].

### ه) نتایج داخلی و مسائل منبعث از نظریهٔ زایبرگ‌ـ‌ویتن

در اینجا به منظور تکمیل فهرست، تعدادی از قضایا و حدسهایی را که در خود نظریهٔ زایبرگ‌ـ‌ویتن جالب‌توجه‌اند ذکر می‌کنیم. در برخی موارد، حتی بیان صورت حکم مستلزم دانستن نظریه است، که خواننده را برای ملاحظهٔ توضیحات بیشتر به ادامهٔ مقاله ارجاع می‌دهیم.

• قضایای چسباندن[7] و فرمولهای ضربی. صورت عمومی مسأله چنین است. فرض کنید خمینهٔ چهاربعدی هموار و بستهٔ $M$ توسط زیرخمینهٔ سه‌بعدی $M_0$ به دو قسمت $M^+$ و $M^-$ تقسیم شده است. مراد، کسب اطلاعات در بارهٔ ناورداهای زایبرگ‌ـ‌ویتنِ $M$، یا حتی توصیف جوابها و فضای پیمانهٔ زایبرگ‌ـ‌ویتنِ $M$ است، در صورتی‌که اطلاعات مشابه در مورد $M^\pm$ و $M_0$ داده شده باشد. مسألهٔ در حالتهای خاصی که $M_0$ به‌صورت $S^1 \times \Sigma$ باشد، که $\Sigma$ یک رویهٔ ریمانی است، یا به‌صورت یک فضای تاربندی شده زایفرت[8] باشد، مثلاً بر اثر جراحی روی یک گره[9] یا زنجیر[10] به‌دست آمده باشد، یا در حالت کلی در نقاط تحویل‌ناپذیر فضای پیمانهٔ $M_0$، تحت شرایط معقولی حل شده است [MST]، [MOY]، [S].

• فرمولهای «گذار از دیوار»[11]. در حالتی که $b_2^+(M) = 1$، ناوردای خمینهٔ $M$ به متریک روی آن نیز وابسته است. فضای متریکهای روی $M$ به‌طور مشخص به حجره[1]هایی تقسیم‌بندی می‌شود و گذار از یک «دیوار» برای عبور از یک حجره به حجرهٔ مجاور، ناوردا را طبق فرمول معینی یک واحد کم یا زیاد می‌کند، حداقل در حالتی که $b_1(M) = 0$ و بعد فضای زمینه مثبت و زوج باشد [M]. نتایجی نیز برای حالت $b_1(M) > 0$ در دست است.

• قضایای «صفر»[2]. موارد متعددی وجود دارد که ناوردای زایبرگ‌ـ‌ویتن صفر می‌شود، مثلاً اگر خمینهٔ مورد نظر به‌صورت جمع همبند[3] دو خمینهٔ $M_1$ و $M_2$ باشد که $b_2^+(M_i) > 0$؛ چنین خمینه‌ای خمینهٔ تحویل‌پذیر[4] خوانده می‌شود. حکم فوق خود نتیجه‌ای از قضیهٔ چسباندن است.

• حدس مینیمال. با توجه به قضیهٔ صفر خمینه‌های تحویل‌پذیر، و این نکته که خمینه‌های همتافته ناورداهای ناصفر دارند، تاوبز حدس زد که خمینه‌های همتافته، عناصر ساختمانی خمینه‌های چهاربعدی باشند، دست‌کم در حالت خمینه‌های ساده‌ـ‌همبند، یعنی هر چنین خمینه‌ای جمع همبند خمینه‌های همتافته باشد. سابو برای این حدس مثال ناقضی ساخت، یعنی خمینهٔ ساده‌ـ‌همبند تحویل‌ناپذیری که همتافته هم نیست [Sz]. درواقع خانواده‌ای از این مثالها وجود دارد [FS]. برخلاف بعد سه — که حدس هندسی‌سازی[5] ترستن[6] تجزیه‌ای از خمینه‌های سه‌بعدی به اجزای شناخته‌شده پیش می‌نهد — هنوز هیچ حدس صورت‌بندی‌شده‌ای برای شناسایی خمینه‌های چهاربعدی دوام نیافته است.

• هم‌ارزی نظریهٔ دانلدسن با نظریهٔ زایبرگ‌ـ‌ویتن و برنامهٔ فیهان[7]‌ـ‌لِنِس[8]. ویتن با توجه به استدلالها و قراین فیزیکی معتقد است که نظریهٔ ایشان با نظریهٔ دانلدسن هم‌ارز است و محاسبات یک نظریه را می‌توان متناظراً در نظریهٔ دیگر انجام داد. برنامه‌هایی برای اثبات ریاضی این مطلب در پیش است که شاید فعالترین آنها در حال حاضر برنامهٔ فیهان‌ـ‌لنس باشد [FL]. در این برنامه، تعمیمهایی از نظریهٔ زایبرگ‌ـ‌ویتن به گروههای پیمانه‌ای ناآبلی‌ای همچون $PU(2)$ در نظر گرفته می‌شوند.

• حدس «ساده نوع»[9]. تصور بر این است که ناوردای زایبرگ‌ـ‌ویتن برابر صفر است مگر آنکه بعد فضای پیمانه صفر باشد. تمام مثالهای بررسی‌شده تاکنون این حدس را تأیید می‌کنند ولی هنوز اثبات یا ردی برای حکم کلی وجود ندارد. در حال حاضر این شاید مهمترین و سرسخت‌ترین حدس موجود در نظریهٔ زایبرگ‌ـ‌ویتن باشد.

## معادلات زایبرگ‌ـ‌ویتن

فرض کنید $M$ یک خمینهٔ چهاربعدی هموار، جهتدار، بسته (فشرده و بی‌لبه) و همبند باشد. $M$ را به یک متریک ریمانی نیز مجهز کنید. (بعداً خواهیم دید که اگر $b_2^+(M) > 1$، ناورداهای تعریف‌شده مستقل از متریک ریمانی خواهند بود، یعنی ناورداها فقط به ساختار هموار $M$ بستگی دارند.) همچنین یک ساختار $\mathrm{Spin}^c(4)$ برای $M$ تثبیت کنید [تابلوی اسپین]. معادلات و ناورداهای زایبرگ‌ـ‌ویتن برحسب این ساختار $\mathrm{Spin}^c$ بیان می‌شوند. برای مقاصد عملی

1. plurigenera 2. Van de Ven 3. Kodaira 4. Kähler 5. rank 6. signature 7. gluing 8. Seifert-fibered space 9. knot 10. link 11. wall-crossing formulas

1. chamber 2. vanishing 3. connected sum 4. reducible 5. geometrization 6. William Thurston 7. Feehan 8. Leness 9. simple type

# گروه اسپین

فرض کنید $V$ یک فضای ضرب داخلی حقیقی باشد. فضای $\oplus_{n=0}^{\infty}\underbrace{V\otimes\cdots\otimes V}_{n}$ را به ایده‌آل دوطرفهٔ تولیدشده با عناصرِ به شکل $v\otimes v+||v||^{2}$ برای هر $v\in V$ تقسیم کنید. فضای خارج‌قسمتی به همراه ضرب $\otimes$ یک جبر $\mathbb{Z}/2\mathbb{Z}$ـ‌مدرج[1] است که جبر کلیفورد[2] خوانده شده با $\mathrm{Cl}(V)$ نشان داده می‌شود. معمولاً با مختلط‌شدهٔ[3] این جبر سروکار داریم: $\mathbb{C}\mathrm{l}(V)=\mathrm{Cl}(V)\otimes_{\mathbb{R}}\mathbb{C}$. اگر عنصری از $\mathrm{Cl}(V)$ به‌صورت حاصل‌ضرب تعداد زوجی از عناصر $V$ باشد آن‌را یک عنصر زوج می‌خوانیم. عناصر فرد به‌طور مشابه تعریف می‌شوند. زوج و فردبودن با توجه به رابطهٔ هم‌ارزی فوق خوش‌تعریف است و درجه‌بندی مذکور در بالا را به‌دست می‌دهد: $\mathrm{Cl}(V)=\mathrm{Cl}_{0}(V)\oplus\mathrm{Cl}_{1}(V)$. (در ضمن توجه کنید که $\mathbb{C}\mathrm{l}(V)$ به‌عنوان یک فضای برداری با $\Lambda^{*}(V)\otimes\mathbb{C}$ یکریخت است.)

اگر $\{e_1,\dots,e_n\}$ پایه‌ای یکه و متعامد برای $V$ باشد، عنصر $\omega_{\mathbb{C}}=i^{[\frac{n+1}{2}]}e_1e_2\dots e_n$ یک عنصر حجم برای $\mathbb{C}\mathrm{l}(V)$ خوانده می‌شود. محاسبه‌ای سرراست نشان می‌دهد که $\omega_{\mathbb{C}}^{2}=1$. ضرب کردن از چپ در $\omega_{\mathbb{C}}$، یک یکریختی خطی $\mathbb{C}\mathrm{l}(V)\to\mathbb{C}\mathrm{l}(V)$ القا می‌کند که مقادیر ویژهٔ $\pm 1$ دارد و در نتیجه تجزیه‌ای به شکل زیر حاصل می‌شود:

$$\begin{aligned}\mathbb{C}\mathrm{l}(V)&=\mathbb{C}\mathrm{l}^{+}(V)\oplus\mathbb{C}\mathrm{l}^{-}(V)\\&=\mathbb{C}\mathrm{l}^{+}_{0}(V)\oplus\mathbb{C}\mathrm{l}^{+}_{1}(V)\oplus\mathbb{C}\mathrm{l}^{-}_{0}(V)\oplus\mathbb{C}\mathrm{l}^{-}_{1}(V)\end{aligned}$$

می‌توان نشان داد که اگر $V$ زوج بعدی باشد، $\mathbb{C}\mathrm{l}^{+}_{0}(V)$ به‌عنوان یک جبر با $\mathbb{C}\mathrm{l}(W)$ یکریخت است، که $W$ زیرفضایی از $V$ با نقص بعد دو است:

$$V=W\oplus\mathbb{R}^{2}$$

همچنین، $\mathbb{C}\mathrm{l}(V)$ یک جبر ماتریسی $2\times 2$ روی $\mathbb{C}\mathrm{l}^{+}_{0}(V)$ است:

$$\mathbb{C}\mathrm{l}(V)\cong\mathbb{C}\mathrm{l}(W)\underset{\mathbb{C}}{\otimes}\mathbb{C}\mathrm{l}(\mathbb{R}^{2})$$

و $\mathbb{C}\mathrm{l}(\mathbb{R}^{2})\cong M_{2\times 2}(\mathbb{C})$. بنابراین، به کمک قضیهٔ آرتین‌ـودربرن[4] و به‌استقراء می‌توان ثابت کرد که اگر $\dim V=2n$، $\mathbb{C}\mathrm{l}(V)$ یکریخت با جبر ماتریسی $M_{2^{n}\times 2^{n}}(\mathbb{C})$ است و نمایش مختلط تحویل‌ناپذیر یگانه‌ای چون $S(V)$ دارد؛ عمل $\mathbb{C}\mathrm{l}(V)$ روی $S(V)$، یک یکریختی

$$\mathbb{C}\mathrm{l}(V)\cong\mathrm{End}(S(V))=S(V)\otimes S^{*}(V)$$

القا می‌کند. ضرب در $\omega_{\mathbb{C}}$ نیز یک یکریختی خطی

$$S(V)\to S(V)$$

القا می‌کند که مقادیر ویژهٔ $\pm 1$ دارد و لذا تجزیه‌ای به صورت

1. $\mathbb{Z}/2\mathbb{Z}$-graded 2. Clifford algebra 3. complexified

4. Artin-Wedderburn

القا می‌کند که مقادیر ویژهٔ $\pm 1$ دارد و لذا تجزیه‌ای به صورت $S(V)=S^{+}(V)\oplus S^{-}(V)$ حاصل می‌شود. تحت این تجزیه، $\mathbb{C}\mathrm{l}(V)$ به‌صورت ماتریسی زیر تجزیه می‌شود:

$$\mathbb{C}\mathrm{l}(V)=\begin{pmatrix}\mathbb{C}\mathrm{l}^{+}_{0}(V) & \mathbb{C}\mathrm{l}^{-}_{1}(V)\\ \mathbb{C}\mathrm{l}^{+}_{1}(V) & \mathbb{C}\mathrm{l}^{-}_{0}(V)\end{pmatrix}$$

یعنی

$$\begin{aligned}&\mathbb{C}\mathrm{l}^{+}_{0}(V)=\mathrm{Hom}(S^{+}(V),S^{+}(V))=\mathrm{End}(S^{+}(V)),\\&\mathbb{C}\mathrm{l}^{+}_{1}(V)=\mathrm{Hom}(S^{+}(V),S^{-}(V)),\\&\qquad\cdots\end{aligned}$$

علاوه بر این، در حالت خاص $\dim V=4$، تحت یکریختی $\mathbb{C}\mathrm{l}(V)$ و $\Lambda^{*}(V)\otimes\mathbb{C}$ به عنوان فضاهای برداری، $\mathbb{C}\mathrm{l}^{+}_{0}(V)$ به

$$(\Lambda^{2}_{+}(V)\otimes\mathbb{C})\oplus\mathbb{C}(\frac{1+\omega_{\mathbb{C}}}{2})$$

نگاشته می‌شود و جزء دوم این تجزیه درواقع «اثر» تبدیل خطی است.

حال زیرگروهی از گروه ضربی یکال[1]های $\mathrm{Cl}(V)$ را در نظر بگیرید که با بردارهای به طول یک تولید شده‌اند. اشتراک این زیرگروه با $\mathrm{Cl}_{0}(V)$ را $\mathrm{Spin}(V)$ می‌نامیم. اگر به‌جای $\mathrm{Cl}(V)$ از $\mathbb{C}\mathrm{l}(V)$ استفاده کنیم، گروه $\mathrm{Spin}^{c}(V)$ به‌دست می‌آید. داریم:

$$\mathrm{Spin}^{c}(V)=\mathrm{Spin}(V)\underset{\{\pm 1\}}{\times}\mathrm{U}(1)$$

می‌توان نشان داد که $\pi_1(\mathrm{SO}(n))=\mathbb{Z}/2\mathbb{Z}$ و $\mathrm{Spin}(n)$ یگانه گروه پوششی دو لایهٔ همبند $\mathrm{SO}(n)$ است ($n>2$). در ضمن،

$$\begin{aligned}&\mathrm{Spin}(3)=\mathrm{SU}(2)\\&\mathrm{Spin}(4)=\mathrm{SU}(2)\times\mathrm{SU}(2)\\&\mathrm{Spin}^{c}(4)=\mathrm{SU}(2)\times\mathrm{SU}(2)\times\mathrm{U}(1)/\{\pm 1\}\end{aligned}$$

توجه کنید که با توجه به تجزیهٔ بالا از $\mathrm{Spin}^{c}(4)$، افکنشهای متنوعی از این گروه به گروههای لی دیگر قابل تعریف است، مثلاً اگر برای تمیز گذاشتن بین مؤلفهٔ اول و دوم در تجزیهٔ $\mathrm{Spin}^{c}(4)$ به ترتیب از نمادهای $+$ و $-$ استفاده کنیم، خواهیم داشت:

$$\mathrm{Spin}^{c}(4)=\mathrm{SU}(2)^{+}\times\mathrm{SU}(2)^{-}\times\mathrm{U}(1)/\{\pm 1\}$$

$$\begin{array}{l}\nearrow\ \mathrm{SU}(2)^{+}\times\mathrm{U}(1)/\{\pm 1\}=\mathrm{U}(2)^{+}\\ \to\ \mathrm{SU}(2)^{-}\times\mathrm{U}(1)/\{\pm 1\}=\mathrm{U}(2)^{-}\\ \searrow\ \mathrm{U}(1)/\{\pm 1\}=\mathrm{U}(1)\end{array}$$

1. unit

حال اگر به این امر کلی توجه کنیم که هر کلاف اصلی با تعیین دوگان‌دور[1] های تعریف‌کنندهٔ آن مشخص می‌شود، و نیز اینکه هر همریختی گروهی $\rho : G \to H$ را می‌توان با دوگان‌دورهای $g_{\alpha\beta} : U_\alpha \cap U_\beta \to G$ ترکیب کرد تا دوگان‌دورهای $h_{\alpha\beta} = \rho \circ g_{\alpha\beta} : U_\alpha \cap U_\beta \to H$ را به‌دست آورد، نتیجه می‌گیریم که هر $\mathrm{Spin}^c(4)$-کلاف اصلی، به نوبهٔ خود دو $\mathrm{U}(2)$-کلاف و یک $\mathrm{U}(1)$-کلاف به‌طور طبیعی تعریف می‌کند. اگر نمایش استاندارد $\mathrm{U}(2)$ روی $\mathbb{C}^2$ و $\mathrm{U}(1)$ روی $\mathbb{C}^1$ را منظور کنیم، کلافهای برداری وابسته عبارت خواهند بود از $S^+$، $S^-$ و $\mathscr{L}$، که $S^\pm$ کلافهای برداری مختلط با رتبهٔ دو و $\mathscr{L}$ یک کلاف خطی مختلط (با رتبهٔ یک) روی $M$ است. این کلافهای برداری مختلط به‌ترتیب کلاف اسپینورهای مثبت، اسپینورهای منفی و دترمینان نامیده می‌شوند. داریم:

$$\mathscr{L} = \det(\widetilde{P}) = \det(S^+) = \det(S^-)$$

از سوی دیگر، به‌کمک روشهای توپولوژی جبری، و توجه به این نکته که $G$-کلافهای اصلی غیریکریخت روی $M$ متناظرند با عناصر گروه کوهومولوژی $H^1(M;G)$، می‌توان ثابت کرد که هر خمینهٔ چهاربعدی جهتدار دارای یک ساختار $\mathrm{Spin}^c(4)$ است و تعداد این ساختارها متناظر است با $H^2(M;\mathbb{Z})$. (این مطلب در مورد ساختار $\mathrm{Spin}(4)$ الزاماً صحیح نیست.) یک ساختار $\mathrm{Spin}^c(4)$ روی $M$ به‌معنای زیر است. کلاف مماس $M$ را در نظر بگیرید. با توجه به جهتدار بودن $M$ و مجهزبودن آن به یک متریک ریمانی، $\mathrm{SO}(4)$-کلاف اصلی مماس متناظر، $P$، را بسازید. دوگان‌دورهای چنین کلافی درواقع همان دوگان‌دورهای کلاف مماس $TM \to M$ هستند: $g_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{SO}(4)$. یک ساختار $\mathrm{Spin}^c(4)$ روی $M$ یعنی ترفیعهایی از دوگان‌دورهای مزبور به $\mathrm{Spin}^c(4)$ به‌نحوی که ترفیعها هم در شرط دوگان‌دور صدق کنند.

1. cocycle

$$\begin{tikzcd} & \mathrm{Spin}^c(4) \arrow[d, "p"] \\ U_\alpha \cap U_\beta \arrow[ur, dashed, "\tilde{g}_{\alpha\beta}"] \arrow[r, "g_{\alpha\beta}"'] & \mathrm{SO}(4) \end{tikzcd}$$

(دراینجا، $p : \mathrm{Spin}^c(4) \to \mathrm{SO}(4)$ افکنش طبیعی است که به صورت $\mathrm{Spin}(4) \times \mathrm{U}(1)/\{\pm 1\} \to \mathrm{Spin}(4)/\{\pm 1\}$ به‌دست می‌آید.)

معادلاً، یک ساختار $\mathrm{Spin}^c(4)$ روی $M$ یعنی ترفیعی از کلاف اصلی مماس به یک $\mathrm{Spin}^c(4)$-کلاف اصلی:

$$\begin{tikzcd} \mathrm{Spin}^c(4) \arrow[r] \arrow[dr, "p"'] & \tilde{P} \arrow[dr, dashed] & \\ & \mathrm{SO}(4) \arrow[r] & P \arrow[d] \\ & & M \end{tikzcd}$$

**یادداشت.** خمینه‌های مختلط، یک ساختار $\mathrm{Spin}^c$ طبیعی دارند: فضای مماس یک $\mathrm{U}(n)$-کلاف است و نگاشت

$$\begin{aligned} \mathrm{U}(n) &\to \mathrm{SO}(2n) \times \mathrm{U}(1) \\ A &\mapsto (\underset{\substack{\uparrow \\ \text{(شمول طبیعی)}}}{\imath(A)}, \det(A)) \end{aligned}$$

ترفیعی به $\mathrm{U}(n) \to \mathrm{Spin}^c(2n)$ دارد:

$$\begin{tikzcd} & \mathrm{Spin}^c(2n) = \mathrm{Spin}(2n) \times_{\{\pm 1\}} \mathrm{U}(1) \arrow[d] \\ \mathrm{U}(n) \arrow[ur, dashed] \arrow[r] & \mathrm{SO}(2n) \times \mathrm{U}(1) \end{tikzcd}$$

تذکر فوق برای خمینه‌های تقریباً مختلط هم برقرار است.

می‌توان فرض کرد که این ساختار توسط یک کلاف خطی مختلط $\mathscr{L}$ روی $M$ مشخص شده است. در ضمن $S^+$ و $S^-$ را کلافهای اسپینورهای مثبت و منفی روی $M$ وابسته به ساختار $\mathrm{Spin}^c$ مذکور بگیرید [تابلوی اسپین]. معادلات زایبرگ-ویتن درواقع یک دستگاه دو معادلهٔ دومجهولی است. این دستگاه از دو معادلهٔ دیفرانسیل پاره‌ای مرتبهٔ یک به‌صورت زیر تشکیل شده است.

$$(\mathrm{SW}) \begin{cases} F_A^+ = \psi \otimes \psi^* - \frac{1}{2}|\psi|^2 I & (1) \\ \partial\!\!\!/_A(\psi) = 0 & (2) \end{cases}$$

مجهولات این دستگاه عبارت‌اند از یک هموستار $A$ روی کلاف $\mathscr{L} = \det(\widetilde{P})$ و یک اسپینور مثبت $\psi \in \Gamma(S^+)$. اکنون به شرح مختصر معادلات مزبور می‌پردازیم.

(۱) در این معادله، $F_A \in \Omega^2(M; i\mathbb{R})$ خمیدگی هموستار $A$ است و $F_A^+ = F_A + *F_A$ جزء خودْدوگان[1] آن [تابلوی هموستار]. توجه کنید که $i\mathbb{R}$ جبر لی وابسته به گروه کلاف خطی $\mathscr{L}$ است که عبارت است از $\mathrm{U}(1)$، و $*$ ستارهٔ هاج است. بنابراین سمت چپ معادلهٔ (۱) در $\Omega^2_+(M; i\mathbb{R})$ قرار دارد. برای توضیح سمت راست معادله، بنویسید $\psi = \binom{\psi_1}{\psi_2}$. بنابراین

$$\begin{aligned} \psi \otimes \psi^* &= \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \otimes (\overline{\psi}_1 \ \overline{\psi}_2) = \begin{pmatrix} \psi_1\overline{\psi}_1 & \psi_1\overline{\psi}_2 \\ \psi_2\overline{\psi}_1 & \psi_2\overline{\psi}_2 \end{pmatrix} \\ &= \begin{pmatrix} |\psi_1|^2 & \psi_1\overline{\psi}_2 \\ \psi_2\overline{\psi}_1 & |\psi_2|^2 \end{pmatrix} \end{aligned}$$

1. self-dual

و در ضمن

$$\psi\otimes\psi^* \in\Gamma(S^+\otimes(S^+)^*)=\Gamma(\mathrm{Hom}(S^+,S^+))=\Gamma(\mathrm{Cl}^+_0(\widetilde{P}))$$

از نحوهٔ تعریف $\mathrm{Cl}^+_0$ مشخص می‌شود که این کلاف برداری یکریخت است با $\Lambda^2_+(M)\oplus\mathbb{C}(\frac{1+\omega}{2})$، و بررسی این یکریختی نشان می‌دهد که مؤلفهٔ اول این جمع همبند متناظر است با همریختیهایی در $\mathrm{Hom}(S^+,S^+)$ که اثر[1] آنها صفر است. با توجه به اینکه $\mathrm{tr}(\psi\otimes\psi^*-\frac{1}{2}|\psi|^2I)=0$، سمت راست معادلهٔ (۱) هم در $\Omega^2_+(M;i\mathbb{R})$ قرار دارد و لذا هر دو طرف معادله از یک سنخ هستند. در ضمن دقت کنید که این معادله نسبت به $\psi$ از درجهٔ دوم است.

(۲) در این معادله $\partial\!\!\!/_A:\Gamma(S^+)\to\Gamma(S^-)$ عملگر دیراک وابسته به هموستار $A$ است. این معادله درواقع بیان می‌کند که $\psi$ یک اسپینور همساز[2] است. [تابلوی هموستار].

## فضای پیمانه و ناوردای زایبرگ-ویتن

ابتدا به فضای جوابهای معادلات (SW) توجه می‌کنیم، یعنی زیرمجموعه‌ای از $\mathscr{C}(\widetilde{P})=\mathscr{A}(\mathscr{L})\times\Gamma(S^+)$ که متشکل از عناصر $(A,\psi)$ با شرط $\mathrm{SW}(A,\psi)=0$ است، که

$$\mathrm{SW}(A,\psi)=(F_A^+-\psi\otimes\psi^*+\frac{1}{2}|\psi|^2I,\ \partial\!\!\!/_A(\psi))$$

در حالت کلی این فضا می‌تواند حتی تهی باشد یا بینهایت بعدی. درواقع توجه کنید که اگر $(A,\psi)$ جوابی برای این معادلات باشد، هر تبدیل پیمانه‌ای، یعنی هر نگاشت کلافی
$$\sigma:\ \begin{tikzcd}[column sep=small]\widetilde{P} \arrow[rr] \arrow[dr] & & \widetilde{P} \arrow[dl] \\ & M & \end{tikzcd}$$
، این جواب را به جواب دیگری تبدیل خواهد کرد. به بیان دیگر، $((\det\sigma)^*A,S^+(\sigma^{-1})(\psi))$ هنوز هم در معادلات (SW) صدق می‌کند. به این ترتیب از هر جواب $(A,\psi)$ می‌توان به تعداد نگاشتهای کلافی جواب جدید تولید کرد. به عبارتی، گروه تبدیلات پیمانه‌ای $\mathscr{G}(\widetilde{P})=\mathrm{Aut}(\widetilde{P})$، که متناظر است با نگاشتهای $M\to Z(\mathrm{Spin}^c(4))=S^1$، روی $\mathscr{C}(\widetilde{P})$ عمل می‌کند. مدار عمل این گروه بُعد نامتناهی دارد. به منظور انجام عملیات هندسی و محاسبات آنالیزی مجاز در این فضاهای بینهایت‌بعدی، ترجیح می‌دهیم که با تتمیم[3] فضاهای مزبور کار کنیم. پس فضاهای $\mathscr{A}(\mathscr{L})$ و $\Gamma(S^+)$ و در نتیجه $\mathscr{C}(\widetilde{P})=\mathscr{A}(\mathscr{L})\times\Gamma(S^+)$ را به نُرمهای سوبولف $L^2_2$ مجهز کرده، و تتمیم آنها را با نُرم فوق در نظر می‌گیریم [تابلوی سوبولف]. همچنین تتمیم $\mathscr{G}(\widetilde{P})=\mathrm{Maps}(M,S^1)$ را با نُرم $L^2_3$ در نظر می‌گیریم. اینها خمینه‌های هیلبرت هستند و بسیاری از قضایای مهم آنالیز ــ ازجمله قضایای تابع ضمنی و تابع وارون ــ برای آنها برقرار است.

درواقع فضای مورد علاقهٔ ما، که فضای پیمانه نامیده می‌شود، فضای جوابهای معادلات (SW) است با تقریب عمل گروه پیمانه‌ای. این فضا، با توپولوژی خارج‌قسمت، یک فضای هاسدورف است [M]. پایاگر[4] جواب

1. trace 2. harmonic 3. completion 4. stabilizer

$(A,\psi)$ عبارت است از

$$\mathrm{Stab}(A,\psi)=\begin{cases}\{1\} & \psi\neq 0\\ S^1=\{M\to S^1\ \text{نگاشتهای ثابت}\} & \psi=0\end{cases}$$

جواب $(A,\psi)$ را که $\psi\equiv 0$ «تحویل‌پذیر»[1] می‌خوانیم و در غیر این‌صورت «تحویل‌ناپذیر»[2]. (این نامگذاری در قیاس با نظریهٔ دانلدسُن صورت گرفته است و محتوای مفهومی ندارد.) فضای پیمانهٔ معادلات SW، یعنی فضای جوابهای معادلات زایبرگ-ویتن، با تقریب یکی‌گرفتن جوابهایی که با یک تبدیل پیمانه‌ای به هم تبدیل می‌شوند، در حالت کلی یک خمینهٔ هموار با بُعد متناهی نیست. برای رفع این مشکل، اختلالی[3] جزئی در معادلات به نحو زیر ایجاد می‌کنیم.

$$(\mathrm{SW}_h)\begin{cases}F_A^+=\psi\otimes\psi^*-\frac{1}{2}|\psi|^2I+ih & (1)_h\\ \partial\!\!\!/_A(\psi)=0 & (2)\end{cases}$$

که در آن $h\in\Omega^2_+(M;\mathbb{R})$ دو-فرمی خودْدوگان است. فضای جوابهای این معادلات، $\mathscr{S}_h(\widetilde{P})$، نیز تحت عمل گروه تبدیلات پیمانه‌ای ناورداست و در ضمن برای $h$ نوعی[4]، جواب تحویل‌پذیر ندارد. فضای خارج‌قسمتی، یعنی فضای پیمانهٔ $\mathscr{M}_h(\widetilde{P}):=\mathscr{S}_h(\widetilde{P})/\mathscr{G}(\widetilde{P})$ نیز مانند $\mathscr{M}(\widetilde{P})$ فضایی هاسدورف است. به هر جواب $(A,\Psi)\in\mathscr{S}_h(\widetilde{P})$ یک همبافت بیضوی[5] نظیر می‌کنیم.

$$\begin{aligned}0\to\Omega^0(M;i\mathbb{R})&\xrightarrow{\mathscr{D}^0}\Omega^1(M;i\mathbb{R})\oplus\Gamma(S^+)\\ &\xrightarrow{\mathscr{D}^1}\Omega^2_+(M;i\mathbb{R})\oplus\Gamma(S^-)\to 0\end{aligned}\qquad(\mathscr{E})$$

در این همبافت بیضوی، $\mathscr{D}^0$ درواقع مشتق عمل گروه $\mathscr{G}(\widetilde{P})$ روی $\mathscr{C}(\widetilde{P})$ و $\mathscr{D}^1$ بخش خطی‌شدهٔ معادلات $\mathrm{SW}_h$ است.

$$\mathscr{D}^0(f)=(2df,-f.\Psi)$$

$$\mathscr{D}^1(\alpha,\psi)=\begin{pmatrix}d^+\alpha & -B(\Psi,\psi)\\ \frac{1}{2}\alpha.\Psi & \partial\!\!\!/_A(\psi)\end{pmatrix}$$

در روابط بالا، . نشان‌دهندهٔ ضرب کلیفرد، $d^+\alpha$ نمایانگر جزء خودْدوگان $d\alpha$ و $B$ فرم دوخطی متناظر با فرم درجهٔ دوم ظاهر شده در معادلات SW است:

$$B(\Psi,\psi)=\Psi\otimes\psi^*+\psi\otimes\Psi^*-(\mathrm{Re}<\Psi,\psi>)I$$

همان‌طور که مشاهده می‌شود، نگاشتهای $\mathscr{D}^0$ و $\mathscr{D}^1$ مستقل از اختلال $h$ هستند. این همبافت دارای سه گروه کوهومولوژی است. صفر بودن $H^0(\mathscr{E})$ به‌معنای یک‌به‌یک بودن $\mathscr{D}^0$ و معادل با تحویل‌ناپذیربودن جواب $(A,\Psi)$ است. $H^2(\mathscr{E})$ فضای مانع[6] نامیده می‌شود و صفرشدن آن به معنی عادی[7] بودن نقطهٔ $\xi=(A,\Psi)$ در $\mathscr{M}_h(\widetilde{P})$ است. هر دوی این موارد،

1. reducible 2. irreducible 3. perturbation 4. generic
5. elliptic complex 6. obstruction space 7. regular

## هموستار و مشتق همموردا، خمیدگی، عملگر دیراک

یک همومستار روی کلاف اصلی $P \xrightarrow{\pi} M$ یعنی یک یک-فرم $\omega \in \Omega^{1}(P;\mathrm{g})$ که در دو خاصیت زیر صدق می‌کند:

الف) اگر $R_g : P \to P$ با $R_g(p) = p \cdot g$ داده شده باشد در این صورت $R_g^{*}\omega = g^{-1}\omega g$؛

ب) برای $p \in P$ داده شده، اگر $R_p : G \to P$ با $R_p(g) = p \cdot g$ تعریف شده باشد آنگاه $R_p^{*}\omega = \omega_{MC}$، که

$$\omega_{MC} = g^{-1}dg \in \Omega^{1}(G;\mathrm{g})$$

یگانه یک-فرمی است که تحت ضرب از سمت چپ در عناصر $G$ ناورداست و در $e \in G$ همانی است.

فضای تمام همومستارهای روی $P$، فضایی است مستوی که مدل آن فضای برداری $\Omega^{1}(M;\mathrm{ad}P)$ است. $\mathrm{ad}P = P \underset{G \circlearrowright \mathrm{ad}}{\times} \mathrm{g}$ کلاف برداری وابسته به کلاف اصلی $P$ و نمایش الحاقی گروه $G$ بر جبر لی خود $\mathrm{g}$ است:

$$\mathrm{ad} : G \times \mathrm{g} \to \mathrm{g}$$

$$(g,v) \mapsto gvg^{-1}$$

فرض کنید $E = P \underset{G \circlearrowright \rho}{\times} V$ کلاف برداری وابسته به کلاف اصلی $P$ و نمایش $\rho$ گروه $G$ روی فضای برداری $V$ باشد، یعنی $\rho : G \to GL(V)$ یک همریختی گروهی است و $E = P \times V/\sim$، که

$$(p \cdot g, v) \sim (p, \rho(g) \cdot v)$$

در این صورت، همومستار $\omega \in \Omega^{1}(P;\mathrm{g})$ یک مشتق همموردا $\nabla : \Omega^{i}(M;E) \to \Omega^{i+1}(M;E)$ تعریف می‌کند. تعاریفی برای همومستار، انتقال موازی، ترفیع خمها و ... روی یک کلاف برداری وجود دارد که به آنها نمی‌پردازیم، ولی آنها همگی متناظرند با مشتقهای همموردا روی آن کلاف. فضای همومستارهای روی کلاف برداری $E$ را با $\mathscr{A}(E)$ و فضای مقاطع آن کلاف را با $\Gamma(E)$ نمایش می‌دهیم.

خمیدگی یک همومستار $\omega \in \Omega^{1}(P;\mathrm{g})$، دو فرمی $\Omega \in \Omega^{2}(M;\mathrm{ad}P)$ است به‌طوری که $\pi^{*}\Omega \in \Omega^{2}(P;\mathrm{g})$ در رابطهٔ

$$\pi^{*}\Omega = d\omega + \omega \wedge \omega$$

صدق می‌کند.

اکنون به توصیف عملگر دیراک می‌پردازیم. فرض کنید $M$ به یک ساختار $\mathrm{Spin}^{c}(4)$ مانند $\widetilde{P}$ مجهز شده باشد و $\mathscr{L} = \det(\widetilde{P})$ به یک همومستار $A$. با تثبیت یک متریک ریمانی برای $M$، همومستار یگانهٔ لوی-چیویتا برای $TM$ به‌دست می‌آید. با توجه به اینکه

$$\begin{array}{ccc} & \mathrm{Spin}^{c}(4) = \mathrm{Spin}(4) \underset{\{\pm 1\}}{\times} \mathrm{U}(1) & \\ \swarrow & & \searrow \\ \mathrm{Spin}(4)/\{\pm 1\} = \mathrm{SO}(4) & & \mathrm{U}(1)/\{\pm 1\} = \mathrm{U}(1) \end{array}$$

و هر یک از دو کلاف اصلیِ حاصل از افکنشهای مزبور مجهز به یک همومستار است (($\mathrm{U}(1)$-کلاف همان $\mathscr{L}$ و $\mathrm{SO}(4)$-کلاف همان $TM$ است)، همومستاری نیز برای $\mathrm{Spin}^{c}(4)$-کلاف $\widetilde{P}$ تعریف می‌شود. بالاخص همومستارهایی برای کلافهای برداری $S^{+}$، $S^{-}$ و $S = S^{+} \oplus S^{-}$ به‌دست می‌آید. اگر $\nabla : \Omega^{\cdot}(M;S^{\pm}) \to \Omega^{1}(M;S^{\pm})$ مشتقگیری همموردای وابسته باشد، عملگر دیراک $\partial\!\!\!/_{A} : \Gamma(S^{\pm}) \to \Gamma(S^{\mp})$ را به‌صورت زیر تعریف می‌کنیم. فرض کنید $\{e_1, e_2, e_3, e_4\}$ پایه‌ای یکه و متعامد برای $T_xM$ باشد. $e_j \in T_xM$ را می‌توان به‌عنوان عنصری از جبر کلیفرد $\mathrm{Cl}(4)$ در نظر گرفت. از آنجا که $S$ یک $\mathrm{Cl}(4)$-مدول است، $e_j$ را می‌توان در عناصر $S$ ضرب کلیفرد کرد. برای $\psi \in \Gamma(S)$ تعریف می‌کنیم

$$\partial\!\!\!/_{A}(\psi) = \sum_{j=1}^{4} e_j \cdot \nabla_{e_j}(\psi)$$

محاسبهٔ سرراستی نشان می‌دهد که این تعریف مستقل از انتخاب پایه است.

---

یعنی صفرشدن $H^{\cdot}(\mathscr{E})$ و $H^{2}(\mathscr{E})$، را می‌توان با انتخاب نوعیِ $h$ محقق ساخت. «عادی بودن» نقطهٔ $\xi$ به این معنی است که فضای مماس $\mathscr{M}_h$ در این نقطه درواقع همان فضای مماس زاریسکی $H^{1}(\mathscr{E})$ است. بنابراین نقطهٔ $\xi$ نقطهٔ تکین[1] نیست. درواقع با انتخاب نوعیِ $h$ می‌توان مطمئن بود که $\mathscr{M}_h(\widetilde{P})$ خمینه‌ای هموار است و بُعد آن همان اندیس[2] همبافت بیضوی $\mathscr{E}$، که از طریق قضیهٔ اتیا-سینگر[3] قابل محاسبه است و برابر است با

$$d = \frac{1}{4}(c_1(\mathscr{L})^{2} - 2\chi(M) - 3\sigma(M))$$

که در آن $c_1(\mathscr{L})$ ردهٔ چرن اول کلاف خطی $\mathscr{L} = \det(\widetilde{P})$، $\chi(M)$ شاخص اویلر خمینه و $\sigma(M)$ نشان تقاطعی آن است [M]. به‌کمک روشهای آنالیزی و فرمولهای از نوع بوکنر[1] در هندسهٔ ریمانی می‌توان ثابت کرد که $\mathscr{M}_h(\widetilde{P})$ فشرده نیز هست [M].

توجه کنید که $\mathscr{M}_h(\widetilde{P})$ ممکن است تهی باشد، حتی اگر بُعد صوری[2] آن یعنی $d$ مثبت باشد. ما هنوز هیچ جوابی برای معادلات نیافته‌ایم. درواقع به‌کمک تخمینهای خمیدگی می‌توان نشان داد که $\mathscr{M}(\widetilde{P})$ فقط برای تعدادی متناهی $\mathrm{Spin}^{c}$-ساختارِ $\widetilde{P}$ ناتهی است و بعد صوری نامنفی دارد [M].

1. singular 2. index 3. Atiyah-Singer

1. Bochner type 2. formal dimension

## فضاهای سوبولف

نُرم سوبولف روی فضای مقاطع یک کلاف برداری $E$ به‌صورت زیر تعریف می‌شود.

$$||s||_{L^p_k} = \left( \int_M \sum_{i=0}^{k} || \nabla^i s ||^p \right)^{1/p}, \quad k \in \mathbb{N}, 1 < p < \infty$$

که $s \in \Gamma(E)$ و $\nabla$ یک هموستار روی $E$ است.

تتمیم فضای $\Gamma(E)$ با نرم $L^p_k$ را با $\Gamma_{L^p_k}(E)$ یا $L^p_k(E)$ نمایش می‌دهیم. احکام زیر برای فضاهای سوبولف برقرارند:

- مقاطع $C^\infty$ در $L^p_k(E)$ چگال‌اند.
- (سوبولف) نشانندهٔ پیوسته‌ای چون $L^p_k \hookrightarrow L^{p'}_{k'}$ وجود دارد اگر $k' \le k$ و $k' - \frac{n}{p'} \le k - \frac{n}{p}$، که در آن $n$ بعد خمینهٔ $M$ است. نشانندهٔ مزبور عملگری فشرده است اگر $k' < k$ و $k' - \frac{n}{p'} < k - \frac{n}{p}$.
- نشانندهٔ پیوسته‌ای چون $L^p_k \hookrightarrow C^r$ وجود دارد اگر $r \le k - \frac{n}{p}$. در اینجا $r > 0$. نشانندهٔ مزبور فشرده است اگر $r < k - \frac{n}{p}$.
- ضرب پیوسته‌ای چون $L^{p_1}_{k_1} \otimes L^{p_2}_{k_2} \to L^p_k$ وجود دارد اگر $k \le \min(k_1, k_2)$ و $k - \frac{n}{p} \le (k_1 - \frac{n}{p_1}) + (k_2 - \frac{n}{p_2})$.

در حالاتی خاص، مثلاً وقتی‌که M یک رویهٔ کِی‌لِر باشد، می‌توان معادلات زایبرگ‌ـویتن را به صراحت حل کرد و مشاهده کرد که $\mathscr{M}(\widetilde{P})$ عملاً می‌تواند ناتهی باشد.

حال می‌توان کلافی خطی روی $\mathscr{M}_h(\widetilde{P})$ به شرح زیر ساخت. به‌جای گروه تبدیلات پیمانه‌ای $\mathscr{G}(\widetilde{P})$، گروه تبدیلات پیمانه‌ای «پایه‌دار»[1] $\mathscr{G}^\circ(\widetilde{P})$ را در نظر بگیرید و فضای خارج‌قسمتی $\widetilde{\mathscr{M}}_h(\widetilde{P}) := \mathscr{S}_h(\widetilde{P})/\mathscr{G}^\circ(\widetilde{P})$ را بسازید. یک تبدیل پیمانه‌ای پایه‌دار تبدیلی است که در یک نقطهٔ مشخص و تثبیت‌شدهٔ $p_0 \in M$ به‌صورت همانی عمل می‌کند. به‌عبارت دیگر، تابع متناظر $g : M \to S^1$ در شرط $g(p_0) = 1$ صدق می‌کند. از آنجا که $\mathscr{G}(\widetilde{P})/\mathscr{G}^\circ(\widetilde{P}) \simeq \mathrm{U}(1)$، $\widetilde{\mathscr{M}}_h(\widetilde{P})$ یک $\mathrm{U}(1)$ـکلاف روی $\mathscr{M}_h(\widetilde{P})$ است و در نتیجه دارای ردهٔ چرن اولی چون $\mu \in H^2(\mathscr{M}_h, \mathbb{Z})$ است.

اکنون می‌خواهیم ناوردای زایبرگ‌ـویتن خمینهٔ $M$ و $\mathrm{Spin}^c$ـساختار $\widetilde{P}$ را تعریف کنیم. اگر $d = \dim(\mathscr{M}_h)$ فرد باشد، ناوردای مزبور را صفر می‌گیریم. اگر $d = \dim(\mathscr{M}_h)$ زوج باشد، $\mu^{d/2} \in H^{\mathrm{top}}(\mathscr{M}_h, \mathbb{Z})$ و قرار می‌دهیم

$$\mathrm{SW}(M, \widetilde{P}) := \int_{\mathscr{M}_h} \mu^{d/2}$$

$\mathscr{M}_h$ دارای یک جهت طبیعی است هر گاه جهتی برای $H^0(M; i\mathbb{R}) \oplus H^1(M; i\mathbb{R}) \oplus H^2_+(M; i\mathbb{R})$ تثبیت شود.) ثابت می‌شود که اگر $b^+_2(M) > 1$، ناوردای بالا خوش‌تعریف است و به انتخابهای مختلف انجام شده (متریک ریمانی، دوـفرم $h$، ...) بستگی ندارد و فقط به ساختار هموار $M$ و ساختار $\mathrm{Spin}^c$ معین وابسته است. اگر $b^+_2(M) = 1$، در بسیاری موارد «فرمولهای گذار از دیوار» وجود دارند که نحوهٔ تغییر ناوردای فوق را بر اثر تغییر متریک ریمانی محاسبه می‌کنند. در این حالت، ناوردا برای متریکهای ریمانیی محاسبه می‌شود که $[c_1(\mathscr{L})]$ در $H^2(M; \mathbb{R})$ بر $\mathscr{H}^2_+$ تحت حاصلضرب ناوی[1] عمود نیست، که در آن $\mathscr{H}^2_+$ فضای دوـفرمهای خودْدوگانِ همساز تحت متریک ریمانی مزبور است. در چنین حالتی، معادلات SW و اختلالات کوچک آن جوابِ تحویل‌ناپذیر ندارند. حال مقدار ناوردای زایبرگ‌ـویتن به‌صورتی که در قبل تعریف شد بستگی به حجره‌ای دارد که متریک ریمانی در آن واقع است. دیوارهای این حجره عبارت‌اند از فضای متریکهایی چون $g$ که فرم خودْدوگان وابسته به آن، $\omega(g)$، بر $c_1(\mathscr{L})$ عمود است. اگر $H_1(M; \mathbb{Z}) = 0$، گذار از یک دیوار مقدار ناوردا را به‌اندازهٔ $(-1)^{d/2}$ عوض می‌کند، که $d$ بعد $\mathscr{M}(\widetilde{P})$ است.

## مثالها

۱. ناوردای زایبرگ‌ـویتن خمینه‌های کِی‌لِر به‌صورت زیر به‌دست می‌آید. فرض کنید $\widetilde{P}_0$ ساختار $\mathrm{Spin}^c$ طبیعی چنین خمینه‌ای باشد [تابلوی اسپین]. در این صورت $\det(\widetilde{P}_0) = K^{-1}_M = \Lambda^{2,0} M$ و ناورداهای زایبرگ‌ـویتن چنین هستند: $\mathrm{SW}(\widetilde{P}_0) = 1$، $\mathrm{SW}(-\widetilde{P}_0) = \pm 1$ و ناورداهای سایر ساختارها صفرند. (علامت $\mathrm{SW}(-\widetilde{P}_0)$ به‌طور مشخص قابل تعیین است و به دو ناوردای هندسی جبری $M$ بستگی دارد: $(-1)^{1+p_g-q}$.) درواقع، فضای پیمانه به‌طور کاملاً صریح قابل توصیف است. فضای جوابهای SW، با تقریب پیمانه، با فضای زیر متناظر است:

- اگر $\deg(\mathscr{L}) > 0$، که $\mathscr{L} = \det(\widetilde{P})$: با فضای زوجهای $(\overline{\partial}_{\mathscr{L}}, \alpha)$، که $\overline{\partial}_{\mathscr{L}}$ یک ساختار هولومورف روی $\mathscr{L}$ است و $\alpha$ مقطع هولومورف ناصفری از $\mathscr{L}_0 = \sqrt{K \otimes \mathscr{L}}$؛ با تقریب یکی بودن ساختار هولومورف و ضرب مقاطع در عددی ثابت.
- اگر $\deg(\mathscr{L}) < 0$: با فضای زوجهای $(\overline{\partial}_{\mathscr{L}}, \beta)$، که در اینجا $\beta$ مقطع هولومورف ناصفری از $\sqrt{K \otimes \mathscr{L}^{-1}}$ است؛ با همان تقاریب بالا.
- اگر $\deg(\mathscr{L}) = 0$: با فضای هموستارهای پادْخودْدوگان[2] روی $\mathscr{L}$؛ با تقریب پیمانه.

به‌علاوه، اگر $\mathscr{L}_0$ تاب کلاف $\mathrm{Spin}^c$ (پیچشِ افزون بر کلاف طبیعی) و $A_0$ هموستار روی آن باشد،

$$\mathscr{L} = K^{-1} \otimes \mathscr{L}_0^2$$

$$S^+(\widetilde{P}) = \Lambda^0(M; \mathscr{L}_0) \oplus \Lambda^{0,2}(M; \mathscr{L}_0)$$

$$S^-(\widetilde{P}) = \Lambda^{0,1}(M; \mathscr{L}_0)$$

$$\partial\!\!\!/ = \sqrt{2}(\overline{\partial} + \overline{\partial}^*)$$

$$\partial\!\!\!/_A = \sqrt{2}(\overline{\partial}_{A_0} + \overline{\partial}^*_{A_0})$$

۲. تاوبز ثابت می‌کند که برای خمینه‌های همتافته با $b^+_2 > 1$، مقدار ناوردای SW برای ساختار $K^{-1}$ برابر یک است (تعمیم قضیهٔ فوق) و در ضمن

1. based

1. cup product 2. anti self-dual (ASD)

برابر ناوردای گروموف است (که از شمارش تعداد خمهای شبه‌هولومورف به‌دست می‌آید).

**۳.** فضای پیمانهٔ $S^4$ فقط از نقاط تحویل‌پذیر تشکیل شده است (به موجب اتحاد بوکنر و مثبت بودن خمیدگی اسکالر $S^4$).

**۴.** $\mathbb{C}P^2\#\mathbb{C}P^2$ یک خمینهٔ همتافته نیست. چون بُعد فضای پیمانه به پیمانهٔ دو همنهشت است با $1-b_1+b_2^+$، که برای این خمینه عدد $3=2+0-1$ را به‌دست می‌دهد، پس بُعد فضای پیمانه فرد است و ناوردای زایبرگ‌ـویتن صفر.

## سپاسگزاری

نگارنده از حمایت مالی پژوهشگاه دانشهای بنیادی (IPM) در هنگام تهیهٔ این مقاله سپاسگزار است.

## مراجع

[ک] کاشانی، سیدمحمدباقر؛ خمینه‌های چهاربعدی؛ نشر ریاضی ۱۲ (دی ۱۳۷۳) ۸ـ۱۲.

[F] Furuta, "Monopole equation and the $\frac{11}{8}$ conjecture", preprint, 1995.

[FL] P. M. N. Feehan, T. G. Leness, et al, Series of articles on $PU(2)$ monopoles, preprints available at
http://www. math. ohio-state. edu/~feehan/preprints. html.

[FM] Robert Friedman and John W. Morgan, "Algebraic surfaces and Seiberg-Witten invariants", *J. Algebraic Geom.*, (3) **6** (1997) 445-479.

[FQ] Robert Friedman and Zhenbo Qin, "On complex surfaces diffeomorphic to rational surfaces", *Invent. Math.*, (1) **120** (1995) 81-117.

[FS] Ronald Fintushel and Ronald J. Stern, "Knots, links, and 4-manifolds", *Invent. Math.*, (2) **134** (1998) 363-400.

[G] M. Gromov, "Pseudoholomorphic curves in symplectic manifolds", *Invent. Math.*, (2) **82** (1985) 307-347.

[KM] P. B. Kronheimer and T. S. Mrowka, "The genus of embedded surfaces in the projective plane", *Math. Res. Lett.*, (6) **1** (1994) 797-808.

[KM1] P. B. Kronheimer and T. S. Mrowka, "Gauge theory for embedded surfaces. I.", *Topology*, (4) **32** (1993) 773-826.

[KM2] P. B. Kronheimer and T. S. Mrowka, Gauge theory for embedded surfaces. II. "*Topology* , (1) **34** (1995) 37-97.

[KM3] P. B. Kronheimer and T. S. Mrowka, "Embedded surfaces and the structure of Donaldson's polynomial invariants. ", *J. Differential Geometry*, (3) **41** (1995) 573-734.

[M] John W. Morgan, *The Seiberg-Witten Equations and Applications to the Topology of Smooth Four-manifolds*, volume 44 of Mathematical Notes, Princeton University Press, Princeton, NJ (1996).

[MOY] Tomasz Mrowka, Peter Ozsváth, and Baozhen Yu, "Seiberg-Witten monopoles on Seifert fibered spaces", *Comm. Anal. Geom.*, (4) **5** (1997) 685-791.

[MST] John W. Morgan, Zoltán Szabó, and Clifford Henry Taubes, "A product formula for the Seiberg-Witten invariants and the generalized Thom conjecture", *J. Differential Geom.*, (4) **44** (1996) 706-788.

[N] Liviu I. Nicolaescu, *Notes on Seiberg-Witten Theory*, volume 28 of Graduate Studies in Mathematics, Amer. Math. Soc., Providence, RI (2000).

[OSz] Peter Ozsváth and Zoltán Szabó, "The symplectic Thom conjecture", *Ann. of Math. (2)*, (1) **151** (2000) 93-124.

[S] Pedram Ṣafari, *A Gluing Theorem for Seiberg-Witten Moduli Spaces*, PhD thesis, Columbia University (2000).

[Sz] Zoltán Szabó, "Simply-connected irreducible 4-manifolds with no symplectic structures", *Invent. Math.*, (3) **132** (1998) 457-466.

[T1] Clifford Henry Taubes, "The Seiberg-Witten invariants and symplectic forms", *Math. Res. Lett.*, (6) **1** (1994) 809-822.

[T2] Clifford Henry Taubes, "The Seiberg-Witten Gromov invariants", *Math. Res. Lett.*, (2) **2** (1995) 221-238.

[T3] Clifford Henry Taubes, "More constraints on symplectic forms from Seiberg-Witten invariants", *Math. Res. Lett.*, (1) **2** (1995) 9-13.

[T4] Clifford Henry Taubes, *Seiberg Witten and Gromov Invariants for Symplectic 4-manifolds*, edited by Richard Wentworth, International Press, Somerville, MA (2000).

[T5] Clifford Henry Taubes, "Gr=SW: counting curves and connections, *J. Differential Geom.*, (3) **52** (1999) 453-609,

[T6] Clifford Henry Taubes, "Gr→SW: from pseudo-holomorphic curves to Seiberg-Witten solutions", *J. Differential Geom.*, (2) **51** (1999) 203-334.

[T7] Clifford Henry Taubes, "SW→Gr: from the Seiberg-Witten equations to pseudo-holomorphic curves", *J. Amer. Math. Soc.*, (3) **9** (1996) 845-918.

[T8] Clifford Henry Taubes, "Counting pseudo-holomorphic submanifolds in dimension 4", *J. Differential Geom.*, (4) **44** (1996) 818-893.

[W] Edward Witten, "Monopoles and four-manifolds", *Math. Res. Lett.*, (6) **1** (1994) 769-796.

******

* پدرام صفری، دانشگاه صنعتی شریف

safari@sharif.edu

# کارآیی نامعقول ریاضیات در زیست‌شناسی مولکولی*

آرتر لسک*
ترجمهٔ افرا علیشاهی

عنوان مقالهٔ من اقتباسی است از عنوان مقالهٔ مشهور ویگنر، «کارآیی نامعقول ریاضیات در علوم طبیعی [۱]». البته این عنوان در فیزیک و در زیست‌شناسی مولکولی، از دو بابت مخالفْ توجه‌برانگیز است. در فیزیک، بدیهی است که ریاضیات کارآمد است ــ بسیاری از غولهایی که فیزیکدانها بر شانه‌های آنها ایستاده‌اند ریاضیدان‌اند ــ و مایهٔ شگفتی است که ویگنر این را نامعقول می‌داند. در زیست‌شناسی مولکولی، نقش واقعی ریاضیات بدیهی نیست، و بیم آن می‌رود که انتظار کارآیی از ریاضیات نامعقول باشد، بیمی که در این مورد از فیزیک بسیار موجه‌تر است. البته، بسیاری از ابزارهای متداول در زیست‌شناسی مولکولی محاسباتی ــ مثلاً جستجو در پایگاه داده‌ها برای یافتن دنباله‌های شبیه به یک دنبالهٔ یافت‌شده ــ مسلماً مبتنی بر ریاضیات و علوم کامپیوتر هستند. اما اینکه آیا درک غایی ما از فرآیندهای حیاتی به زبان ریاضیات بیان شود ــ آن‌چنان‌که، مثلاً، مفاهیم تقارن زمینه‌ساز بیان قوانین فیزیک هستند ــ یا به زبان سنتیِ توصیفیِ «روایی» زیست‌شناسی، هنوز محل بحث است.

چرا تردید در کارآیی ریاضیات در زیست‌شناسی می‌تواند معقول باشد؟ خصوصیات مشاهده‌شدهٔ سیستمهای زنده به‌وسیلهٔ ترکیبی از اینها تعیین می‌شوند:

- قوانین فیزیک و شیمی
- سازوکار تکامل
- تصادف تاریخی

تفکیک تأثیر اینها از هم دشوار است، و کشاکش سازنده‌ای که بین آنهاست، تحقیقات ما را تحت تأثیر قرار می‌دهد. بسیاری از قوانین فیزیک، جهان طبیعی را ــ که شامل سیستمهای زنده هم می‌شود ــ با تعیین روابط بین شرایط اولیه و نهایی توصیف می‌کنند. در زیست‌شناسیْ پیچیدگی مجموعهٔ شرایط اولیهٔ ممکن، دشواریهایی پدید می‌آورد. نقش عظیم تصادف تاریخی ما را به تردید می‌اندازد و به زانو درمی‌آورد: حتی اگر قوانین بنیادی فیزیک و شیمی نتایج ساده‌ای به دنبال داشته باشد که توصیفات مبسوطی از فرآیندهای حیات در اختیار ما قرار دهد، ما ممکن است قادر به کشف آنها نباشیم، زیرا پدیده‌های مورد مطالعهٔ ما پیچیده‌ترند، در برابر ایده‌آل‌سازیهای ساده‌کننده مقاومت می‌کنند، و خصوصیاتی از خود نشان می‌دهند که تحت تأثیر شدید انتخاب شرایط اولیه از بین مجموعهٔ بسیار عظیم و پراکنده‌ای از شرایط ممکن است. سیبها در زیست‌شناسی نقشی بسیار مهمتر از افتادن بر سر مردم ایفا می‌کنند.

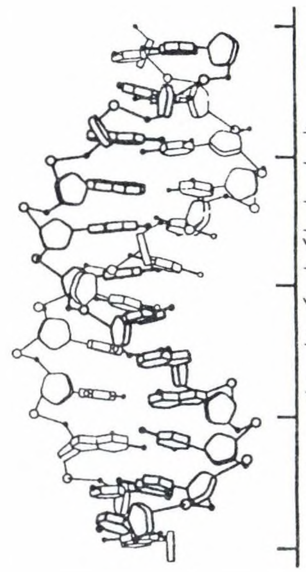
## موضوع بحث در زیست‌شناسی مولکولی

اشیاء مورد مطالعهٔ ما حداقل شکلی دارند که می‌توانیم سعی کنیم ریاضیات را در مورد آنها به‌کارگیریم. این اشیاء عبارت‌اند از:

- توالیهای ژنها در DNA
- توالیهای آمینواسیدها در پروتئینها
- ساختار پروتئینها
- کارکرد پروتئینها

احتمالاً خوانندگان در بارهٔ پروژهٔ ژنوم شنیده‌اند؛ هدف این پروژه تعیین توالی کامل DNA در موجودات زنده است: یعنی مجموعهٔ نقشه‌ها. توالیهای DNA در ژنومها حاوی همهٔ اطلاعاتی هستند که یک موجود زنده برای تولد، بزرگ شدن و رشد، و مرگ به آنها نیاز دارد. با تکمیل تعیین توالی ژنوم مخمر در سال ۱۹۹۶، ما همانقدر در بارهٔ یک سلول مخمر می‌دانیم که خود سلول مخمر می‌داند. این عبارت آن‌قدر که به‌نظر می‌رسد متکبرانه نیست. ما واقعاً همهٔ اطلاعات را داریم. باید پذیرفت که ما نمی‌توانیم این اطلاعات را به همان کارآمدی تفسیر کنیم که یک سلول مخمر می‌تواند، اما ما مجموعهٔ کامل نقشه‌ها را در اختیار داریم. ولی این نقشه‌ها تنها یک توصیف ایستا از ساختار و عملکرد بالقوه به‌دست می‌دهند؛ می‌ماند اینکه مشاهدات خود را تا مرحلهٔ یکپارچه‌سازی توصیف و کارکرد پروتئین در زمان و فضا درون یک موجود

**توالی‌ای از بازها در DNA . . .** **به توالی‌ای از آمینواسیدها در پروتئین ترجمه می‌شود . . .** **که خودبه‌خود تا می‌شود تا ساختار سه‌بعدی دقیقی بسازد**



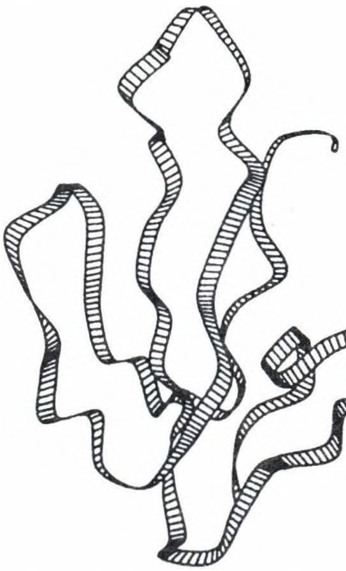

سه باز

| | | | |
|---|---|---|---|
| UUU F | UCU S | UAU Y | UGU C |
| UUC F | UCC S | UAC Y | UGC C |
| UUA L | UCA S | UAA "EOF" | UGA "EOF" |
| UUG L | UCG S | UAG "EOF" | UGG W |
| CUU L | CCU P | CAU H | CGU R |
| CUC L | CCC P | CAC H | CGC R |
| CUA L | CCA P | CAA Q | CGA R |
| CUG L | CCG P | CAG Q | CGG R |
| AUU I | ACU T | AAU N | AGU S |
| AUC I | ACC T | AAC N | AGC S |
| AUA I | ACA T | AAA K | AGA R |
| AUG M | ACG T | AAG K | AGG R |
| GUU V | GCU A | GAU D | GGU G |
| GUC V | GCC A | GAC D | GGC G |
| GUA V | GCA A | GAA E | GGA G |
| GUG V | GCG A | GAG E | GGG G |

یک آمینواسید

**کد ژنتیک**

**«جدول ترجمه»**



**شکل ۱** جریان اطلاعات در خلال «خوانش» از روی یک ژن. ژنها، یا نقشه‌های بنیادی موجودات زنده، در ساختار DNA گنجانده شده‌اند (راست). در سمت راست شکل، مارپیچ مضاعف DNA را می‌بینیم، که حاوی دو نوار به‌هم پیچیده است که یکی با خطوط باریک و دیگری با خطوط پهن رسم شده است. «اثر پلکانی» توسط مجموعه‌ای از زیرواحدهای شیمیایی به‌نام «باز» پدید می‌آید. در هر نقطه روی هر نوار یکی از چهار باز A، T، G یا C ممکن است قرارگیرند. خوانندهٔ تیزبین می‌تواند ببیند که این بارها ــ «لبه‌های قائم» پله‌های پلکان مارپیچی ــ به شکلهای مختلف ظاهر می‌شوند. بازهای هم‌سطح هر دو نوار با هم ارتباط برقرار می‌کنند، و این ارتباط مستلزم مکمل بودن مطلق است: حضور یک A روی یک نوار مستلزم حضور یک T روی نوار دیگر است، و یک G روی یک نوار به یک C روی نوار دیگر نیاز دارد، و برعکس، یک T روی یک نوار مکمل یک A و یک C مکمل یک G روی نوار دیگر است. به این ترتیب، هر نوار حاوی اطلاعات کافی برای ساختن نوار همراه خود است. به لحاظ منطقی، راه تکثیر DNA آن است که نوارها را از هم جدا کنیم و مکمل هرکدام از نوارهای مجزا را دوباره بسازیم. توالیهای بازها در ژنها، به‌کمک یک جدول ترجمهٔ مستقیم به‌نام «کد ژنتیک» توالیهای آمینواسید در پروتئینها را رمزگشایی می‌کنند (وسط). پیام ژنتیکی با الفبای چهار حرفی A، T، G و C نوشته می‌شود. پروتئینها هم پلیمرهایی هستند که توالی‌ای از سازه‌های شیمیایی را در خود دارند، چنانکه در هر نقطه یکی از بیست آمینواسید ممکن قرار می‌گیرد، که با نمادهای A، C، D، E، F، G، H، J، K، L، M، N، P، Q، R، S، T، V، W، Y مشخص می‌شوند. برای مشخص کردن مجموعهٔ ۲۰ آمینواسید، هرکدام به بیش از دو باز نیاز دارد؛ درواقع از روی توالی DNA در یک زمان سه باز خوانده می‌شود، با افزونگی‌ای که برای تکامل مطلقاً ضروری است (سه سه‌تایی از بازها برای علامتهای انتهایی «پایان حیات» نگهداری شده‌اند). پروتئینها خود به خود تا می‌شوند تا ساختارهای طبیعی فعال سه‌بعدی بسازند (چپ). این نقطه‌ای است که در آن طبیعت از توالیهای یک‌بعدی ژنها به جهان سه‌بعدی‌ای که ما در آن زندگی می‌کنیم جهش می‌کند. مثال فوق یک توکسین از یک مار آبی است، یکی از ساختارهای پروتئین متعددی که به‌وسیلهٔ بلورنگاری با اشعهٔ X تعیین شده است. هر ژن توالی‌ای از بازها دارد، که ابتدا توالی آمینواسیدهای یک پروتئین، و سپس ساختار سه‌بعدی آن، و سپس کارکرد آن را تعیین می‌کند.

زنده بسط دهیم. جمع‌آوری این اطلاعات به «پروژهٔ پروتئوم» معروف است. پروژه‌ای که در حال جمع‌آوری قواست تا در عصر پس از ژنوم ایفای نقش کند.

میزان اندازه‌گیری توالیهای ژنها بسیار زیاد و فزاینده است. در سال ۱۹۹۸ توالی کامل DNAی یک نوع کرم موسوم به Caenorhabditis elegans تکمیل شد ($۹٫۷ \times ۱۰^{۷}$ باز). محتمل است که در سالهای ۱۹۹۹ و ۲۰۰۰ به‌ترتیب شاهد تکمیل توالی DNAی مگس سرکه ($۱٫۸ \times ۱۰^{۸}$ باز) و ژنوم انسان ($۳٫۴ \times ۱۰^{۹}$ باز) و بسیاری موجودات زندهٔ کوچک و بزرگ دیگر باشیم. لویی پانزدهم می‌توانست بگوید: «بعد از من، طوفان»، اما نوح نمی‌توانست؛ ما هم نمی‌توانیم.

توالیها و ساختارهایی که ما مطالعه می‌کنیم ارتباطات مهمی با هم دارند. در سطح مولکول، توالیهای ژنهای DNA، توالیهای آمینواسیدهای پروتئینها را رمزگشایی می‌کنند. سپس توالیهای آمینواسید پروتئینها ساختار سه‌بعدی پروتئینها را تعیین می‌کنند. آنگاه ساختار پروتئینها، کارکرد آنها را تعیین می‌کنند (شکل ۱). برای تعیین کارکرد پروتئین، یک ساختار سه‌بعدی دقیق ضروری است، زیرا تعاملهای مورد نیاز به‌کنار هم قرارگرفتن بخشهای مختلف مولکولها در روابط فضایی دقیق بستگی دارند. نهایتاً پسخورد کارکرد پروتئین به دنبالهٔ ژنها ــ از طریق تکامل به روش انتخاب طبیعی ــ این حلقه را کامل می‌کند.

توالیهای DNA درکامپیوترهای ما، رشته‌های حرفی هستند، یعنی اشیاء یک‌بعدی. ژنها یا زیررشته‌هایی از توالیهای ژنوم، براساس یک کدنویسی تقریباً جهان‌شمول به توالیهای آمینواسید پروتئینها ترجمه می‌شوند. این توالیها هم به‌وسیلهٔ رشته‌های حرفی یک‌بعدی بازنمایی می‌شوند. سپس پروتئینها خود به خود تا می‌شوند تا ساختارهای سه‌بعدی «طبیعی» یکتا بسازند. (شاهد این مدعا آن است که می‌توان آنها را به‌وسیلهٔ حرارت تخریب کرد ــ ساختار سه‌بعدی را ویران کرد ــ و وقتی سرد شوند مجدداً شکل اولیهٔ خود را پیدا می‌کنند، مانند آلیاژهای شکل‌ـحافظه[1]). تاخوردگی خود به‌خودِ پروتئینها

1. shape-memory

همان نقطه‌ای است که طبیعت موجب می‌شود توالیهای یک‌بعدی ژنها به جهان سه‌بعدی‌ای که ما در آن زندگی می‌کنیم جهش کنند.

## اهداف زیست‌شناسی مولکولی محاسباتی

اهداف ما از پرداختن به این مبحث چیست؟ نخست آنکه بتوانیم مشابهتها و تفاوتهای موجود بین توالیها و بین ساختارها را به‌سادگی توصیف و طبقه‌بندی کنیم. توپولوژی فضای توالی، فضای ساختار، و فضای کارکرد پروتئین چیست؟ نگاشتهای بین این فضاها چگونه‌اند؟ ما می‌خواهیم که با استفاده از تکامل به‌عنوان اصل سازماندهی، بتوانیم روابط بین توالی، ساختار و کارکرد پروتئین را توصیف و پیشگویی کنیم.

از کجا باید کار را شروع کرد؟ سیدنی برنر یک بار با افسوس گفت: «مسأله در زیست‌شناسی این است که هیچ نوسانگر همسازی وجود ندارد.» منظور او از این گفته این بود که در زیست‌شناسی، برخلاف فیزیک، راه گریزی از پیچیدگی نیست، حتی از طریق ایده‌آل‌سازی. در فیزیک، نوسانگر همساز یک مسألهٔ ساده است که می‌توان به روشهای متعددی آن‌را به دقت حل کرد؛ این مسأله در مورد برخی از پدیده‌ها دقیقاً قابل اعمال است، و برای سایر پدیده‌ها تخمین مؤثری است. نوسانگر همساز در فیزیک یک بستر آزمون سنتی برای روشهای جدید است. در حقیقت، زیست‌شناسی مولکولی محاسباتی دو «نوسانگر همساز» به تعبیر برنر دارد: هم‌ردیف‌سازی توالیها، و برهم‌نهی ساختارها. این اعمال، که می‌توان آنها را با دقت و کارآیی انجام داد، پایهٔ بسیاری از تحلیلهای روابط توالی-ساختار در زیست‌شناسی مولکولی هستند. اکنون دانستن این نکته که جهان واقعی اغلب ناهمساز است شگفت‌انگیز نخواهد بود. با این‌همه، بسیاری از این ابزارهای ارزشمند به‌کمک این موارد ساده ساخته شده‌اند.

اما، ابزارها پاسخ می‌سازند نه سؤال. پژوهش در این حوزه همچنان بر تعامل بین دانشمندان و داده‌ها، با کمک روشهای ریاضیاتی و محاسباتی، متکی است.

## توالیها و هم‌ردیف‌سازی آنها

توالیهای ژن و پروتئین شکل رشته‌های حرفی به‌خود می‌گیرند. برای توالیهای ژن، حروف از مجموعهٔ چهار عضوی $\{\mathrm{A, T, G, C}\}$ انتخاب می‌شوند که نشان‌دهندهٔ نوکلئوتیدهای آدنین، تایمین، گوانین و کایتوزین هستند. برای توالیهای پروتئین، حروف از یک مجموعهٔ بیست عضوی انتخاب می‌شوند که نشان‌دهندهٔ بیست آمینواسید معمولی هستند.

### هم‌ردیف‌سازی

هم‌ردیف‌سازی دو رشتهٔ حرفی به معنی تعیین یک تناظر معنی‌دار بین مؤلفه‌های آنهاست. برای دو رشتهٔ حرفی زیر

c t a t a a t c و g c t g a a c

دو هم‌ردیف‌سازی ممکن عبارت‌اند از

g c t g − a a − c  g c t g a − a − − c

و

− c t a t a a t c  − c t − a t a a t c

چگونه می‌توان تعیین کرد که کدام‌یک از این دو، یا کدام‌یک از بسیاری از هم‌ردیف‌سازیهای ممکن دیگر، بهترین هم‌ردیف‌سازی است؟ آیا ما می‌توانیم متریکی برای رشته‌های حرفی طرح کنیم و فاصلهٔ بین آنها را تعریف کنیم؟ شاخصهای عدم شباهت بین رشته‌های حرفی عبارت‌اند از:

(۱) فاصلهٔ همینگ، که بین دو رشتهٔ هم‌طول تعریف می‌شود، یعنی تعداد مکانهایی که حاوی حروف نامشابه هستند.

(۲) فاصلهٔ لونشتاین[1] بین دو رشته که الزاماً هم‌طول نیستند، یعنی کمترین تعداد «عملیات ویرایش» مورد نیاز برای تبدیل یک رشته به رشتهٔ دیگر، که در اینجا عمل ویرایش عبارت است از حذف، درج، یا تغییر یک حرف در هر توالی. هر توالی داده شده از عملیات ویرایش، یک هم‌ردیف‌سازی یکتا را القا می‌کند. اما عکس این مطلب صحیح نیست.

در زیست‌شناسی مولکولی، می‌دانیم که درج و حذف در توالیهای ژن و پروتئین رخ داده‌اند. بنابراین فاصلهٔ همینگ به‌اندازهٔ کافی عام نیست. به‌علاوه، شواهدی وجود دارد که نشان می‌دهد وقوع برخی تغییرات محتملتر از تغییرات دیگر است. بنابراین حتی فاصلهٔ لونشتاین نیز باید تعمیم داده شود تا براساس مدل تکاملی زیربنایی ما، وزنهای متفاوتی برای عملیات مختلف ویرایش در آن منظور گردد. برای مثال، به‌نظر می‌آید که جهشها محافظه‌کارانه رخ می‌دهند؛ جایگزینی یک آمینواسید در یک پروتئین با آمینواسید دیگری که اندازه یا خواص فیزیکی-شیمیایی یکسان دارد محتملتر است تا جایگزینی آن با آمینواسید دیگری که خصوصیات آن بیشتر متفاوت است. برای انعکاس این مطلب، به‌جای شمارش گسستهٔ عملیات ویرایش، به هر تغییر در توالی یک «هزینه» (متعلق به $\mathbb{R}$) نسبت می‌دهیم. همچنین شواهدی وجود دارد که نشان می‌دهد هزینهٔ یک فاصلهٔ خالی، مانند مدل لونشتاین، با طول آن متناسب نیست؛ اگرچه انتخاب مناسب وزنهای فواصل خالی به شکل تابعی از طول دقت زیادی می‌خواهد، در بسیاری از طرحها از یک تابع خطی با یک پارامتر $\alpha$ برای مقدار اولیه دادن به فاصلهٔ خالی و یک پارامتر کوچکتر $\beta$ برای بسط فاصله، به‌منظور محاسبهٔ هزینهٔ فاصلهٔ خالی به شکل

$$\alpha + \beta \times (\text{۱} - \text{طول فاصلهٔ خالی})$$

استفاده می‌شود. الگوریتمهایی وجود دارند که با کمینه کردن مجموع هزینه‌های عملیات ویرایشی که یک رشته را به رشتهٔ دیگر تبدیل می‌کند، بهترین هم ردیف‌سازی را تعیین می‌کنند.

مسألهٔ هم‌ردیف‌سازی بهینهٔ توالی را به شکل صوری می‌توان چنین بیان کرد. دو رشتهٔ حرفی $A = a_1a_2\dots a_n$ و $B = b_1b_2\dots b_m$ داده شده‌اند، که هر $a_i$ و $b_j$ عضوی از یک مجموعهٔ الفبای $\mathcal{A}$ است. فرض کنید $\mathcal{A}^+ = \mathcal{A} \cup \{\emptyset\}$. یک توالی از عملیات ویرایش، مجموعه‌ای از زوجهای مرتب $(x, y)$ است، که $x, y \in \mathcal{A}^+$. عملیات ویرایش منفرد عبارت‌اند از:

جایگزینی $b_j$ با $a_i$، که با $(a_i, b_j)$ نشان داده می‌شود،

حذف $a_i$ از رشتهٔ $A$، که با $(a_i, \phi)$ نشان داده می‌شود،

حذف $b_j$ از رشتهٔ $B$، که با $(\phi, b_j)$ نشان داده می‌شود.

تابع هزینهٔ $d$ بر عملیات ویرایش به شکل زیر تعریف می‌شود:

$$d(a_i, b_j) = \text{هزینهٔ یک جهش}$$

1. Levenshtein

هزینهٔ یک حذف یا درج $= d(\emptyset, b_j)$ یا $d(a_i, \emptyset)$

و کمترین فاصلهٔ وزن‌دار بین رشته‌های $A$ و $B$ عبارت است از

$$D(A,B) = \min_{A \to B} \sum d(x,y)$$

که در آن $x, y \in \mathcal{A}^+$ و مقدار مینیمم به‌ازای همهٔ توالیهای عملیات ویرایش که $A$ را به $B$ تبدیل می‌کنند محاسبه می‌شود. اگر $d(x,y)$ یک متریک بر $\mathcal{A}^+$ باشد، $D(A,B)$ متریکی بر رشته‌های حروف از $\mathcal{A}^+$ خواهد بود. (در این بیان از مسأله فرض می‌شود که هزینهٔ فاصله‌های خالی مستقل از طول آنهاست؛ طرحهای واقع‌بینانه‌تر که به فاصله‌های خالی وزن می‌دهند، تعمیمی از این طرح هستند.)

مسألهٔ یافتن $D(A,B)$ است و یک یا چند هم‌ردیف‌سازی که با آن متناظرند. الگوریتمی که این مسأله را در زمان $\mathcal{O}(mn)$ حل می‌کند زمان درازی است که شناخته شده است، و در بسیاری از مسائل از قبیل ویرایش متن، تشخیص گفتار، و تحلیل آواز پرندگان به‌کارگرفته شده است [۲]. این الگوریتم توسط مقاله تأثیرگذار نیدلمن و وونش [۳]، به زیست‌شناسی وارد شد.

چند خصوصیت این الگوریتم قابل توجه‌اند.

• این الگوریتم یک بهینهٔ مطلق به‌دست می‌دهد: توجه داشته باشید که این یکی از دو «نوسانگر همساز» زیست‌شناسی مولکولی محاسباتی است. ما روشی در اختیار داریم که مطمئنیم که در مینیممهای موضعی به دام نخواهد افتاد.

• این خبرِ خوب بود. خبر بد اینکه تعبیر نتایج چندان سرراست نیست. اگرچه توالی‌ای از عملیات ویرایش که از یک هم‌ردیف‌سازی بهینه ناشی شده است ممکن است با یک مسیر تکاملی واقعی متناظر باشد، اما اثبات اینکه چنین است ممکن نیست. هر چه فاصلهٔ ویرایش بیشتر باشد، تعداد مسیرهای تکاملی معقول بیشتر است. نه تنها هم‌ردیف‌سازی‌های بهینه ممکن است یکتا نباشند، بلکه ممکن است هم‌ردیف‌سازیهای زیربهینهٔ بسیاری وجود داشته باشند که ارزش آنها کاملاً به مقدار بهینه نزدیک باشد. مثلاً، فیچ و اسمیت ژنهای جوجه را برای هموگلوبینهای $\alpha$ و $\beta$ آزمایش کردند [۴]. آنها ۱۷ هم‌ردیف‌سازی بهینه یافتند، که یکی از آنها با هم‌ردیف‌سازی مبتنی بر ساختارهای هموگلوبین شناخته شده مطابقت داشت، و بیش از هزار هم‌ردیف‌سازی یافت شد که با مقدار بهینه کمتر از ۵٪ اختلاف داشت.

## مشکلات ناشی از هم‌ردیف‌سازیِ جفت‌جفتِ توالیها

مشاهده شده است که با تکامل پروتئینها، توالیهای آمینواسیدها بسیار سریعتر از ساختار واگرا می‌شوند. در بسیاری از موارد می‌توان یک رابطهٔ تکاملی بین دو ساختار پروتئین یافت، حتی اگر هیچ شباهت قابل درکی بین توالیهای ژنها یا توالیهای آمینواسیدها وجود نداشته باشد. آنچه اتفاق می‌افتد این است: ژنها فضای توالیهای DNA را می‌کاوند، اما انتخاب طبیعی به مثابهٔ ترمزی در برابر تغییر ساختار عمل می‌کند تا کارکرد را حفظ کند. افزونگی[1] در کد ژنتیک ــ این واقعیت که تعداد زیادی از بازهای سه‌گانه یک آمینواسید یکسان را کدنویسی می‌کنند، و بسیاری از تغییرات تک‌بازی منجر به حصول آمینواسیدهایی با

1. redundancy

```
TYLWEFLLKLLQDR.EYCPRFIKWTNREKGVFKLV..DSKAVSRLWGMHKN.KPD
VQLWQFLLEILTD..CEHTDVIEWVG.TEGEFKLT..DPDRVARLWGEKKN.KPA
IQLWQFLLELLTD..KDARDCISWVG.DEGEFKLN..QPELVAQKWGQRKN.KPT
IQLWQFLLELLSD..SSNSSCITWEG.TNGEFKMT..DPDEVARRWGERKS.KPN
IQLWQFLLELLTD..KSCQSFISWTG.DGWEFKLS..DPDEVARRWGKRKN.KPK
IQLWQFLLELLQD..GARSSCIRWTG.NSREFQLC..DPKEVARLWGERKR.KPG
IQLWHFILELLQK..EEFRHVIAWQQGEYGEFVIK..DPDEVARLWGRRKC.KPQ
VTLWQFLLQLLRE..QGNGHIISWTSRDGGEFKLV..DAEEVARLWGLRKN.KTN
ITLWQFLLHLLLD..QKHEHLICWTS.NDGEFKLL..KAEEVAKLWGLRKN.KTN
LQLWQFLVALLDD..PTNAHFIAWTG.RGMEFKLI..EPEEVARLWGIQKN.RPA
IHLWQFLKELLASP.QVNGTAIRWIDRSKGIFKIE..DSVRVAKLWGRRKN.RPA
RLLWDFLQQLLNDRNQKYSDLIAWKCRDTGVFKIV..DPAGLAKLWGIQKN.HLS
RLLWDYVYQLLSD..SRYENFIRWEDKESKIFRIV..DPNGLARLWGNHKN.RTN
IRLYQFLLDLLRS..GDMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKGNRKK
LRLYQFLLGLLTR..GDMRECVWWVEPGAGVFQFSSKHKELLARRWGQQKGNRKR

 L  fl  lL          i W       F         a  WG  K
```

**شکل۲** هم‌ردیف‌سازی چندگانهٔ توالیهای ناتمام یک خانواده از پروتئینها که دامنه‌های ETS نامیده می‌شوند. هر خط با توالی آمینواسیدهای یک پروتئین متناظر است، که با توالی‌ای از حروفی که هر کدام نمایندهٔ یک آمینواسید است مشخص می‌شود. با نگاه به هر ستون می‌توان آمینواسیدهایی را که در آن مکان در هر یک از پروتئینهای خانواده پدیدار می‌شوند مشاهده کرد. به این طریق الگوهای اولویت‌دار قابل رؤیت می‌شوند. مثلاً مکان سوم حاوی یک لوسین یا L در هر توالی است ــ این امر دلالت بر این دارد که برخی محدودیتهای ساختاری یا کارکردی، تکامل را از تغییر وضعیت این مکان بازداشته‌اند. حروف زیر جدول مکان سازه‌های نامتغیر (حروف بزرگ) یا نامتغیر با یک استثنا (حروف کوچک) را مشخص می‌کنند. به توزیع نابرابر تغییر در ستونهای مختلف توجه کنید. تناوب سازه‌های ماندگار (۳، ۴ یا ۸) از وجود مارپیچها در پروتئین خبر می‌دهد، که خبر صحیحی است. الگوهای دیگر عمیقتر پنهان شده‌اند، و تعیین آنها نیاز به تحلیل محاسباتی دارد. چنین الگوهایی ممکن است حاوی همبستگیهایی میان توزیع آمینواسیدها در مکانهای مختلف باشند. مثلاً، در ستون چهارم از سمت چپ آمینواسید تیروزین یا Y، تنها در دو توالی آخر پدیدار می‌شود، سایر توالیها دارای تریپتوفان یا W هستند. یک همبستگی تقریبی در الگوی تغییر بین این ستون و ستونهای چهارم و پنجم از راست وجود دارد. بسیاری عقیده دارند (یا حداقل امیدوارند) که **همبستگیهای** الگوهای تغییر در مکانهای مختلف چنین جدولی از توالیها، سرنخهایی در بارهٔ محلهایی که در فضای سه‌بعدی بر هم اثر می‌گذارند به ما بدهد. متأسفانه قرائن بسیار ضعیف است.

خواص فیزیکی-شیمیایی یکسان می‌شود ــ نتایج ساختاری تغییرات توالی را تعدیل می‌کند.

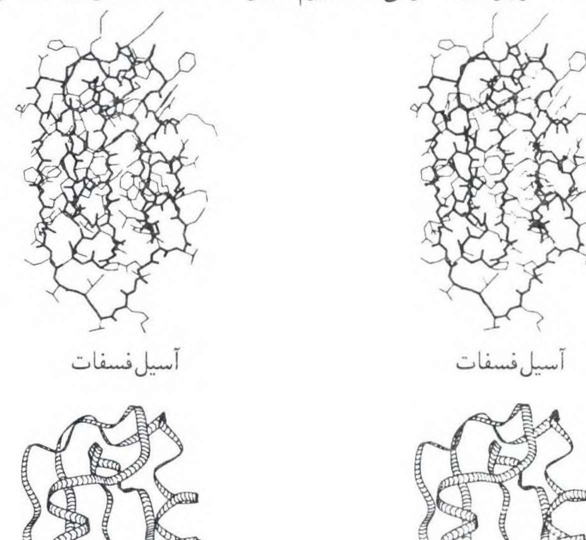
حتی اگر مشابهت در سطح توالی قابل تشخیص باشد، برای پروتئینهایی که ارتباط دوری با هم دارند، به کمک مقایسهٔ ساختارها که آخرین راه چاره است، می‌توان فهمید که هم‌ردیف‌سازی جفت‌جفتِ بهینهٔ توالیها اغلب منجر به جواب غلط می‌شود.

اما اگر تعداد زیادی توالی مرتبط در دسترس باشند، هم‌ردیف‌سازی چندگانهٔ توالیها نتایج با ارزشتر و دقیقتری نسبت به هم‌ردیف‌سازی جفت‌جفتِ توالیها به‌دست می‌دهد. چرا هم‌ردیف‌سازیهای چندگانه اطلاعات توالی را گسترش می‌بخشند؟ از اینجا الگوهای ماندگاری ظاهر می‌شوند. گستره و طبیعت تغییر در هر یک از مکانها راهنمای مهمی برای تعیین نقش ساختاری یا کارکرد ناحیه‌های مختلف توالی است (شکل ۲). برای مثال، سازه[1]هایی که در یک خانوادهٔ کامل از پروتئینها ماندگار [=بی‌تغییر] بوده‌اند معمولاً در کارکرد دخیل‌اند، یا دست‌کم اغلب نقشی اساسی در ساختار ایفا می‌کنند. برعکس، مناطقی که عملیات درج و حذف در آنها زیاد صورت می‌گیرد معمولاً با

1. residue

مناطق حاشیه‌ای متناظرند. (یک توالی در قبال ساختار خود منفعل است، یک جفت توالی هم‌ردیف‌شده، ساختار خود را نجوا می‌کنند، سه یا چند توالی ساختار خود را به صدای بلند فریاد می‌کنند).

اگر توالیها تنها به‌صورت غیرمستقیم به ساختار اشاره می‌کنند، چرا مستقیماً به سراغ ساختار نرویم؟ علت این است که مقدار داده‌ها در مورد توالیهای شناخته‌شده بسیار بیشتر از مقدار داده‌های ساختاری است. برای حدود ۲۰ موجود زنده، توالی همهٔ ژنوم تعیین شده است، که توالیهای کامل




**شکل ۳** پروتئینها چنان ساختارهای پیچیده‌ای هستند که لازم آمده است تا ابزارهای ویژه‌ای برای نمایش آنها ساخته شود. این شکل یک پروتئین نسبتاً کوچک به نام آسیل فسفات را با سه درجهٔ مختلف ساده‌سازی نشان می‌دهد. بالا: مدل کلی کامل؛ زنجیرهٔ اصلی پررنگتر از شاخه‌های جنبی است. وسط: مسیر زنجیر، توسط یک خم درون‌یابی‌شدهٔ هموار بازنمایی شده است، فلشها جهت زنجیره را تعیین می‌کنند. پایین: نمودار اجمالی، که در آن استوانه‌ها نمایانگر مارپیچها و فلشها نمایانگر نوارها هستند. با تغییر خطوطی که از پشت اشیاء صلب می‌گذرند به خطوط شکسته، این اشیاء به‌صورت «نیمه‌شفاف» بازنمایی شده‌اند. برای برهم‌نهی بازنماییهای مجاور، صفحه را ۹۰° بچرخانید و به صورت سه‌بعدی نگاه کنید (ولی نه خیلی طولانی!).

همهٔ ژنها را در اختیار ما می‌گذارد. تنها برای اقلیت کوچکی از این ژنها ساختار سه‌بعدی پروتئینهای متناظر را در اختیار داریم.

## تحلیل ساختار پروتئینها

اولین مشکل تحلیل ساختار مولکولهایی به پیچیدگیِ پروتئینها، مشکل بازنمایی است. تکنیکهای گرافیک کامپیوتری برای رسم بازنماییهایی ساده‌شده‌ای از پروتئینها ارائه شده‌اند. شکل ۳ روشن می‌کند که برای یک مولکول کوچک پروتئین، تعبیر یک بازنمایی دقیق با جزئیات کامل چقدر دشوار است، و نوع تصاویر ساده‌شده‌ای را که برنامه‌ها تولید می‌کنند تا امکان دستیابی بصری به‌موضوع را به ما بدهند نشان می‌دهد. آزمایشگاههای کوچک متعدد، تعداد زیادی بازنمایی متفاوت تولید کرده‌اند؛ یعنی افراد بسیاری بازنماییهای سادهٔ‌شدهٔ مختلفی را پیشنهاد کرده‌اند، و این پیشنهادات در بسته‌های گرافیکی عامی جمع‌آوری شده است. یک ترسیمگر مولکول ماهر آنها را با هم ترکیب خواهد کرد تا جنبه‌های مختلف یک ساختار را با درجات دقت قابل تنظیم نمایش دهد. چنین تصاویری، که به‌صورت تمام رنگی درآمده‌اند و در آنها از جلوه‌های سایه‌زنی فانتزی (اما غیرواقعی، با توجه به اندازهٔ مولکولها نسبت به طول موج نور مرئی) استفاده شده است، مجلات، پوسترها، و حتی لباسها و فنجانها را می‌آرایند.

ما اکنون ساختار ۱۰۰۰۰ پروتئین را می‌شناسیم، و طیف وسیعی از الگوهای فضایی را در آنها مشاهده می‌کنیم. در پاسخ به حرف راترفورد که معتقد است «همهٔ علم یا فیزیک است یا جمع کردن تمبر»، من می‌گویم که مطالعهٔ ساختار پروتئینها بهترین جنبه‌های این دو رشته را درهم ادغام کرده است! ما با تنوع چشمگیری مواجهیم، اما در عین حال به وجود اصول جامع در پس‌زمینه ایمان داریم.

هر پروتئین از یک زنجیرهٔ اصلی پلیمر تکرارشوندهٔ خطی (یعنی بدون شاخه) تشکیل شده است که شاخه‌های جنبی آمینواسید در فواصل منظم به آن آویخته‌اند. پروتئین با رشته‌ای از چراغهای درخت کریسمس قابل مقایسه است، که سیم آن با زنجیرهٔ اصلی تکرارشونده و رشته‌های رنگی چراغها با شاخه‌های جنبی متنوعْ متناظرند.

زنجیرهٔ اصلیْ یک خم فضایی را توصیف می‌کند که به‌وسیلهٔ تعاملهای مطلوب میان زنجیره‌های جنبی که به هم متصل شده‌اند پایدار شده است. چنین خم فضایی را در بخش میانی شکل ۳ به‌راحتی می‌توان دید. دو منطقه در جلو تصویر فرم مارپیچی دارند که محورشان تقریباً عمودی است. این یکی از دو ساختار استانداردی است که مناطق محلی زنجیره به خود می‌گیرند. ساختار استاندارد دیگر، نوارِ تقریباً بازشده است: پروتئین شکل ۳ چهار رشته نوار در خود دارد که در راستای تقریباً عمودی قرار دارند. این نوارها از پهلو بر هم اثر می‌گذارند تا اجتماع خود را پایدار کنند. در چارچوب پایینی شکل ۳، مارپیچها و نوارها، با «نماد»هایی نشان داده شده‌اند، مارپیچها با استوانه و نوارها با فلشهای بزرگ. چارچوب بالایی شکل ۳ بازنمایی ساختار را با بیشترین جزئیات، شامل زنجیرهٔ اصلی و شاخه‌های جنبی نشان می‌دهد؛ تضاد رنگها نشان‌دهندهٔ اهمیت ساده‌سازی در تولید تصویر حتی از یک پروتئین کوچک است، به‌نحوی که تصویر به لحاظ بصری قابل فهم باشد.

گام اولیهٔ تجزیهٔ یک ساختار جدید عبارت است از تعیین مناطق مارپیچ و نوار. این اطلاعاتی است که برای تبدیل بازنمایی چارچوب مرکزی شکل ۳

به چارچوب پایینی مورد نیاز است. متداولترین نوع مارپیچ در ساختارهای پروتئینی، در هر پیچ ۳٫۶ سازه را دربرمی‌گیرد. خصوصیاتی از توالی که این تاوب را نشان می‌دهند، مناطق مارپیچی را تداعی می‌کنند.

## برهم‌نهی ساختارها

مانند توالیها، مسألهٔ اساسی در تحلیل ساختارها هم طرح و محاسبهٔ یک شاخص مشابهت است. فرض کنید که مجموعه‌هایی مختصاتی در اختیار داریم که دو ساختار را بازنمایی می‌کنند:

$$p_i = (x_i, y_i, z_i), \qquad i = ۱, \dots N$$

و

$$q_j = (x'_j, y'_j, z'_j), \qquad j = ۱, \dots M$$

درست مانند توالیها، در اینجا هم مسألهٔ هم‌ردیف‌سازی مطرح می‌شود. تقابل بین سه مسألهٔ مرتبط را که در شیمی محاسباتی مطرح‌اند در نظر بگیرید.

(۱) شاخص مشابهت دو مجموعه از اتمها با تناظرهای داده‌شدهٔ

$$p_i \leftrightarrow q_i \qquad i = ۱, \dots, N$$

را تعیین کنید (همتای این شاخص برای توالیها، فاصلهٔ همینگ است.) این مسأله را می‌توان دقیقاً و به نحو کارآیی حل کرد ــ این دومین «نوسانگر همساز» زیست‌شناسی مولکولی محاسباتی است.

(۲) شاخص مشابهت دو مجموعه از اتمها با تناظر نامعلوم را تعیین کنید، اما با این فرض که ساختار مولکولی آنها ــ به‌طور خاص، ترتیب خطی سازه‌ها ــ تناظر را محدود می‌کند. در مورد پروتئینها، هم‌ردیف‌سازی باید ترتیب را در طول زنجیره حفظ کند:

$$p_{i(k)} \leftrightarrow q_{j(k)}, \qquad k = ۱, \dots, K \leq N, M$$

با این قید که $k_۱ > k_۲ \Rightarrow i(k_۱) > i(k_۲)$ و $j(k_۱) > j(k_۲)$. این شاخص را می‌توان متناظر با فاصلهٔ لونشتاین بین رشته‌های حرفی، یا هم‌ردیف‌سازی توالیها با فواصل خالی در نظر گرفت.

(۳) شاخص مشابهت بین دو مجموعه از اتمها با تناظرهای نامعلوم، بدون هیچ محدودیتی بر تناظر، را تعیین کنید:

$$p_{i(k)} \leftrightarrow q_{j(k)}, \qquad k = ۱, \dots, K \leq N, M$$

این مسأله در حالت مهم زیر پیش می‌آید: فرض کنید دو (یا چند) مولکول تأثیرات زیست‌شناختی مشابه دارند، مثلاً دارای خواص داروشناختی یکسان هستند؛ این حالت معمولاً وقتی پیش می‌آید که ساختارها در زیرمجموعه‌ای نسبتاً کوچک از اتمهایشان که مسؤول فعالیت زیستی معینی است، و یک فارماکوفور نامیده می‌شود، مشترک باشند. برای تشخیص فارماکوفور، خوب است بتوانیم زیرمجموعه‌های ماکسیمال اتمهای دو یا چند مولکول را که ساختار مشابه دارند بیابیم.

مسائل (۲) و (۳) نیازمند تعیین نحوهٔ هم‌ردیف‌سازی نقاط هستند. روشهای هم‌ردیف‌سازی که منحصراً بر مختصات (نه بر توالیهای آمینواسید) مبتنی هستند، هم‌ردیف‌سازی ساختاری نامیده می‌شوند. در هم‌ردیف‌سازیهای ساختاری، سازه‌های متناظر یکی‌گرفته می‌شوند زیرا آنها نسبت به کل ساختار مکان مشابهی اختیار می‌کنند. باید به استخراج زیرساختار مشترک ماکسیمال و بنا نهادن هم‌ردیف‌سازی بر آن اندیشید. (برای مثال، زیرساختار مشترک ماکسیمال حروف B و R، حرف P است.) سازه‌های خارج از زیرساختار مشترک ماکسیمال قابل هم‌ردیف شدن نیستند، حقیقتی که با هم‌ردیف‌سازی جفت‌جفتِ توالیها مشخص نمی‌شود؛ این یکی از ضعفهای آن است.

کلیترین رویکرد به این سه مسأله بر حل مسألهٔ (۱)، یعنی حالت تناظر معلوم $p_i \leftrightarrow q_i$ مبتنی است. دو شیء یکسان را می‌توان با انتقال و دوران صلب یکی از آنها، برهم‌نهاد. دو شیء که مشابه هستند از طریق دوران و انتقال به برهم‌نهی تقریبی می‌رسند. اگر اشیاء مجموعه‌های مرتب نقاط باشند، شاخص مشابهت آنها برابر با جذر میانگین مربع انحرافها، $\Delta$، پس از برهم‌نهی بهینه خواهد بود:

$$\Delta^۲ = \min_{\mathbf{R},\mathbf{t}} \left\{ \sum_{i=۱}^{N} \| \mathbf{R} p_i + \mathbf{t} - q_i \|^۲ \right\}$$

که در آن $\mathbf{R}$ ماتریس دوران مناسب ($\det\mathbf{R} = ۱$) و $\mathbf{t}$ بردار انتقال است. در برهم‌نهی بهینه، مکانهای میانی (به‌زبان محاوره، «مراکز ثقل») دو مجموعه برهم منطبق می‌شوند. مسألهٔ تعیین جهت نسبی صحیح به‌عنوان «مسأله پروکروتس متعامد» شناخته شده است و راه‌حلهایی مبتنی بر روشهای استاندارد جبر خطی برای آن وجود دارد [۵].

حل مسألهٔ زیرساختار مشترک ماکسیمال، مبنای تعریف یک متریک را برای ساختارها فراهم می‌کند. بر این مبنا می‌توان مشابهتهای مقطعی و جزئی را پیدا کرد، و یک درخت طبقه‌بندی برای کل مجموعهٔ ساختارهای پروتئین ارائه داد.

رویکردهای موجود به محاسبهٔ زیرساختار مشترک ماکسیمال بر دو بازنمایی از ساختارها متکی بوده‌اند: (۱) به‌صورت فهرستهایی از مختصات $p_i = (x_i, y_i, z_i)$، $i = ۱, \dots, n$، یا (۲) به‌صورت ماتریسهای فاصلهٔ $D_p(i,j) = |p_i - p_j|$. مزیت عمدهٔ ماتریسهای فاصله این است که یک بازنمایی مستقل از مبدأ و راستا برای ساختار ارائه می‌کنند. برحسب ماتریسهای فاصله، مؤلفهٔ ماکسیمال تفاضل ماتریسهای فاصله $\max_{i,j}\{|D_p(i,j) - D_q(i,j)|\}$، **شاخصی برای اختلاف ساختاری بین دو مجموعهٔ نقطهٔ هم‌ردیف‌شده به‌دست می‌دهد.**

مختصات و ماتریسهای فاصله، بازنماییهایی تقریباً معادل از یک مجموعهٔ نقاط هستند. محاسبهٔ ماتریس فاصله از روی مختصات بدیهی است. اینکه آیا می‌توان مختصات را دقیقاً و مستقیماً از روی ماتریس فاصله بازسازی کرد کمتر واضح است، اما این‌کار را می‌توان به‌کمک قطری‌سازی ماتریس انجام داد [۶]. البته، ماتریس فاصله هم ساختار اولیه و هم تصویر آینه‌ای آن را به‌طور معادل مشخص می‌کند (پس دسته‌های چپ و راست متناظر، دو تصویر آینه‌ای هستند)، اما این ابهام مشکلی جدی برای کاربردهای زیست‌شناسی مولکولی نیست. اطلاعات مربوط به موقعیت و جهت هم مسلماً مفقود می‌شوند.

دشواری اصلی در محاسبات زیرساختار مشترک ماکسیمال از نوع (۲) و (۳)، پیچیدگی ترکیباتی ناشی از در نظرگرفتن همهٔ هم‌ردیف‌سازیهای ممکن است. معلوم شده است که الگوریتمهای مبتنی بر ماتریسهای فاصله در مواجهه با این مسأله بسیار کارآتر از الگوریتمهای مبتنی بر مجموعه‌های مختصات عمل می‌کنند. بازنماییهای ماتریسی مرتبط که به‌جای مختصات اتمی، بر مؤلفه‌های ساختاری نظیر مارپیچها و نوارها مبتنی هستند، بازنمایی فشرده‌ای از الگوهای تاخوردگی پروتئینها به‌دست می‌دهند. استخراج زیرماتریسهای مشترک ماکسیمال، بزرگترین زیرساختارهای دارای الگوی تاخوردگی مشترک را آشکار می‌کند. به‌علاوه چنین بازنماییهایی، امکان شمارش **همهٔ** الگوهای تاخوردگی پروتئین را در اختیار ما می‌گذارند. به‌صورت تجربی برآورد شده است که همهٔ پروتئینهای طبیعی کمتر از حدود ۱۰۰۰ الگوی تاخوردگی دارند. شمارش کامل به ما امکان می‌دهد که انتخابهای طبیعت را آزمایش کنیم، و سعی کنیم تصادف تاریخی و ضرورت معماری را از هم تمیز دهیم.

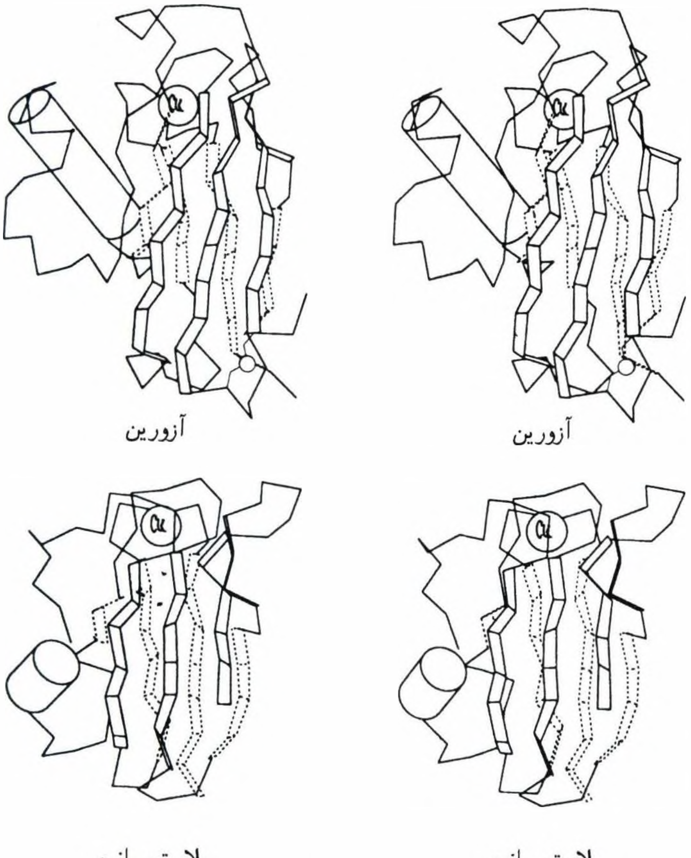
## تکامل پروتئین

مطالعهٔ تکامل پروتئین یعنی بررسی اینکه چگونه توالیهای آمینواسید و ساختارهای پروتئین متناظر در گونه‌های وابسته با هم اختلاف پیدا می‌کنند. این تحقیقی از نوع اطلاع‌دهنده است، که به ما در فهم روابط توالی-ساختار کمک می‌کند. زیر اگرچه ما می‌دانیم که یک توالی آمینواسید منفرد همهٔ اطلاعات لازم برای مشخص کردن ساختار پروتئین را در خود دارد، هنوز نمی‌فهمیم که چگونه باید از توالی به ساختار رسید. این را «صورت انتگرالی» مسألهٔ تاخوردگی پروتئین در نظر بگیرید، که مسألهٔ حل نشده‌ای است. در بررسی تکامل پروتئین، مشاهده می‌کنیم که چگونه تغییرات توالیها در تغییرات ساختار منعکس می‌شوند؛ این را آسانتر می‌توان فهمید. این مسأله را «صورت دیفرانسیلی» مسأله تاخوردگی پروتئین در نظر بگیرید.

| موضوع: | تاخوردگی پروتئین | تکامل پروتئین |
|---|---|---|
| مشاهده: | توالی ← ساختار | تغییر در توالی ← تغییر در ساختار |
| صورت مسأله: | «صورت انتگرالی» | «صورت دیفرانسیلی» |
| وضعیت مسأله: | حل‌نشده | حل‌نشده اما باید ساده‌تر باشد |

استدلال ساده‌ای بر این دلالت می‌کند که ساختار باید تابعی تقریباً «پیوسته» از توالی باشد، حداقل برای توالیها و ساختارهایی که به‌صورت طبیعی شکل گرفته‌اند. فرض کنید پروتئینی وجود می‌داشت که در آن هر جهشی (هر تغییری در توالی آمینواسید) یک ساختار ناپایدار تولید می‌کرد. در این صورت طبیعت هرگز نمی‌توانسته است با فرآیندهای تکاملی به چنین ساختاری برسد، زیرا هیچ صورت اولیهٔ پایداری نمی‌توانسته برای آن وجود داشته باشد. پس چنین نتیجه می‌شود که ساختارهای طبیعی باید مستحکم باشند. اغلب تغییرات کوچک در توالی باید تغییری در ساختار ندهند. (این شرط برای ساختارهای پروتئینی که به‌روشهای مصنوعی ساخته می‌شوند برقرار نیست.)

درواقع، پروتئینهای طبیعی با توالیهای بسیار شبیه به‌هم، ساختارهای بسیار شبیه به‌هم دارند. پیش از آنکه انسولین مصنوعی انسانی در دسترس قرارگیرد، انسولین خوک درمان بالینی مؤثری برای انسانهای مبتلا به مرض قند بود، با




**شکل ۴** در حین تکامل، جهشهای توالیهای ژن برروی هم انباشته می‌شوند، و در نتیجه توالیها و ساختارهای پروتئین از هم فاصله می‌گیرند. این شکل دو پروتئین مرتبط انتقال‌دهندهٔ الکترون، یعنی پلاستوسیانین برگ سپیدار و آزورین باکتریایی را نشان می‌دهد. بخشی از ساختارها که در نیمهٔ راست تصویر واقع است و شامل آرایه‌های پر و نقطه‌چین مناطق «روبانی‌شکل» ــ که نوار نامیده می‌شوند ــ و محل پیوند با مس است، به‌خوبی در حین تکامل ماندگار است، در حالی‌که بخشی از ساختارکه در نیمهٔ چپ تصویر واقع است به‌شدت تغییر کرده است.

آنکه توالیهای آمینواسید انسولین خوک و انسان یکسان نیستند. اعتماد به چنین مشابهتهایی روشی برای پیش‌بینی ساختار پروتئینها از روی ساختارهای بسیار نزدیکِ شناخته‌شده به‌دست می‌دهد، که به‌عنوان «مدلسازی مانستگی»[1] شناخته شده است. اما، همچنان‌که تکامل ادامه می‌یابد، توالیها و ساختارها نهایتاً به شکل بنیادیتری واگرا می‌شوند. شکل ۴ دو پروتئین پلاستوسیانین و آزورین را نشان می‌دهد که با هم ارتباط کمی دارند، در منطقهٔ راست شکل، دو نوار وجود دارد که روبه‌رو بسته شده‌اند، و «هستهٔ» ماندگار ساختار را تشکیل می‌دهند، در حالی‌که منطقهٔ حلقوی بلند سمت چپ از شکلی کاملاً متفاوت در دو ساختار برخوردار است.

## پیش‌بینی ساختار پروتئین

طبیعت الگوریتمی در اختیار دارد که به کمک آن ساختار سه‌بعدی پروتئین را فقط از روی توالی آمینواسید آن تعیین می‌کند. ما باید قادر به کشف آن

1. homology madelling

باشیم. سپس باید بتوانیم ساختار پروتئینها را که در توالیهای ژن انسان و سایر ژنومها به‌صورت ذاتی قرار دارد پیش‌بینی کنیم، و آنها را در مسائل عملی نظیر طراحی دارو به‌کار گیریم. معلوم شده است که مسألهٔ پیش‌بینی ساختار پروتئین مسألهٔ بسیار دشواری است. رویکردهای بسیاری اختیار شده‌اند، و ادعاهای بسیاری مطرح شده‌اند. اما در حال حاضر هیچ روش محاسباتی که بتواند به‌صورت سازگار حتی یک پیش‌گویی صحیح کیفی از ساختار پروتئین براساس توالی آمینواسید ارائه کند وجود ندارد، مگر آنکه یک پروتئین بسیار مشابه موجود باشد.

فرض کنید که توالی آمینواسیدهای یک پروتئین جدید به شما داده شده بود، و از شما خواسته بودند که ساختار آن را پیش‌بینی کنید. شما باید سعی می‌کردید چه چیزی را پیش‌بینی کنید؟ کاملترین اطلاعاتی که یک پیش‌بینی ممکن است به‌دست دهد مجموعهٔ کاملی از مختصات سه‌بعدی مدل یک پروتئین نهایی است ــ یعنی یک پیش‌بینی سه‌بعدی. یک هدف کمتر بلندپروازانه می‌تواند پیش‌بینی این باشد که مناطق مارپیچ و نوار در کجای توالی پدیدار می‌شوند ــ یعنی یک پیش‌بینی یک‌بعدی. جایی بین این‌دو، پیش‌بینی‌هایی قرار دارند که از پیش‌بینیهای ساختار ثانویهٔ یک‌بعدی فراتر می‌روند، اما تنها برخی اطلاعات کیفی در بارهٔ آرایش فضایی کلی الگوی تاخوردگی ارائه می‌کنند ــ بیایید اینها را پیش‌بینیهای دوبعدی بخوانیم.

از چه نوع اطلاعاتی می‌توان در پیش‌بینی ساختار پروتئین استفاده کرد؟ هدف نهایی، رویکردِ ابتدا به ساکن[1] «محض» است ــ تنها از توالی پروتئین مقصد استفاده کنید و نه هیچ چیز دیگر. باید به‌خاطر داشت که این همان کاری است که طبیعت می‌کند ــ پروتئینها وقتی می‌خواهند تا شوند در پایگاههای دادهٔ وب جستجو نمی‌کنند. اما ما می‌توانیم جستجو کنیم، و موفقیتهایی هم در استفاده از اطلاعات بانکهای داده برای شناسایی تاخوردگی یک پروتئین مقصد از روی ساختارهای شناخته‌شده حاصل شده است. این مسأله به نام تشخیص تاخوردگی معروف شده است. البته این روش تنها در صورتی مؤثر است که ساختار یک یا چند پروتئین با تاخوردگی مشابه پروتئین مقصد در پایگاه دادهٔ شما موجود باشد.

چه کسی را باید متقاعد کرد؟ فهرست زیر به‌گونه‌ای مرتب شده است که تقریباً از بالا به پایین سختگیری کم می‌شود. اغلب دانشمندان پذیرفته‌اند که اقناع «حامیان مالی طرح» از همه مهمتر است!

چه‌کسی را باید متقاعد کرد؟

۱. بلورشناسان
۲. متخصصان طیف‌شناسی NMR
۳. حامیان مالی طرح
۴. داوران مقاله‌ها
۵. همکاران
۶. مادرتان

1. ab initio

درواقع شما چگونه می‌توانید مردم را متقاعد کنید که روش موفقی برای پیش‌بینی ساختار پروتئین در اختیار دارید؟ دو نوع از ادعاها از اساس غیرقابل آزمون هستند. یکی آنکه شما می‌توانید ساختار پروتئینی را که ساختار آن از قبل دانسته شده است پیش‌بینی کنید. دیگر آنکه شما ساختار پروتئینی را پیش‌بینی کرده‌اید که ساختار تجربی آن ناشناخته است و به احتمال زیاد تا مدت طولانی ناشناخته خواهد ماند. باید در حوزهٔ بینابینی میان شناخته‌شده‌ها و آنچه تا زمان طولانی غیرقابل شناسایی است کار کرد، و پیش‌بینی ساختار را با تعیین ساختارهای در دست اقدام هماهنگ نمود.

برای نظم بخشیدن به این فعالیت، برای پاداش دادن به کسانی که موجب پیشرفتهای اصیل شده‌اند، و برای تکذیب ادعاهای آنانی که مصراً ادعا می‌کردند «مسأله پیش‌بینی ساختار پروتئین را حل کرده‌اند»، جان مالت[1] ایدهٔ آزمونهای کورِ سازمان‌یافته را مطرح کرد. ایده این است که دانشمندان در حین فرایند کشف ساختارها، توالیهای آمینواسید را به همه اعلام کنند، اما قول بدهند که ساختار را تا زمان مورد توافقی مخفی نگه دارند. همهٔ کسانی که باور دارند که روشی برای پیش‌بینی ساختار پروتئین در اختیار دارند می‌توانند پیش‌بینیهای خود را تا قبل از تاریخ اعلام ساختار ارائه دهند. پس از اعلام، می‌توان پیش‌بینیها را با تجربه مقایسه کرد ــ که این باعث خرسندی عدهٔ اندکی و تأسف گروه کثیری خواهد شد. این ایده تحت برنامهٔ CASP ــ برآورد انتقادی پیش‌بینی ساختار[2] ــ براساس یک دورهٔ دوساله ارائه شده است.

روشهای پیش‌بینی به دو طبقهٔ کلی تقسیم می‌شوند، استقرایی و استنتاجی. در روشهای استقرایی مستقیماً از بانکهای دادهٔ توالیها و ساختارها استفاده می‌شود. روشهای استنتاجی، رویکردهای واقعاً ابتدا به ساکن هستند ــ حالت جزیرهٔ بی‌آب و علف ــ که هدف آنها پیش‌بینی ساختار پروتئین براساس اصول عام فیزیک، شیمی و زیست‌شناسی، بدون ارجاع صریح به توالیها و ساختارهای شناخته‌شده است. البته پیشرفت روشهای ابتدا به ساکن بستگی به چیزهایی دارد که از بررسی توالیها و ساختارها آموخته‌ایم. تمایز این روشها در این است که درک حاصل از این بررسی به شکل اصول عامی خلاصه شده است که می‌توان آنها را بدون مراجعه به اطلاعات خاص در پایگاههای داده، به‌کار برد.

روشهایِ پیش‌بینی ابتدا به ساکن را می‌توان به دو رویکرد تقسیم کرد، که من آنها را «طبیعی» و «زوری» می‌نامم. رویکرد طبیعی به دنبال درک فرآیند تاخوردگی طبیعی می‌گردد و سپس سعی می‌کند آن‌را شبیه‌سازی کند. رویکرد «زوری» هر فرایندی را که بتواند زنجیره را به شکل [=ترکیب، ساختمان] مناسب برساند مجاز می‌داند، حتی اگر این فرایند در راستای مسیری انجام شود که طبیعی نباشد یا حتی از نظر فیزیکی غیرقابل تحقق باشد.

شواهدی هست که انتخاب طبیعی نه تنها وضعیت طبیعی نهایی پروتئینها، بلکه مسیر تاخودگی آنها را نیز شکل داده است. زیرا نه تنها پروتئینها باید به گونه‌ای تکامل‌یافته باشند که شکل فعال پایداری پیدا کرده باشند، بلکه باید در زمان معقولی، از یک وضعیت تانخوردهٔ اولیه شامل مخلوطی از شکلهای تصادفی به چنین شکلی دست یافته باشند. یک محاسبهٔ ساده براساس

1. John Moult 2. Critical Assessment of Structure Prediction

میزان جهشهای اتمی در محلول نشان می‌دهد که سرعت پیمایش فرسایندهٔ فضای شکلهای ممکن، چند مرتبهٔ بزرگی کمتر از اندازهٔ کافی خواهد بود. (این مطلب گاهی پارادوکس لوینتال[1] نامیده می‌شود.) هیچ مدرکی وجود ندارد که نشان دهد مسیر تاخوردگی واقعاً بر وضعیت نهایی تأثیر می‌گذارد، اگرچه به لحاظ نظری این امر ممکن است. اگر حالات تاخوردگی دیگری ممکن بودند، اما مسیر چنان پیش می‌رفت که به یکی از آنها منجر شود، آنگاه ما پیش‌گویان مجبور بودیم رویکرد طبیعی، نه زوری، را بپذیریم.

مانع پیش‌بینی ساختار کجاست؟ ما فکر می‌کنیم نیروهایی را که شکلهای طبیعی پروتئین را پایدار می‌کنند می‌شناسیم. حتی ممکن است بتوانیم تابع انرژی شکل‌گیری صریحی از مختصات بنویسیم. کاری که باید انجام دهیم کمینه کردن آن است. اما این مهم است که تشخیص دهیم که پروتئینها، به بیان ترمودینامیکی، تنها در حاشیه پایدارند. درواقع انرژی شکل‌گیری یک پروتئین تاخورده برابر با اختلاف بسیار کوچکی بین عبارات متضاد بزرگ است، چیزی که برای تحلیلگران عددی کابوس است.

آیا مشکل این است که نمی‌توانیم تابع انرژی را به‌اندازهٔ کافی دقیق بنویسیم، یا تابع چنان پیچیده است که نمی‌توانیم بهینه‌اش کنیم؟ یک آزمون، کمینه کردن انرژیهای شکل‌گیری پروتئینهاست، با شروع از حالتهای طبیعی شناخته‌شدهٔ آنها. چنین محاسباتی به سمت شکلهای با کمترین انرژی نزدیک به نقطهٔ شروع همگرا می‌شوند، و نشان می‌دهند که توابع انرژی در همسایگی پاسخ صحیح رضایت‌بخش هستند. (این مطلب چندان شگفت‌انگیز نیست، زیرا توابع براساس تنظیم پارامترها تعریف شده‌اند به‌گونه‌ای که حالات طبیعی مشاهده شده را باز تولید کنند.) اما این کافی نیست. تابعی که در همسایگی نقطهٔ مینیمم صحیح است الزاماً مجموعهٔ کاملی از مسیرهای پیموده‌شده در فضای شکلها را ــ که یک برنامه را قادر می‌سازند تا با شروع از یک نقطهٔ دلخواه مینیمم سراسری را بیابد ــ در اختیار ما قرار نمی‌دهد.

دو مسأله وجود دارد. نخست اینکه بسیاری از نیروهایی که پروتئینها را پایدار می‌کنند کوتاه‌برد هستند. حتی اگر تابع انرژی را دقیقاً می‌دانستیم، و اگر کمینه‌سازی را از یک شکلِ به‌طور تصادفی گسترش‌یافتهٔ نافشرده شروع می‌کردیم، ممکن بود به این نتیجه برسیم که هیچ نیروی بلندبردی وجود ندارد که سیستم را به سمت ساختار صحیح براند. دوم اینکه، حتی اگر با فروریزی سیستم به یک حالت فشرده برسیم، نمای انرژی به‌عنوان تابعی از مختصات چندین مینیمم موضعی را شامل می‌شود که توسط حایلهای بلند از هم جدا شده‌اند. بسیاری از این مینیممهای موضعی کاندیداهایی برای حالت طبیعی خواهند بود. پروتئینهای حقیقی به‌وسیلهٔ ترکیبی از (۱) «پردازش موازی» سنگین که در آن همهٔ سازه‌ها همزمان ابعاد موضعی خود را در فضای شکل می‌کاوند، و (۲) تکامل مسیرهای تاخوردگی که سیستم را به سمت جواب صحیح هدایت می‌کنند، بر این مسائل غلبه کرده‌اند. کامپیوترهای ما نمی‌توانند به پردازش موازی دست یابند، توابع انرژی ما نمی‌توانند مسیرهای تاخوردگی بلندبرد را توجیه کنند، و الگوریتمهای ما نمی‌توانند به‌سادگی مینیمم سراسری یک تابع پیچیدهٔ چندمتغیرهٔ غیرخطی را بیابند. (این نوسانگر همساز، وقتی لازمش داریم کجاست؟!)

دشواری حالت پیشینی، همان‌طور که متوجه شده‌ایم، به پیدایش روشهای تجربی مبتنی بر توالیها و ساختارهای شناخته‌شده منجر شده است. روشهای پیش‌بینی که از بانکهای داده استفاده می‌کنند عبارت‌اند از (۱) روشهایی برای مدلسازی مانستگی ــ پیش‌بینی ساختار مقصد از روی یک پروتئین خیلی مشابه که ساختار آن کاملاً شناخته شده است؛ و (۲) روشهای تشخیص تاخوردگی ــ تخمین‌زدن تطابق‌پذیری توالی آمینواسید با مجموعهٔ الگوهای شناخته‌شدهٔ تاخوردگی پروتئین. این روشها تا حدی (اما نه کاملاً) به‌دلیل رشد بانکهای داده کاملتر شده‌اند. هر چه توالیها و ساختارهای بیشتری شناخته شوند، احتمال آنکه یک پروتئین جدید مشابه پروتئینی باشد که قبلاً شناخته شده است بیشتر می‌شود. برعکس، روشهای ابتدا به ساکن کندتر پیش می‌روند. بعد از یکی از رقابتهای اخیر CASP، یکی از اظهارنظرهای توأم با دلخوری در بارهٔ این روشها این بود که حداقل «شکست از این پس تضمین‌شده نیست [۸].» فرد بدبین ممکن است پیش‌بینی کند که رشد بانکهای داده به این معنی خواهد بود که روشهای مبتنی بر اطلاعات راه‌حلهای عمل‌گرایانه‌ای برای چنان اکثریت بزرگی از سؤالات ارائه خواهند کرد، که اشتیاق نسبت به حمایت از توسعهٔ روشهای ابتدا به ساکن نقصان خواهد یافت. بنابراین مایه شرمساری خواهد بود که زیست‌شناسی محاسباتی، یکی از جالبترین محاسبات زیست‌شناختی را از دست داده باشد!

## مراجع


1. Wigne, E.P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13, 1-14.
2. Sankoff, D. and Kruskal, J.B., eds. (1983). *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison.* Addison-Wesley, Reading, Mass.
3. Needleman, S.B. and Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48, 443-453.
4. Fitch, W.M. and Smith, T.F. (1983). Optimal sequence alignments. *Proc. Natl. Acad. Sci. USA* 80, 1382-1386.
5. Golub, G. and van Loan, C., *Matrix Computations.* Johns Hopkins Press. Baltimore, 2nd ed. 1989.
6. Young, G. and Householder, A.S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika* 3, 19-22. (For background and history see [7].)
7. Blumenthal, L.M. (1938). *Distance Geometries. A study of the development of abstract metrics.* University of Missouri Studies 13, #2.
8. *New York Times*, March 25, 1997.


* * * * * *


- Arthur M. Lesk, "The unreasonable effectiveness of mathematics in molecular biology", *Math. Intelligencer*, (2), **22** (2000) 28-37.



* آرتر لسک، دانشگاه کیمبریج، انگستان


1. Levinthal

# سیر پیدایش دقت در احتمال ریاضی (۱۹۰۰ـ۱۹۵۰)•

جوزف دوب*
ترجمهٔ عطاءالله تقاء

## ۱. مقدمه

این نوشتارِ کوتاه غیررسمیْ شمّه‌ای است از تاریخچهٔ ظهور دقت در احتمال ریاضی طی نیمهٔ نخست قرن بیستم. نتایج و حکمهای مشخص تنها در صورتی ذکر شده‌اند که در تاریخ پیشرفت منطقی احتمال ریاضی حائز اهمیت بوده‌اند.

نشو و نمای علوم همچون دنباله‌ای از گامهایی که به جلو برداشته می‌شوند نیست. از پسِ واقعه که نظر کنیم، پیشرفتِ کار بطیء بوده، مسیر پرپیچ و خمی را پشت سرنهاده، فراوان کج رفته و سر از بن‌بست‌ها در آورده، و چه‌بسا در جهاتی راه پیموده که قابل قبول دانشوران پیشرو دوران نبوده است. در دههٔ ۱۹۳۰ فضاهای باناخ را به دیدهٔ تمسخر می‌نگریستند چون آنها را مجرداتی مزخرف می‌دانستند، بعد نوبت فضاهایِ کوژِ موضعی فرا رسید و اکنون هم نوبت آنالیز نااستانده است. ریاضیدانان هم در پذیرش ایده‌های جدید شائقتر از دیگر آدمیان نیستند و پذیرشِ تمام و کمال احتمال ریاضی تا نیمهٔ دوم قرن تحقق نپذیرفت. علی‌الخصوص بسیاری از آماردانان و احتمال‌دانان قالب‌بندیِ ریاضیِ احتمال به‌واسطهٔ نظریهٔ اندازه را خوش نمی‌داشتند، و برخی هنوز احتمال ریاضی را بیرونِ آنالیز می‌نهند. اقوال ذیل مؤید این باورند:

**پلانک**: حقیقت علمی جدید این‌گونه غالب نمی‌شود که رقبایش را متقاعد کند و راه راست را به آنان بنماید، بلکه اغلب بدین شیوه غلبه می‌یابد که رقبا سرانجام از میان رفته و صحنه را خالی می‌کنند و نسل جدیدی با آن [شیوهٔ نگرش جدید] بار می‌آید.

**پوانکاره**: سابق بر این موقعی که یک نفر تابع جدیدی ابداع می‌کرد، برای این منظور بود که هدفی عملی را به پیش برد؛ امروزه روز، اما، توابعی ابداع می‌کنند تا اثباتهای آبائمان را ابطال کنند، و از این ابداعات هیچ چیز دیگری هم عاید نمی‌شود.

**اِرمیت**: (در نامه‌ای به اِستیلتیس) من از بلای اسفبار توابعی که مشتق ندارند با نگرانی و وحشت به خود می‌پیچم.

نظریهٔ احتمال در ابتدا، و هم پس از آن برای مدتی طولانی، عبارت بود از صورت آرمانی و تحلیل برخی پدیده‌های زندگی واقعی در خارج از حیطهٔ ریاضیات، اما اندک‌اندک در نیمهٔ نخست این قرن احتمال ریاضی بخشی معمولی از ریاضیات شد. ریاضی‌سازیِ احتمال ایده‌های نو می‌طلبید، و به‌ویژه نیازمند رهیافت جدیدی به ایدهٔ پذیرفتنی بودنِ یک تابع بود. با توجه به نقل‌قول‌های فوق شگفتی ندارد که پذیرش این ریاضی‌سازی آهسته صورت گرفت و با مقاومت روبه‌روگردید. درواقع حتی اینک برخی احتمال‌دانان از آن بیم دارند که قالب‌بندیِ ریاضیِ احتمال جاذبهٔ ذاتی‌اش را از آن ستانده باشد، والبته حق با ایشان است، به این تعبیرکه آن جاذبه و افسونِ احتمال‌ـریاضیاتِ مبهم و قدیمی که مبتنی بر تعاریف غیرریاضی بود، دو شقه شده: افسونِ احتمال جهانِ واقع و جاذبهٔ دقت ریاضی. ولیکن باید تأکید کرد که بسیاری از اساسیترین نتایج مبحث احتمال ریاضی از سرچشمهٔ غیرریاضیِ احتمال جهانِ واقع جاری می‌شوند که هرگز حتی تعریفی نداشته است که همگان در آن متفق باشند. در حقیقت، ارتباط مابین احتمال جهان واقع و احتمال ریاضی هم بلای جان و هم منبع الهامِ رشد احتمال ریاضی بوده است.

## ۲. مسألهٔ (غیرریاضی) جهان واقع چیست؟

آنچه معمولاً احتمال (جهان واقع) می‌نامیم برخاسته از زمینه‌های متعدد است. علاوه بر زمینه‌های آشنای بازیهای قمار، بیمه، و فیزیک آماری زمینه‌های ساده‌ای از این قبیل هم هستند: فرض کنید فردی سوار دوچرخه‌اش سرکار می‌رود. اگر طی ده روز متوالی، هر بارکه این فرد دوچرخه‌اش را پارک می‌کَند والوِ تیوبِ چرخ جلو در نیمهٔ فوقانیِ چرخ باشد، همان قدر مایهٔ شگفتی است که ده پرتاب پی‌درپیِ یک سکه همه به شیر بینجامند. مع‌هذا آشکار است که (در مورد چرخ) اگر مسیر پیموده‌شده خیلی کوتاه باشد، یا (در مورد سکه) اگر سکه از جایی نزدیک به مکانِ فرودش رها شود و سرعتِ گردشِ اولیهٔ آن

کم باشد، از شگفتی کاسته خواهد شد و جنبهٔ احتمالاتیِ قضیه محل تردید خواهد بود. نتیجه اینکه پیش از بیان هر گزارهٔ احتمالاتی باید زمینهٔ مربوطه را نیک بر رسید. اگر هم بحث فلسفی موضوعیت داشته باشد، که این خود جای بحث دارد، باید آن‌را با بررسی زمینهٔ فیزیکی تکمیل کرد.

## ۳. قانون اعداد بزرگ

آنچه در یک رشته آزمایش مستقل تکراری، همچون پرتاب سکه، آدمی را به یکباره خیره می‌سازد چیزی است که به قانون اعداد بزرگ موسوم گشته است. در مورد پرتاب سکه این قانون می‌گوید که به یک معنا اگر تعداد شیرآمدن‌ها در $n$ پرتاب را به $n$ تقسیم کنیم، حاصل به $\frac{1}{2}$ میل می‌کند هرگاه $n$ افزایش یابد. در اینجا کلمات اساسی عبارت‌اند از «به یك معنا». اگر قانون اعداد بزرگ یک حکم ریاضی است، یعنی اگر مدلی ریاضی برای پرتاب سکه داشته باشیم که قانون اعداد بزرگ در آن به‌عنوان قضیه‌ای ریاضی بیان شده باشد، این قضیه یا برحسب یکی از مفاهیم متنوع حدّ ریاضی درست است و یا نه. از سوی دیگر، اگر بنا باشد قانون اعداد بزرگ در یک زمینهٔ غیرریاضی جهان واقع بیان شود، اصلاً واضح نیست که بتوان مفهوم حد را به نحو معقولی فرمولبندی کرد. آشکارترین معضل آن است که در جهان واقع تنها تعدادی متناهی آزمایش می‌توان در زمانی محدود صورت داد. هر کس که می‌کوشد به متعلمان توضیح دهد که هنگامی‌که سکه‌ای به کرّات پرتاب می‌شود چه رخ می‌دهد، مذبوحانه عباراتی همچون در درازمدت، میل می‌کند، به نظر می‌رسد که نزدیك ... تجمع می‌کنند، و نظیر آنها را به‌کار می‌گیرد بلکه شاید به مفهومی شبح‌وار شکلی ببخشد. ولکن واقع امر از این قرار است که هر کس سکه‌ای را به دفعات پرتاب کند به عینه می‌بیند که بعد از دفعاتی نه چندان بسیار به نظر می‌رسد که تعداد شیرها در $n$ پرتاب تقسیم بر $n$، همچنان‌که $n$ افزایش می‌یابد به $\frac{1}{2}$ نزدیک می‌شود. ساده‌ترین راه‌خلاص، که یک آماردان بیزی برجسته برگزیده، این راه‌حل بی‌محتواست: اصلاً در باره اینکه وقتی سکه‌ای پرتاب می‌کنیم چه خواهد شد صحبت ممنوع! یک راه رایج‌تر که به همان اندازه رضایتبخش می‌باشد آن است که این مطلب که آیا بحث موردنظر در محدودهٔ ریاضیات هست یا نه در ابهام گذاشته شود. چه بسا اینکه این گزاره را قانون نامیده‌اند نشانه‌ای از همین ابهام است. در بارهٔ این قانون حرفهای زیر را زده‌اند (تأکید از نگارنده است):

**لاپلاس**: (۱۸۱۴) این قضیه، که شعور متعارف بر آن صحه می‌گذارد، با آنالیز به دشواری اثبات می‌شود.

**ویل**[1]: (۱۹۳۹) آدم دلیلی نمی‌بیند که این گزاره درست باشد؛ ولی از آنجا که اثبات خلاف آن به توسط آزمایش هم غیرممکن است، حداقل می‌شود با آسودگی‌خاطر آن‌را بیان کرد.

**باوئر**[2]: (ترجمه از زمینهٔ مربوط به تاس به زمینهٔ مربوط به سکه): این یك واقعیت است که به تجربه ثابت شده که خارج‌قسمت ... به‌ازای $n$های بزرگ انحرافی از $\frac{1}{2}$ نشان می‌دهد که به صفر نزدیك می‌شود.

این بیانات نشانگر افسونِ مانای مباحثات در بارهٔ احتمال جهان واقع‌اند. لیکن از بخت بد، ریاضیدانان بدین‌سو کشش داشته‌اند که در بارهٔ پرسش زیر تأمل نمایند. یا دست‌کم در بارهٔ آن قلمفرسایی کنند.

## ۴. احتمال چیست؟

نظرهایی را در جهت پاسخگویی به این پرسش و نیز پیرامون نحوهٔ تدریس این موضوع در زیر ملاحظه می‌کنید:

**پوانکاره**: (۱۹۱۲) ارائهٔ تعریفی رضایتبخش برای احتمال تقریباً نشدنی است.

---

1. Ville 2. Bauer

**مازورکیه‌ویچ:** (۱۹۱۵) نظریهٔ احتمال عنصر مستقلی در تعالیم ریاضی نیست؛ مع‌هذا مطلوب است که ریاضیدانان اصول کلی آن را بدانند، مفاهیم بنیادی آن به نحو جامعی معین نگردیده‌اند و شامل دشواریهای حل‌نشدهٔ بسیارند.

**فون میزس:** (۱۹۱۹) درواقع معلوم نیست که وضع فعلی مبحث احتمال از چه قرار است، جز اینکه این مبحث یک رشتهٔ ریاضیاتی نیست. (او در این راستا پیش رفت که آن را به شکل یک رشته ریاضی درآورد، به این شیوه که احتمال ریاضی را بر دنباله‌ای از مشاهدات (Beobachtungen) با خواصی که یک دنبالهٔ خوشتعریف ریاضی فاقد آن است مبتنی نمود. همچنین به شوخی گفته شده که وی احتمال را عددی بین صفر و یک تعریف کرده که در بارهٔ آن هیچ چیز دیگری معلوم نیست.)

**پیرسون:** (۱۹۳۵) (نقل از محاوره) احتمال آنچنان با آمار گره خورده که هر چند تدریس جداگانهٔ آنها مقدور است، ولی چنین برنامه‌ای، کاری است کارستان.

**اوسپنسکی:** (۱۹۳۷) او در یک کتاب درسیِ پراستفاده، تعریف زیر را که زمانی در کتابهای درسی متداول بود ارائه می‌کند: اگر، سازگار با شرطِ $S$، $n$ موردِ دوبه‌دو مجزا و همشانس داشته باشیم و $m$ تا از آن موارد مساعدِ رویداد $A$ باشند، آنگاه احتمال ریاضی $A$ را $\frac{m}{n}$ تعریف می‌کنیم.

آنچه آمد باید اهمیت تفکیک نظریهٔ احتمال ریاضی از کاربردهای آن در جهان واقع را به‌وضوح نموده باشد. ولیکن به این نکته توجه کنید که هیچ‌کس در کاربردپذیریِ احتمال ریاضی در جهان واقع تشکیک نمی‌کند. قمار، ژنتیک، بیمه، و فیزیک آماری سرجایشان هستند.

ذیلاً تنها به احتمال ریاضی خواهیم پرداخت، البته بجز توضیح زیر در باب پرتاب سکّه. مکانیک نیوتنی یک مدل ریاضی محدود برای پرتاب سکه به‌دست می‌دهد. در پرتاب سکه، جسم جامد تحت تأثیر گرانش سقوط می‌کند. حرکت سکه در مدل نیوتنی توسط قوانین نیوتن تعیین می‌شود، و صحبت در بارهٔ آنچه بر سر سکه می‌آید بدون اِعمال این قوانین خالی از نقص نخواهد بود. تنها این قوانین‌اند، و نه افاضات فلسفی، که قادرند تأثیر کمّی و اهمیت شرایطِ آغازی و پایانیِ حرکت سکه را توضیح داده و اظهارات مربوط به هم‌احتمالیِ آمدن شیر و خط را توجیه کنند. البته این قوانین در بهترین حالت می‌توانند تحلیلِ پرتاب سکه را به ملاحظاتی در باب شرایط آغازی و پایانیِ پرتاب تحویل نمایند، اما این شرایط می‌توانند نشانگر این باشند که «هم‌احتمالی» به چه وابسته است و از این رهگذر تعبیر معقول و مناسبی برای آن فراهم آورند.

## ۵. احتمال ریاضی پیش از عصر تعریفهای دقیق

پیش از سال ۱۹۰۰ میلادی پیشرفتهای مهم بسیاری در عرصهٔ احتمال ریاضی صورت پذیرفت، اما این موضوع هنوز ریاضیات نبود. هر چند از زمینه‌های احتمالاتیِ غیرریاضی، مسائلی در ترکیبیات، معادلات تفاضلی و معادلات دیفرانسیل سر برمی‌آورد، حداقلِ توجه به بنیان ریاضی این زمینه‌ها و حداکثرِ توجه به مسائل ریاضیِ محض برآمده از آنها معطوف می‌شد. این توجه نابرابر ناگزیر بود، چه نظریهٔ اندازه که برای مدلسازی ریاضی زمینه‌های احتمالاتی جهان واقع مورد نیاز است هنوز ابداع نشده بود.

همیشه این امر روشن بود که در هر احتمال ریاضی کلاسیکی که عرضه گردد، مفهوم جمعپذیری احتمال آنچنانکه در مورد رویدادهای دو به دو ناسازگار جهان واقع برقرار است، مفهومی اساسی خواهد بود. البته ریاضیدانان بسیار پیش از سال ۱۹۰۰ با تابعهای جمعی مجموعه‌ها که از مفاهیم حجم، جِرم و امثال آنها نشأت گرفته بودند آشنایی داشتند. همچنین دریافته بودند که زمینه‌های مربوط به میانگینها به احتمال می‌انجامد. غالباً آشکار بود که برای طرح مسائل چگونه باید از زمینه‌ها سود جست، ولی اینکه چگونه یک زمینهٔ ریاضی کلی را می‌توان فرمولبندی کرد، یعنی چگونه می‌توان یک ساختار ریاضی تعریف نمود که بشود زمینه‌های گوناگون را در آن نشاند، واضح نبود.

بعداً معلوم شد شرطی ضعیفتر از جمعپذیری که کمتر با آن آشنایی داشتند شرطی اساسی است. در اینجا به زبان غیردقیق معمول سخن خواهیم گفت. اگر $x_1$، $x_2$، ... اعدادی باشند که تصادفی به دست آورده‌ایم و اگر $A$ مجموعه‌ای از اعداد باشد، احتمال این‌را که دستِکم یکی از اعداد آن دنباله در $A$ باشد در نظر بگیرید. به عبارت دیگر احتمال این را که مداری از این حرکت برروی نقاط یک خط، $A$ را قطع کند در نظر بگیرید. با محاسبهٔ معمولی (و چشمپوشی از هر نوع دقت) تابع $\phi$ از $A$ به $\phi(A)$ تعریف می‌شود که در حالت کلی جمعپذیر نیست. درواقع $\phi$ در نابرابری زیر صدق می‌کند.

$$\phi(A)+\phi(B)-\phi(A\cup B)\geq\phi(A\cap B) \qquad (۱.۵)$$

حال آنکه جمعپذیری $\phi$، برابری را در (۱.۵) نتیجه می‌دهد. نکته اینجاست که سمت چپ (۱.۵) این احتمال است که دنبالهٔ $x_\bullet$ها هم $A$ و هم $B$ را قطع کند، احتمالی که حداقل برابر است با $\phi(A\cap B)$ و در حالت کلی بیشتر از آن است، یعنی احتمال اینکه دنبالهٔ $A\cap B$ را قطع نماید. نابرابریِ (۱.۵)، موسوم به نابرابریِ زیرجمعپذیریِ قوی، برای ظرفیت الکتروستاتیک یک جسم در $\mathbf{R}^3$ هم صادق است، و این به ارتباط نزدیک مابین نظریهٔ پتانسیل و احتمال اشاره دارد، که در نیمهٔ دوم قرن بیستم مفصلاً با کمک نظریهٔ شوکه برای ظرفیت الکتروستاتیک رشد یافت.

## ۶. پیدایش نظریهٔ اندازه

یادآوری می‌کنیم که یک میدان بورل ($=\sigma$ـجبر) از زیرمجموعه‌های یک فضا عبارت است از گردایه‌ای از زیرمجموعه‌ها که تحت عملهای متمم‌گیری و تشکیل اجتماعها و اشتراکهای شمارا بسته است. ردهٔ مجموعه‌های بورلِ یک فضای متریک کوچکترین $\sigma$ـجبری است که همهٔ مجموعه‌های باز آن فضا را در خود دارد. فضای اندازه‌پذیر زوجی است چون $(S,\mathbb{S})$، که $S$ فضایی است و $\mathbb{S}$ عبارت است از $\sigma$ـجبری از زیرمجموعه‌های $S$. مجموعه‌های درون $\mathbb{S}$ را مجموعه‌های اندازه‌پذیرِ فضا می‌نامیم. در آنچه در زیر می‌آید، اگر $S$ متریک باشد $\sigma$ـجبرِ الحاقی که آن را به یک فضای اندازه‌پذیر تبدیل می‌کند همواره $\sigma$ـجبرِ مجموعه‌های بورل آن خواهد بود. به‌ویژه $(\mathbf{R}^N,\mathbb{R}^N)$ نماد فضای اقلیدسیِ $N$بُعدی است که با مجموعه‌های بورلش زوج فضای اندازه‌پذیر را تشکیل می‌دهند. هنگامی‌که $N=1$، این

نماد را نمی‌نویسیم. یک تابع اندازه‌پذیر از فضای اندازه‌پذیر $(S_1, \mathbb{S}_1)$ به فضای اندازه‌پذیر $(S_2, \mathbb{S}_2)$ تابعی است از $S_1$ به $S_2$ با این خاصیت که نگارهٔ وارون مجموعه‌ای درون $\mathbb{S}_2$ مجموعه‌ای در $\mathbb{S}_1$ باشد.

نظریهٔ اندازه با رسالهٔ لبگ (۱۹۰۲)، که در آن تعریف حجم در $\mathbf{R}^N$ به مجموعه‌های بورل تعمیم داده شده بود، آغاز گردید. رادون (۱۹۱۳) گام بعدی را برداشت و اندازه‌های عمومیترِ مجموعه‌های بورلِ $\mathbf{R}^N$ را (که روی زیرمجموعه‌های فشرده، متناهی‌اند) معرفی کرد. این اندازه‌ها را معمولاً از طریق تکمیل به رده‌های اندکی بزرگتر از ردهٔ مجموعه‌های بورل تعمیم می‌دهند. سرانجام فرشه (۱۹۱۵)، ۱۳ سال بعد از رسالهٔ دکتریِ لبگ، اعلام کرد که تمام آنچه تعریفها و عملهای نظریهٔ اندازه احتیاج دارند $\sigma$ـجبری از زیرمجموعه‌های یک فضای مجرد است که روی آن یک اندازه، یعنی یک تابع مجموعه‌ایِ شمارا جمعیِ مثبت، تعریف شده باشد. در هر قدم از این سلسله مراحل، توابعی مجموعه‌ای که لزوماً شمارا جمعی مثبت نیستند ــ اندازه‌های علامت‌دار ــ جزو نظریه شدند. همچنان‌که در زیر اشاره شده، معلوم گردید قضیهٔ رادون‌ـنیکودیم (۱۹۳۰)، حاوی شرطهای لازم و کافی برای آنکه یک تابع مجموعه‌ایِ شمارا جمعی از مجموعه‌ها را بتوان به‌صورت انتگرالی روی مجموعه‌ها بیان کرد، آخرین نتیجهٔ اساسی مورد نیاز برای فرمولبندی تعریفهای بنیادی احتمال ریاضی است. این ۲۸ سال پیش از آن بود که نظریهٔ لبگ به اندازهٔ کافی گسترش داده شود که برای بنیان ریاضیِ احتمال بسنده باشد. ولیکن این گسترش به منظور فراهم آوردن بنیانی برای احتمال انجام نپذیرفت. نظریهٔ اندازه به‌عنوان بخشی از آنالیز کلاسیک پا گرفت، و کاربردهای بلافصلی در آنالیز داشت، از قبیل تقریباً همه جا مشتق‌پذیریِ یک تابع یکنوا (نسبت به اندازهٔ لبگ).

ایرادهایی به این موضوع وارد شده است که احتمال ریاضی معمولاً نه تنها جمعی بلکه شماراجمعی باید باشد. این سؤال که آیا احتمال در جهان واقع شماراجمعی است، به فرض اینکه معنایی داشته باشد، به این معناست که آیا مدلهای ریاضی پدیده‌های احتمالاتی جهان واقع لزوماً شامل تابعهای مجموعه‌ای شماراجمعی هستند یا نه. درواقع چه‌بسا زمینه‌های جهان واقع وجود داشته باشند که مدل ریاضی مناسب برای آنها مبتنی بر تابعهای مجموعه‌ای متناهیاً جمعی است و نه شماراجمعی. ولی چنین تابعهای مجموعه‌ای، چه در زمینه‌های ریاضی یا غیرریاضی، کاربردهای بسیار اندکی داشته‌اند، و این‌گونه تابعها بیش از این مورد بحث ما نخواهند بود.

## ۷. کاربردهای نخستینِ نظریهٔ رسمی اندازه در مبحث احتمال

در این بحث به اصطلاحاتی احتمالاتی، که یادگارهای به‌جا مانده از زمینهٔ تاریخی نظریهٔ احتمال‌اند، نیاز خواهیم داشت. فضای احتمال عبارت است از یک سه‌تایی $(S, \mathbb{S}, \mathrm{P})$، که $(S, \mathbb{S})$ یک فضای اندازه‌پذیر و $\mathrm{P}$ اندازه‌ای روی $\mathbb{S}$ است که $\mathrm{P}(S) = 1$. اندازه‌ای با این شرطِ بهنجارسازی را اندازهٔ احتمال می‌خوانیم. متغیر تصادفی عبارت است از تابعی اندازه‌پذیر از یک فضای احتمال $(S, \mathbb{S}, P)$ به یک فضای اندازه‌پذیر $(S', \mathbb{S}')$. فضای $S'$، یا، اگر دقیقتر بنویسیم، $(S', \mathbb{S}')$، یک فضای حالتِ متغیر تصادفی است. استقلالِ دوبه‌دوی متغیرهای تصادفی به شیوهٔ کلاسیک تعریف می‌گردد.

توزیع یک متغیر تصادفی $x$ عبارت است از تابع اندازهٔ $\mathrm{P}_x$ روی $\mathbb{S}'$ که این‌گونه تعریف می‌شود:

$$\mathrm{P}_x(A') = \mathrm{P}\{s \in S : x(s) \in A'\}$$

توزیع توأم تعداد متناهی متغیر تصادفی روی یک فضای احتمال واحد با تبدیل $x$ به یک بردار و تعیینِ $\mathbb{S}'$ و $S'$ به نحو مقتضی حاصل می‌شود. فرایند تصادفیْ خانواده‌ای از متغیرهای تصادفی همچون $\{x(t, \bullet), t \in I\}$ است از یک فضای احتمال $(S, \mathbb{S}, \mathrm{P})$ به یک فضای حالت $(S', \mathbb{S}')$. مجموعهٔ $I$ را مجموعهٔ اندیس فرایند می‌نامیم. پس هر فرایند تصادفیْ تابعی با دو متغیر، $(t, s) \to x(t, s)$، از $I \times S$ به یک فضای حالت مشخص می‌کند. تابع $x(t, \bullet)$ از $S$ به $S'$، $t$اُمین متغیر تصادفیِ فرایند است؛ تابع $x(\bullet, s)$ از $I$ به $S'$، $s$اُمین تابع نمونه‌ای، یا مسیر نمونه‌ای، یا اگر $I$ یک دنباله باشد، دنبالهٔ نمونه‌ای است.

بورل (در سال ۱۹۰۹) خاطرنشان کرد که در نمایش دودوییِ عدد $x$ای بین صفر و یک، به شکل $0.x_1x_2\ldots$، که هر رقم $x_j$ یا صفر است یا یک، این رقمها توابعی هستند از $x$، و اگر بازهٔ $[0, 1]$ را با اندازهٔ لبگ در نظر بگیریم که یک اندازهٔ احتمال بر این بازه است، این تابعها به شکل معجزه‌آسایی متغیرهایی تصادفی می‌شوند که دقیقاً همان توزیعهایی را دارند که در محاسبهٔ احتمالات پرتاب سکه به‌کار می‌روند. یعنی $2^{-n}$ برابر است با احتمالِ منسوب به این رویداد که، در یک آزمایشِ پرتاب سکه، نخستین $n$ پرتابْ دنبالهٔ معینی از شیر و خط به‌دست بدهد، و $2^{-n}$ همچنین طول کل (= اندازهٔ لبگ) تعدادی متناهی بازه است که نقاط متعلق به آنها بسطهایی دودویی با دنباله‌ای مشخص از صفرها و یکها در $n$ جایگاه خاص دارند. پس یک صورت ریاضیِ قانون اعداد بزرگ در پرتابِ سکه، عبارت است از نوعی وجود یک حد برای دنبالهٔ میانگینهای تابعیِ $\{(x_1 + \cdots + x_n)/n, n \geq 1\}$. محاسباتِ مبتنی بر احتمالات مقدماتی کلاسیک نشان توانند داد که این دنباله از میانگینها در اندازه به $\frac{1}{2}$ میل می‌کند، ولی بیانِ ریاضی قویتری از قانون اعداد بزرگ حکمی بود که بورل به‌دست آورد ــ طی یک برهانِ اشتباه و غیرقابل تصحیح ــ که این دنباله از میانگینها به‌ازای تقریباً هر $x$ به $\frac{1}{2}$ میل می‌کند (اندازهٔ لبگ). یک سال بعد فیبر[1] برهان درستی برای این حکم ارائه کرد و از آن هنگام برهانهای بسیار ساده‌تری هم به‌دست آمده‌اند. [فرشه حرمت بورل را نگه داشت: «برهان بورل بیش از حدْ کوتاه است. در آن چندین استدلالِ میانی حذف شده است و نیز احکامی بدون برهان فرض شده‌اند.»] این قضیه قدم مهمی بود، نمونه‌ای از نوع جدیدی از قضایای همگرایی در نظریهٔ احتمال. توجه کنید که (خوشبختانه) ریاضیدانان محض به تعبیر این قضیه در دنیای واقعیِ آدمهای واقعیِ در حال پرتابِ سکه‌های واقعی نیازی ندارند. برخی از قولهای نقل شده نشانگرِ این هستند که نه تنها نیازی ندارند بلکه اصلاً نباید به چنین تعبیری دست یازند.

دانیل (۱۹۱۸) رهیافتی ژرف به نظریهٔ اندازه را به‌کار گرفت که در آن انتگرال پیش از اندازه تعریف می‌شود، و رهیافتی (نه چندان سرراست) به دنباله‌های نامتناهیِ متغیرهای تصادفی از طریق اندازه‌هایِ تعریف شده در یک فضای اقلیدسی بینهایت بُعدی به‌دست آورد.

1. Faber

فرایندِ تصادفیِ حرکتِ براونی در $\mathbf{R}^۳$ مدلِ ریاضیِ حرکتِ براونی است، یعنی حرکت ذره‌ای میکروسکوپی در یک سیّال در حالی‌که مولکولهای آن سیّال به آن ذره اصابت می‌کنند. فرض می‌کنیم حرکت از مبدأ یک دستگاه مختصات دکارتی در $\mathbf{R}^۳$ آغاز می‌شود تا فرایند بهنجار شود، و یک حرکت براونی (بهنجارشده) در $\mathbf{R}$ عبارت است از فرایندِ تابعی مختصاتی از یک فرایند بهنجارشده در $\mathbf{R}^۳$، که از صفر آغاز می‌شود. فرایندِ حرکتِ براونی (بهنجارشده) در $\mathbf{R}^N$ فرایندی است که توسط $N$ فرایند حرکت براونیِ دوبه‌دو مستقل در $\mathbf{R}$ تعریف می‌شود. معلوم بود که توزیعهای توأم متغیرهای تصادفیِ یک فرایند حرکت تصادفی چه باید باشند، و فرض می‌شد که در یک الگوی ریاضیِ مناسب، ردهٔ مسیرهای پیوسته دارای احتمالِ ۱ باید باشد. پیش از آغاز قرنِ بیستم باشه‌لیه[1] حتی توزیعهای مهم متعددی استخراج کرده بود که همگی به فرایند حرکت براونی در $\mathbf{R}$ مربوط بودند، ازجمله توزیع مربوط به تغییر بیشینه در طول یک بازهٔ زمانی. بدین منظور وی توزیعهای متناظر با یک قدم‌زدن تصادفیِ گسسته را پیدا می‌کرد و سپس حد را هنگامی‌که طول قدمها به سمتِ صفر میل می‌کرد به‌دست می‌آورد. دقیقتر بگوییم، آنچه باشه‌لیه استخراج نمود توزیعهایی بودند که برای فرایند حرکت براونی کارایی داشتند، به فرض آنکه اصلاً چیزی تحت عنوان فرایند حرکت براونی وجود داشته باشد، و به فرض اینکه بشود آن‌را با آن قدم زدن‌های تصادفی تقریب زد. توجه کنید که شکی در وجود حرکت براونی نیست؛ حرکت براونی را می‌شود زیر میکروسکوپ نظاره کرد. ولی هنوز برهانی برای وجود یک فرایند تصادفی، یک ساختِ ریاضی، با خواص مطلوب در دست نبود. وینر (۱۹۲۳) فرایند مطلوب حرکت براونی را که امروزه گاه فرایند وینر نامیده می‌شود ساخت. بدین منظورْ وی از رهیافتِ دانیل به نظریهٔ اندازه استفاده کرد تا اندازه‌ای با خواص ذیل بر فضایِ $S$ از توابع پیوسته به‌دست آورد: اگر $x(t,\bullet)$ متغیری تصادفی باشد که با مقدارِ یک تابع در $S$ در زمان $t$ تعریف شده باشد، فرایند تصادفی این متغیرهای تصادفی فرایندی تصادفی است با اعضای $S$ به‌عنوان توابع نمونه‌ای، و با توزیعهای توأمی که برای فرایند حرکت براونی داشتیم به‌عنوان توزیعهای توأم متغیر تصادفی.

نتایجی که باشه‌لیه به‌دست آورده بود سالهای سال مغفول ماندند، و درواقع چندین بار از نو کشف شدند. کارِ وینر، همانند کار بُنیادی او در نظریهٔ پتانسیل، تأثیر مستقیم چندانی بر جای نگذاشت زیرا در مجله‌ای چاپ شد که انتشار گسترده‌ای نداشت. این یکی از جلوه‌های نبوغ او بود که پژوهش خود در باب حرکت براونی را هم در آن زمان و هم بعداً بدون اطلاع از آنچه اغلب می‌دانستند و برخی از تکنیکهای ریاضی مقدماتی و مفید نظریهٔ احتمال، به پیش بُرد.

اشتاینهاوس (۱۹۳۰) نشان داد که استدلالات کلاسیک برای استخراجِ قضیه‌های متعارف احتمال را می‌توان با پذیرفتن اندازهٔ لبگ برروی بازه‌ای بر محور حقیقی به طولِ ۱، به‌عنوان اندازهٔ احتمال پایه، و تعبیر متغیرهای تصادفی به‌عنوان توابع اندازه‌پذیر لبگ بر آن بازه و تعبیر امید ریاضی متغیرهای تصادفی به‌عنوان انتگرال آنها، در یک قالب دقیق جای داد. هیچ برهان تازه‌ای لازم نبود؛ آنچه لازم بود تنها ترجمهٔ مناسبی از واژگان کلاسیک به قالبِ تازهٔ او بود. اگر همهٔ آنچه ریاضی‌سازیِ احتمالات توسط نظریه اندازه می‌توانست به ارمغان آورد همین می‌بود، دیدگاه تحقیرآمیز غیرریاضیدانان نسبت به ریاضیات دقیق موجه می‌بود.

1. Bachelier

## ۸. تکنگاشت کولموگوروف به تاریخ ۱۹۳۳

کولموگوروف (در ۱۹۳۳) بنیان ریاضی ذیل را برای نظریهٔ احتمال پدید آورد.

(الف) قالب احتمال ریاضی عبارت است از یک فضای احتمالِ $(S,\mathbb{S},\mathrm{P})$. مجموعه‌های عضوِ $\mathbb{S}$ متناظر ریاضیِ رویدادهای جهان واقع‌اند؛ نقاطِ $S$ متناظر رویدادهای مقدماتی، یعنی تک‌مشاهده‌های (ممکن) مربوط به جهانِ واقع، هستند.

(ب) متغیرهای تصادفی روی $(S,\mathbb{S},\mathrm{P})$، متناظر توابع مشاهدات جهان واقع‌اند. فرض کنید $\{x(t,\bullet), t\in I\}$ فرایندی تصادفی بر یک فضای احتمال $(S,\mathbb{S},\mathrm{P})$ با فضایِ حالتِ $S'$ است. مجموعه‌ای از $n$ متغیر تصادفی فرایند توزیع احتمالی روی $S'^n$ دارد. این چنین توزیعهای متناهیْ بُعدی دوبه‌دو سازگار هستند بدین معنی که اگر $۱ \le m < n$، آنگاه توزیع توأم $x(t_۱,\bullet)$، $\cdots$، $x(t_m,\bullet)$ روی $S'^m$، توزیع $m$بُعدیِ القاءشده توسط توزیع $n$بعدیِ $x(t_۱,\bullet)$، $\cdots$، $x(t_n,\bullet)$ روی $S'^n$ است.

(پ) از سوی دیگر، کولموگوروف ثابت کرد که به‌ازای هر مجموعهٔ اندیس مفروضِ $I$، و یک فضای اندازه‌پذیر $(S',\mathbb{S}')$ با شرایط مناسب (مثلاً فضای اندازه‌پذیر می‌تواند یک فضای متریِ جدایی‌پذیر و کامل با $\sigma$ـجبرِ مجموعه‌های بورلِ آن باشد) و مجموعه‌ای از توزیعهای دوبه‌دو سازگار روی $S'^n$ (که $n$ عدد صحیح مثبتی است) اندیسگذاری‌شده با زیرمجموعه‌های متناهیِ $I$، یک اندازهٔ احتمال و یک فرایند تصادفیِ $\{x(t,\bullet); t\in I\}$ وجود دارد که بر آن فضای احتمال تعریف شده است، که فضای حالت آن $S'$ است با توزیعهای متغیر تصادفی توأمِ از پیش داده شده. برای اثبات این حکمْ او یک اندازهٔ احتمال رویِ یک $\sigma$ـجبر از زیرمجموعه‌های فضای حاصلضرب $S'^I$، یعنی فضای تمام توابع از $I$ به $S'$، ساخت و متغیرهای تصادفیِ لازم را به شکل توابع مختصاتیِ $S'^I$ به‌دست آورد.

(ث) امید ریاضی یک متغیر تصادفیِ انتگرال‌پذیر و با مقادیر عددی، عبارت است از انتگرالِ آن نسبت به اندازهٔ احتمالِ مفروض.

(ج) تعریف کلاسیک احتمال شرطیِ یک رویداد (= مجموعهٔ اندازه‌پذیر) $A$، به فرض یک رویداد $B$ با احتمالِ اکیداً مثبت، عبارت است از $\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}$. بدین‌شیوه، برای $B$ی مشخص، احتمالات جدیدی به‌دست می‌آیند، و امیدهای ریاضیِ متغیرهای تصادفی برای $B$ی مفروض برحسب این احتمالاتِ شرطیِ جدید محاسبه می‌گردند. عامتر از آن، به فرض در دست داشتن گردایهٔ دلخواهی از متغیرهای تصادفی، احتمالات شرطی و امیدهای ریاضی نسبت به مقادیر داده شدهٔ آن متغیرهای تصادفی مورد نیازند که توابعی از مقادیر منسوب به متغیرهای تصادفیِ شرطی‌کننده هستند. اگر $(S,\mathbb{S},\mathrm{P})$ یک فضای احتمال باشد، و اگر گردایه‌ای از متغیرهای تصادفی داده شده باشد، $\mathbb{F}$ را کوچکترین زیر$\sigma$ـجبر $\mathbb{S}$ بگیرید که همه متغیرهای تصادفی مفروض نسبت به آن اندازه‌پذیر هستند. این $\sigma$ـجبر را $\sigma$ـجبرِ تولیدشده توسط شرایط حاکم بر متغیرهای تصادفی مفروض می‌نامند. به یک تابع اندازه‌پذیر حقیقیْ‌مقدار و تعریف‌شده روی گردایهٔ مفروضی از

متغیرهای تصادفی می‌توان به‌عنوان تابعی اندازه‌پذیر از $(S, \mathbb{F})$ به $\mathbf{R}$ نگریست. امید ریاضی شرطی کولموگوروف برای یک متغیر تصادفی $x$ بر $(S, \mathbb{S}, \mathrm{P})$ که انتگرال‌پذیر و حقیقی‌مقدار باشد، نسبت به $\sigma$-جبر $\mathbb{C}$ از مجموعه‌های اندازه‌پذیر، عبارت است از متغیری تصادفی که نسبت به $\mathbb{C}$ اندازه‌پذیر است و همان انتگرالی را دارد که $x$ روی هر یک از مجموعه‌های درون $\mathbb{C}$ دارد. وجود چنین متغیری تصادفی و یکتایی آن با تقریب مجموعه‌های $\mathrm{P}$-پوچ را قضیهٔ رادون-نیکودیم تضمین می‌کند. امید ریاضی شرطیِ $x$ نسبت به گردایه‌ای از متغیرهای تصادفی را چنین تعریف می‌کنند: امید ریاضی $x$ نسبت به $\sigma$-جبرِ تولیدشده با شرایطیِ روی متغیرهای تصادفی. احتمال شرطی یک مجموعهٔ اندازه‌پذیر $A$ طبق تعریف عبارت است از امید ریاضی شرطی یک متغیر تصادفی که روی $A$ برابر ۱ و جاهای دیگر صفر است.

مقالهٔ توصیفی کولموگورف به تاریخ ۱۹۳۳ تصویر مأیوس‌کننده‌ای از پیشرفت ریاضی ترسیم می‌کند. وی در نخستین صفحاتِ این مقاله صراحتاً می‌گوید که متغیرهای تصادفی حقیقی‌مقدار همان توابع اندازه‌پذیرند و امیدهایِ ریاضیْ انتگرالهای آنها. معهذا، به نظر می‌رسد که وی حتی در سال ۱۹۳۳ هم می‌پنداشته که ریاضیدانان با نظریهٔ اندازهٔ آشنا نیستند. در حقیقت، وقتی‌که در خلالِ تکنگاشتش به تعریفِ متغیر تصادفیِ حقیقیْ‌مقدار می‌رسد، راحت به صفحاتِ اول تکنگاشت ارجاع نمی‌دهد و نمی‌گوید که متغیر تصادفی همان تابع اندازه‌پذیر است، بلکه اندازه‌پذیریِ یک تابع حقیقی‌مقدار را عملاً تعریف می‌کند و نیز وقتی‌که می‌خواهد امید ریاضی یک متغیر تصادفی را تعریف کند صاف و ساده نمی‌گوید که این مقدار برابر است با انتگرال متغیر تصادفی نسبت به اندازهٔ احتمال مفروض، بلکه انتگرال را عملاً تعریف می‌کند. بعداً باز در تکنگاشت هنگامی‌که به قضیهٔ لبگ می‌رسد که گرفتن حد از دنباله‌های تابعیِ همگرا را تحت علامت انتگرال مجاز می‌دارد، به لبگ ارجاع نمی‌دهد بلکه برهان مفصلّی برای آنچه نیاز دارد می‌آورد. نویسندهٔ این سطور به خاطر می‌آورد که در هنگام تحصیل در سال ۱۹۳۲، صحبتهایی حاکی از عدم تأیید عمومیتِ قضیهٔ مورد بحث در سمیناری که ساکس[1] سخنران آن بود از طرف اساتید ابراز می‌شد، همان قضیه‌ای که اکنون قضیهٔ ویتالی-هان-ساکس خوانده می‌شود و از آن زمان به ابزار مهمی در نظریهٔ احتمال بدل شده است، که این خاطره مؤید احتیاط کولموگوروف در به کارگیری نظریهٔ اندازه است. [نویسنده همچنان به یاد می‌آورد که منظور کولموگوروف از اندازه روی فضای تابعی را تا زمانی دراز پس از آنکه تکنگاشت وی را خوانده بود، درنمی‌یافته است.]

مدتی طول کشید تا بنیانِ پیشنهادیِ کولموگوروف را احتمال‌دانان پذیرفتند. این ایده که یک متغیر تصادفی (ریاضی) چیزی جز یک تابع نیست، که فاقد هر فحوای رمانتیک است، برای برخی احتمال‌دانان تحقیرآمیز بود. یک آماردان برجسته در سال ۱۹۳۵ پرسید که آیا دو متغیر تصادفیِ حقیقی‌مقدارِ متعامد با میانگینها (انتگرالها)ی صفر لزوماً مستقل از یکدیگرند، [می‌دانیم] با این فرض اضافه که دارای یک توزیع گاوسیِ دومتغیره باشند، چنین هستند. مثال توابع سینوس و کسینوس بر بازهٔ $[0, 2\pi]$ با اندازهٔ احتمال برابر با اندازه لبگ تقسیم بر $2\pi$ برای ایشان غیرمترقبه بود. این دو تابع، که متعامدند و میانگینهای صفر دارند ولی مستقل نیستند، از قماش متغیرهای تصادفیی که به آنها عادت داشتند نبودند. شاید بعضی آنالیزدانان از اینکه در یابند که در بحث سریهای فوریه می‌شود به بحث در احتمال و امید ریاضی متهمشان کرد، خشنود شوند، و برخی از ایشان هم ممکن است احساس تحقیرشدگی کنند.

## ۹. بسطِ قهقهرایی بنیانِ کولموگوروف

بنیانی که کولموگوروف برای احتمال ریاضی ارائه کرده می‌تواند بسط یابد و به نظر بعضی از احتمال‌دانان لازم است که بسط یابد ــ احتمال‌دانانی که می‌خواهند اطمینان مشاهده‌گران به رخ دادن برخی رویدادها را مبنا قرار دهند بدون آنکه این اطمینان لزوماً وجه ریاضی عددی داشته باشد و به شیوهٔ اصل موضوعی به ارزیابی عددی این اطمینان و نهایتاً به جمع‌پذیری برسند. چنین تحلیلی در بحث راجع به مناسب بودن احتمال ریاضی به‌عنوان مدلی برای پدیده‌های جهان واقع روشنگر تواند بود، اما هر رهیافتی به موضوع که به توجیهی از محاسبات کلاسیک منتهی شده و از نظر ریاضی قابل استفاده باشد به بنیان پیشنهادی کولموگوروف ختم می‌گردد، حال هر طور که به بیان درآید، چه بنیان نظریهٔ اندازه‌ایِ احتمال کاری صورت نمی‌دهد جز آنکه یک چارچوب ریاضیِ دقیقِ صوری برای محاسبات کلاسیک و صورتهای ظریفترکنونی آن به دست می‌دهد. این چارچوب سبب شده است که احتمال ریاضی را بشود در بسیاری حوزه‌های دیگر ریاضی، مثلاً نظریهٔ پتانسیل و معادلات دیفرانسیل جزئی، به‌کار بُرد. هر چند چنین کاربردهایی درگذشته و پیش از پذیرش نظریهٔ اندازه به‌عنوان بنیانی برای احتمال معمول بود، ولی چارچوب احتمالاتی تنها نوعی حضور ریاضیات را تداعی می‌کرد نه اینکه خود جزء لاینفکی از ریاضیات باشد. معنی جوابها به‌عنوان احتمالات و امیدهای ریاضی را نمی‌شد فرمولبندی کرد و مورد بهره‌برداری قرار داد.

## ۱۰. مجموعهٔ ناشمارای اندیس

اگر مجموعهٔ اندیس $I$ مربوط به فرایند تصادفی $\{x(t, \bullet), t \in I\}$ بازه‌ای در خط حقیقی باشد، و اگر فضای حالتِ متغیرهای تصادفی $\mathbf{R}$ باشد، ردهٔ توابع نمونه‌ای پیوسته ممکن است اندازه‌پذیر نباشد. این مشکل، به‌عنوان مثال، در فرایندهایی پیش می‌آید که به توسط ساختِ کولموگوروفیِ یک اندازه روی یک فضای تابعی استخراج می‌شوند، حال توزیع توأم متغیرهای تصادفی آن فرایند هر چه می‌خواهد باشد. برای درک این مشکل، توجه کنید که اگر مجموعهٔ اندیس $I$ مربوط به یک فرایند تصادفی با فضای حالت $\mathbf{R}$، بازه‌ای باشد و $J$ زیرمجموعه‌ای از $I$، آنگاه تابع $s \to \sup_{t \in J} x(t, s)$ اندازه‌پذیر است به شرط آنکه $J$ شمارا باشد، ولی نه لزوماً وقتی‌که $J$ ناشمارا است. اگر بنا باشد کرانداری و پیوستگی توابع نمونه‌ای مورد بحث واقع شوند، باید در رابطه‌های احتمالاتیِ متغیرهای تصادفیِ یک فرایند تصادفی تغییر لازم داده شود تا چنین کوچکترین کرانهای بالایی، توابع اندازه‌پذیر ازکار در بیایند. دوب رهیافتِ شکسته‌بسته‌ای در سال ۱۹۳۷ ارائه کرد، ولی رهیافتِ کارامدتری تا پس از ۱۹۵۰ طرح نشد.

---

1. Saks

## ۱۱. اکراه احتمال‌دانان از پذیرش نظریهٔ اندازه

احتمال‌دانان مقاومت زیادی در مقابل پذیرش و به‌کارگیری نظریهٔ اندازه نشان دادند، هم در زمان کولموگوروف و هم پس از او. قول زیر نشانی است از اکراه برخی از ریاضیدانان نسبت به جداسازی ریاضیات از منشأ الهام آن.

**کَتس[1] (۱۹۵۹)**: اینکه چقدر وسواس داشته باشیم از نظریهٔ اندازه در نظریهٔ احتمال استفاده کنیم بستگی به سلیقه و دیدگاه دارد، من شخصاً طرفدار حداقلِ استفاده هستم چرا که اعتقاد راسخ دارم که نظریهٔ احتمال بیش از آنکه به نظریهٔ اندازه به معنای دقیق آن مربوط باشد با آنالیز، فیزیک و آمار ارتباط تنگاتنگ دارد.

## ۱۲. روابط جدیدی بین توابع که با ریاضی‌سازیِ احتمال ممکن شده است

نظریهٔ احتمال روابط جدیدی بین توابع به‌دست داد. برای نمونه دنبالهٔ $x_1$، $x_2$، ... از متغیرهای تصادفی حقیقی‌مقدار انتگرال‌پذیر روی یک فضای احتمال $(S, \mathbb{S}, \mathrm{P})$ را در نظر بگیرید و فرض کنید که امید ریاضی شرطی $x_n$ با فرض $x_1$، ...، $x_{n-1}$ (که $n > 1$) تقریباً همه جا به صفر میل می‌کند، یعنی انتگرال $x_n$ روی هر مجموعهٔ مشخص‌شده با شرایطِ روی متغیرهای تصادفیِ ماقبل، به صفر می‌گراید. اگر مربع این متغیرهای تصادفی انتگرال‌پذیر باشد، این شرط معادل با شرطی است بس قویتر از تعامدِ دوبه‌دو، چه $x_n$ بر هر تابعِ مربعاً انتگرال‌پذیر از $x_1$، ...، $x_{n-1}$ عمود خواهد بود. به‌نظر می‌رسد برنستاین (۱۹۲۷) نخستین کسی بوده باشد که نظام‌مندانه به چنین دنباله‌هایی پرداخته است. این شرط روی دنباله‌ای از توابع بدین معنی است که دنبالهٔ مجموعهای جزئی یک دنبالهٔ مفروض، به مفهومی معقول، متناظر است با یک بازیِ عادلانه! درواقع مجموعهای جزئی $y_1$، $y_2$، ... با این خاصیت مشخص می‌گردند که امید ریاضی $y_n$ نسبت به $y_1$، ...، $y_{n-1}$ تقریباً همه جا روی فضای احتمال برابر است با $y_{n-1}$. فرایندهای بهره‌مند از این خاصیت، موسوم به مارتینگل را نخستین بار ویل[2] (۱۹۳۹) به‌طور صریح به‌کار گرفت. این فرایندها کاربُردهای فراوان داشته‌اند، ازجمله در [حل] معادلات دیفرانسیل جزئی، مشتقگیری و نظریهٔ پتانسیل. ردهٔ مهم دیگری از دنباله‌های متغیرهای تصادفیْ ردهٔ دنباله‌هایِ دارای خاصیت مارکوف است. این دنباله‌ها این‌گونه مشخص می‌شوند که وقتی $n \geq 1$، احتمالهای شرطی برای $x_n$ نسبت به $x_1$، ...، $x_{n-1}$ تقریباً همه جا برابر با احتمالهای شرطی $x_n$ نسبت به $x_{n-1}$ هستند. به‌عبارت نه چندان دقیق، تأثیر حال، با علم به گذشته، تنها وابسته به نزدیکترین گذشته است. خاصیت مارکوف که در حالت خیلی خاصش در سال ۱۹۰۶ توسط مارکوف ارائه شد (و بعدها دیگران به افتخار او آن‌را به اسم او نامگذاری کردند) بسیار پرثمر از آب درآمده است. برای مثال در نیمهٔ دوم قرن به نظریهٔ پتانسیل احتمالاتی منجر شده است و از این رهگذر نظریهٔ پتانسیل کلاسیک را تعمیم داده و شامل آن شده است.

1. Kaç 2. Ville

## ۱۳. جایگاه نظریهٔ احتمال در نظریهٔ اندازه، و فراتر از آن در آنالیز چیست؟

برخی ریاضیدانان برآن‌اند که هرگاه با خواص تحلیلی احتمال وامیدهای ریاضی سروکار داشته باشیم، موضوعْ بخشی از آنالیز است، ولی اگر با دنباله‌های نمونه‌ای و توابع نمونه‌ای سروکار داشته باشیم، موضوعْ عبارت است از احتمال، نه آنالیز. این مؤلفان در موقعیت جالب‌توجهی هستند از این‌رو که در نظرکردن به تابع دومتغیرهٔ $(t,s) \to x(t,s)$ ــ مثلاً در فرایندهای تصادفی ــ اگر خانوادهٔ توابع $x(t,\bullet)$ هنگامی‌که $t$ تغییر می‌کند مورد مطالعه باشد آن‌را آنالیز می‌خوانند، ولی اگر خانوادهٔ توابع $x(\bullet,s)$ هنگامی‌که $s$ تغییر می‌کند مورد بررسی باشد آن‌را احتمال می‌نامند و قطعاً آنالیز به حساب نمی‌آورند. دقیقتر بگوییم، ایشان بحث پیرامون توزیعها و پرسشهای مربوطه را آنالیز می‌دانند، بحثهایِ به زبان توابع نمونه‌ای را نه. این دیدگاه در قول ذیل بیان شده است:

**پروتر**: ایتو[1] در سال ۱۹۴۴، با ارائهٔ انتگرالش که در آن فرایندهای تصادفی انتگران بودند، توانست پخش چندبُعدی را با تکنیکهای احتمالاتیِ محض مورد بررسی قرار دهد، که نسبت به روشهای آنالیزیِ فِلِر بهتر است.

نکتهٔ ذیل در باب همگراییِ مجموع توابع متعامد دشواریِ تفکیکِ احتمال (ریاضی) از باقیِ آنالیز را باز می‌نماید. فضای احتمال یک فضای اندازه است، ولی این بحث با تغییراتِ ساده‌ای برای هر فضایِ با اندازهٔ متناهی معتبر است.

اگر $x_\bullet$ دنباله‌ای از توابع متعامد بر یک فضای اندازهٔ احتمال باشد و $x_n^2$ دارای انتگرال $\sigma_n^2$ باشد، آنگاه، $\sum x_n$ در میانگین همگراست (ریس-فیشر) هرگاه

$$\sum \sigma_n^2 < +\infty \qquad (۱.۱۳)$$

سریِ متعامد تقریباً همه جا همگراست اگر یا شرطِ (۱.۱۳) به شرط قویترِ

$$\sum \sigma_n^2 \log^2 n < +\infty \qquad (۲.۱۳)$$

تغییر یابد (منشوف[2]-رادِماخر[3]) و یا شرطِ (۱.۱۳) محفوظ بماند ولی شرط تعامد به شرط قویترِ مذکور در بخشِ ۱۰ تغییر یابد (لِوی، ۱۹۳۷).

دیگر بر عهدهٔ خواننده است که داوری کند کدامیک از این نتایج نظریهٔ اندازه‌ای هستند وکدام‌یک احتمالاتی، و آیا اصلاً بیرون راندن احتمال ریاضی از قلمرو آنالیز فایده‌ای دارد، و اگر دارد، آیا نظریهٔ اندازه را هم نباید بیرون راند؟

******

• Joseph L. Doob, "The development of rigor in mathematical probability, (1900-1950)" in *The Development of Mathematics (1900-1950)*, edited by Jean-Paul Pier, Birkhauser (1994).

* جوزف دوب، دانشگاه ایلینویی، آمریکا

1. Itô 2. Menšov 3. Rademacher

# مقالهٔ کلاسیک

# ملاحظاتِ معناشناختی در منطقِ وجهی

سول کریپکی

این مقاله شرحی از بعضی ویژگی‌های نظریه‌های معناشناختیِ منطق‌های وجهی به‌دست می‌دهد [۱]. در موردِ یک توسیعِ خاصِ S5، این نظریه در «یک قضیهٔ تمامیت در منطقِ وجهی» [۲]. عرضه شده است و در «تحلیلِ معناییِ منطقِ وجهی» [۳] خلاصه شده است. مقالهٔ حاضر به یک جنبهٔ این نظریه ــ معرفیِ سورها ــ می‌پردازد و خود را عمدتاً به یک روشِ حصولِ این هدف محدود می‌کند. این مقاله صرفاً معناشناختی است، و لذا کاربردِ تابلوهای معنایی را، که برای عرضهٔ کاملِ نظریه ضروری است، نادیده می‌گیرد [۴]. نیز، اثبات‌ها عمدتاً کنار گذاشته خواهد شد.

چهار نظامِ وجهی را بررسی می‌کنیم. فرمول‌های $A$، $B$، $C$، ... از فرمول‌های اتمیِ $P$، $Q$، $R$، ... با استفاده از ادات‌های $\wedge$، $\sim$، و $\Box$ ساخته می‌شوند. نظامِ $M$ این قاعده‌ها و قالب‌های اصولِ موضوع را دارد:

A0. همان‌گویی‌های قدرِ صدقی[1]

A1. $\Box A \supset A$

A2. $\Box(A \supset B) \supset . \Box A \supset \Box B$

R1. $A, A \supset B / B$

R2. $A / \Box A$

1. truth-functional tautologies

اگر این قالبِ اصولِ موضوع را اضافه کنیم، S4 را به‌دست می‌آوریم:

$$\Box A \supset \Box\Box A$$

نظامِ براؤری[1] را به‌دست می‌آوریم اگر به $M$ اضافه کنیم:

$$A \supset \Box\Diamond A$$

S5 را، اگر اضافه کنیم:

$$\Diamond A \supset \Box\Diamond A$$

نظام‌های وجهی‌ای که قضایاشان تحتِ قاعده‌های R1 و R2 بسته‌اند، و شاملِ همهٔ قضایای $M$اند، «نرمال» نامیده می‌شوند. اگرچه ما نظریه‌ای پرورانده‌ایم که بر نظام‌های غیرنرمالی چون S2 و S3ی لوئیس هم قابلِ اِعمال‌اند، در اینجا توجه‌مان را به نظام‌های نرمال محدود می‌کنیم.

برای به‌دست آوردنِ معناشناسی‌ای برای منطقِ وجهی، مفهومِ ساختارِ مدلی (ی نرمال) را معرفی می‌کنیم. هر ساختارِ مدلی (س. م.) سه‌تاییِ مرتبِ $(\mathbf{G}, \mathbf{K}, R)$ای است که در آن $\mathbf{K}$ یک مجموعه است، $R$ رابطه‌ای بازتابی روی $\mathbf{K}$ است، و $\mathbf{G} \epsilon \mathbf{K}$. شهوداً به موضوع به این صورت نگاه می‌کنیم: $\mathbf{K}$ مجموعهٔ همهٔ «جهان‌های ممکن» است؛ $\mathbf{G}$ «جهانِ واقع» است. اگر $\mathbf{H}_1$

1. *Brouwersche*

عصرِ جدیدِ بررسی‌های ریاضی‌ـمنطقیِ موجهات از سالِ ۱۹۵۹ با انتشارِ قضیه‌های تمامیتِ کریپکی آغاز شد. مقالهٔ حاضر که در آثارگانِ فلسفی و منطقی بسیار به آن ارجاع می‌شود در ۱۹۶۳ منتشر شده است و بحثی توصیفی در بعضی موضوعاتِ منطق‌های وجهی است. ترجمه بر پایهٔ این متن است:

Saul A. Kripke, “Semantical considerations on modal logic”, Leonard Linsky, ed., *Reference and Modality*, Oxford University Press, 1971 (reprinted 1979), pp. 63-72, 172.

از میانِ پژوهشهای کریپکی در سایرِ مباحثِ منطق ریاضی شاید کارهای او در مجموعه‌های پذیرفتنی و نظامِ معروف به «کریپکی‌ـپلاتک» (KP) از بقیه مشهورتر باشد.

کریپکی امریکایی و متولدِ ۱۹۴۰ است، و سی سالی است (از زمانِ چاپِ اولِ نام‌گذاری و ضرورت، ۱۹۷۲) که از تأثیرگذارترین فیلسوفانِ تحلیلی‌مشرب قلمداد می‌شود. هر چند او از ۱۹۸۲ عملاً چیزی منتشر نکرده است، افسانه‌ها حاکی از آن است که آثارِ زیادی تولید کرده است.

و $\mathbf{H}_2$ دو جهان باشند، $\mathbf{H}_1 R \mathbf{H}_2$ شهوداً یعنی اینکه $\mathbf{H}_2$ «نسبت به» $\mathbf{H}_1$ «ممکن» است؛ یعنی اینکه هرگزارهٔ درست در $\mathbf{H}_2$ در $\mathbf{H}_1$ ممکن است. در این صورت، به‌وضوح، رابطهٔ $R$ مسلماً باید بازتابی باشد؛ هر جهانِ $\mathbf{H}$ نسبت به خودش ممکن است، چون هرگزارهٔ درست در $\mathbf{H}$ بطریقِ اولی در $\mathbf{H}$ ممکن است. بازتابی بودن بدین‌گونه شهوداً قیدی طبیعی است. می‌توانیم، متناظر با «اصولِ موضوعِ تحویلی»[1]ی گوناگونِ منطقِ وجهی، قیدهای بیشتری وضع کنیم: اگر $R$ متعدی باشد، $(\mathbf{G}, \mathbf{K}, R)$ را یک S4-س.م. می‌گوییم؛ اگر $R$ متقارن باشد، $(\mathbf{G}, \mathbf{K}, R)$ یک س.م.ی براؤری است؛ و اگر $R$ رابطه‌ای هم‌ارزی باشد، $(\mathbf{G}, \mathbf{K}, R)$ را یک S5-س.م. می‌گوییم. ساختارِ مدلیِ بدونِ محدودیت یک $M$-ساختارِ مدلی هم خوانده می‌شود.

برای تکمیلِ این تصویر، مفهومِ مدل را نیاز داریم. با مفروض بودنِ یک ساختارِ مدلیِ $(\mathbf{G}, \mathbf{K}, R)$، هر مدل به هر فرمولِ اتمی (متغیر گزاره‌ای)ی $P$ یک ارزشِ صدقِ $\mathbf{T}$ یا $\mathbf{F}$ در هر جهانِ $\mathbf{H}\epsilon\mathbf{K}$ تخصیص می‌دهد. به‌طورِ صوری، هر مدلِ $\varphi$ روی یک س.م.یِ $(\mathbf{G}, \mathbf{K}, R)$ یک تابع دوتاییِ $\varphi(P, \mathbf{H})$ است، که در آن $P$ در فرمول‌های اتمی تغییر می‌کند و $\mathbf{H}$ در اعضای $\mathbf{K}$ تغییر می‌کند، که بُردش مجموعهٔ $\{\mathbf{T}, \mathbf{F}\}$ است. با مفروض بودنِ یک مدل، می‌توانیم تخصیص‌های ارزش‌های صدق به فرمول‌های غیراتمی را با استقراء تعریف کنیم. فرض کنید نقداً $\varphi(A, \mathbf{H})$ و $\varphi(B, \mathbf{H})$ به‌ازای همهٔ $\mathbf{H}\epsilon\mathbf{K}$ها تعریف شده باشد. در این صورت اگر $\varphi(A, \mathbf{H}) = \varphi(B, \mathbf{H}) = \mathbf{T}$، تعریف کنید $\varphi(A \wedge B, \mathbf{H}) = \mathbf{T}$؛ در غیر این صورت، $\varphi(A \wedge B, \mathbf{H}) = \mathbf{F}$. $\varphi(\sim A, \mathbf{H})$، $\mathbf{F}$ تعریف می‌شود اگر و تنها اگر $\varphi(A, \mathbf{H}) = \mathbf{T}$؛ در غیر این صورت، $\varphi(\sim A, \mathbf{H}) = \mathbf{T}$. نهایتاً، تعریف می‌کنیم $\varphi(\Box A, \mathbf{H}) = \mathbf{T}$ اگر و تنها اگر به‌ازای هر $\mathbf{H}'\epsilon\mathbf{K}$ که $\mathbf{H}R\mathbf{H}'$، $\varphi(A, \mathbf{H}') = \mathbf{T}$؛ در غیر این صورت، $\varphi(\Box A, \mathbf{H}) = \mathbf{F}$. شهوداً، این می‌گوید که $A$ در $\mathbf{H}$ ضروری است اگر و تنها اگر $A$ در همهٔ جهان‌های $\mathbf{H}'$ـِ نسبت به $\mathbf{H}$ ممکنْ درست باشد.

قضیهٔ تمامیت، در $M$ (S4، S5، نظامِ براؤری) $\vdash A$ اگر و تنها اگر به‌ازای هر مدلِ $\varphi$ روی هر $M$-(S4-، S5، براؤری-) ساختارِ مدلیِ $(\mathbf{G}, \mathbf{K}, R)$، $\varphi(A, \mathbf{G}) = \mathbf{T}$ [5].

این قضیهٔ تمامیتْ مفهومِ نحویِ اثبات‌پذیری در نظام‌های وجهی را با یک مفهومِ معناشناختیِ اعتبار معادل می‌کند.

بقیهٔ اینِ مقاله، به استثنای برخی مطالبِ پایانی، به معرفیِ سورها می‌پردازد. برای این کار، باید به هر جهانْ یک حوزهٔ افرادْ مربوط کنیم ــ افرادی که در آن جهان وجود دارند. به‌طورِ صوری، ساختارِ مدلیِ سوری (س.م.س.) را به‌صورتِ یک ساختارِ مدلیِ $(\mathrm{baG}, \mathbf{K}, R)$ تعریف می‌کنیم، همراه با یک تابعِ $\psi$ که به هر $\mathbf{H}\epsilon\mathbf{K}$ یک مجموعهٔ $\psi(\mathbf{H})$ تخصیص می‌دهد، که حوزه‌ی $\mathbf{H}$ خوانده می‌شود. شهوداً $\psi(\mathbf{H})$ مجموعهٔ همهٔ افرادِ موجود در $\mathbf{H}$ است. توجه کنید که، البته، $\psi(\mathbf{H})$ لازم نیست به ازای مقادیرِ مختلفِ $\mathbf{H}$ مجموعهٔ واحدی باشد، درست همان‌طور که، شهوداً، در جهان‌های دیگری غیر از جهانِ واقعی، بعضی افرادِ واقعاً موجودْ ممکن است غایب باشند در حالی که افرادِ جدیدی، مثلِ پگاسوس[2]، ظاهر شوند.

پس به نمادهای منطقِ وجهیْ یک فهرستِ نامتناهیِ $x$، $y$، $z$، ... از متغیرهای فردی اضافه می‌کنیم، و، به‌ازای هر عددِ صحیحِ نامنفیِ $n$، فهرستی از حرف‌های محمولیِ $n$موضعیِ $P^n$، $Q^n$، ...، که بالانویس‌ها گاهی از سیاقِ متنْ فهمیده می‌شوند. متغیرهای گزاره‌ای (فرمول‌های اتمی) را حرف‌های محمولیِ «۰ موضعی» به‌شمار می‌آوریم. سپس فرمول‌های خوش‌ساخت را به روشِ معمول می‌سازیم، و اکنون می‌توانیم خودمان را برای تعریفِ مدلِ سوری آماده کنیم.

برای تعریفِ مدلِ سوری، باید مفهوم اولیه را، که به هر فرمولِ اتمی در هر جهانْ ارزشِ صدقی نسبت می‌داد، گسترش دهیم. به‌طریقِ مشابه، باید فرض کنیم که در هر جهانْ هر حرفِ محمولیِ $n$موضعیْ مجموعهٔ خاصی از $n$تایی‌ها ــ مصداق‌اش در آن جهان ــ را معین می‌کند. مثلاً موردِ یک حرفِ محمولیِ تک‌موضعیِ $P(x)$ را در نظر بگیرید. دوست داریم بگوییم که، در جهانِ $\mathbf{H}$، محمولِ $P(x)$ در موردِ بعضی از افراد در $\psi(\mathbf{H})$ درست و در موردِ بقیه غلط است؛ به‌طورِ صوری، می‌گوییم که، نسبت به بعضی تخصیص‌های اعضای $\psi(\mathbf{H})$ به $x$، $\varphi(P(x), \mathbf{H}) = \mathbf{T}$ و نسبت به بقیه $\varphi(P(x), \mathbf{H}) = \mathbf{F}$. مجموعهٔ همهٔ افرادی که $P$ در موردِ آنها درست است مصداقِ $P$ در $\mathbf{H}$ نامیده می‌شوند. اما مسأله‌ای وجود دارد: آیا وقتی به $x$ مقداری در حوزهٔ یک جهانِ دیگرِ $\mathbf{H}'$ و نه در حوزهٔ $\mathbf{H}$ تخصیص داده می‌شود باید به $\varphi(P(x), \mathbf{H})$ ارزشِ صدقی داده شود؟ شهوداً، فرض کنید $P(x)$ به معنای «$x$ تاس است» باشد ــ آیا باید به نمونهٔ جایگزین‌شدهٔ «شرلوک هومز تاس است» ارزشِ صدقی تخصیص دهیم؟ هومز وجود ندارد، اما او در اوضاع و احوالِ دیگری وجود می‌داشت. آیا باید به این حکم که او تاس است ارزشِ صدقِ معینی تخصیص دهیم، یا نه؟ فرگه [6] و ستراؤسن [7] به این حکمْ ارزشِ صدقی تخصیص نخواهند داد؛ راسل خواهد داد [8]. برای اهدافِ منطقِ وجهی ما بر آن‌ایم که پاسخ‌های متفاوت به این پرسشْ نمایانگرِ قراردادهای مختلفی هستند. همگی قابلِ دفاع‌اند. تنها بحث‌های موجودی در موردِ این مسأله که من دیده‌ام ــ مالِ هینتیکا [9] و پرایر [10] ــ نظرِ فرگه-ستراؤسن را برمی‌گزینند. این نظر ضرورتاً باید به تغییراتی در منطقِ وجهیِ معمول منجر شود. علت‌اش این است که معناشناسیِ منطقِ گزاره‌ایِ وجهی، که قبلاً به‌دست داده‌ایم، این را مفروض می‌گرفت که هر فرمول باید درهر جهانْ ارزشِ صدقی اختیار کند؛ و حالا، به‌ازای هر فرمولِ $A(x)$ شاملِ یک متغیرِ آزادِ $x$، نظرِ فرگه-ستراؤسن الزام می‌کند که در یک جهانِ $\mathbf{H}$، وقتی به متغیرِ $x$ فردی تخصیص داده می‌شود که در حوزهٔ آن جهان نیست، به فرمول ارزشِ صدقی داده نشود. بدین ترتیب دیگر نمی‌توانیم انتظار داشته باشیم که قوانینِ اولیهٔ منطقِ گزاره‌ایِ وجهی در موردِ احکامِ شاملِ متغیرهای آزاد برقرار باشند، و با گزینشی روبه‌رو می‌شویم: یا در منطقِ گزاره‌ایِ وجهی بازنگری کنیم یا قاعدهٔ جایگزینی را محدود کنیم. پرایر اولی را انجام می‌دهد، هینتیکا دومی را. انتخابِ فرگه-ستراؤسن متضمنِ شقوقِ دیگری هم هست: آیا باید $\Box A$ را (در $\mathbf{H}$) به این معنا بگیریم که $A$ در همهٔ جهان‌های ممکن (نسبت به $\mathbf{H}$) درست است، یا فقط در هر چنین جهانی غلط نیست؟ شقِّ دوم صرفاً می‌خواهد که در هر جهانی $A$ یا درست باشد یا فاقدِ ارزشِ صدق باشد. پرایر، در نظامِ Qاش، در عمل هر دو نوعِ ضرورت را می‌پذیرد،

1. reduction axioms

۲. Pegasus؛ اسبِ بال‌دار-م.

یکی با عنوانِ '$L$' و دیگری با عنوانِ '$NMN$'. پرسشِ مشابهی در موردِ عطف مطرح می‌شود: اگر $A$ غلط باشد و $B$ ارزشِ صدقی نداشته باشد، آیا $A \wedge B$ را باید غلط بگیریم یا بی ارزشِ صدق؟

در گزارشِ کاملی از این نظریهٔ معناشناختی، می‌توانستیم همهٔ این گونه‌های نظرِ فرگه‌ـستراؤسن را بررسی کنیم. در اینجا راهِ دیگر را اختیار می‌کنیم، و فرض می‌کنیم که هر حکمی که شاملِ متغیرهای آزاد باشد در هر جهانی به‌ازای هر تخصیصی به متغیرهای آزادش ارزشِ صدقی دارد [۱۱]. به‌طورِ صوری، مطلب را به این صورت بیان می‌کنیم: قرار دهید $\mathrm{U} = \bigcup_{\mathbf{H}\epsilon\mathbf{K}} \psi(\mathbf{H})$. $\mathrm{U}^n$، $n$اُمین حاصل‌ضربِ دکارتیِ $\mathrm{U}$ با خودش است. هر مدلِ سوری روی یک س.م.س.ی $(\mathbf{G}, \mathbf{K}, R)$ را به‌صورتِ یک تابعِ دوتاییِ $\gamma(P^n, \mathbf{H})$ تعریف می‌کنیم، که در آن متغیر اول در حروفِ محمولیِ $n$تایی، به ازای $n$های دلخواه، تغییر می‌کند، و $\mathbf{H}$ در اعضای $\mathbf{K}$ تغییر می‌کند. اگر $n = ۰$، $\varphi(P^n, \mathbf{H})$ مساویِ $\mathbf{T}$ یا $\mathbf{F}$ است؛ اگر $n \geq ۱$، $\varphi(P^n, \mathbf{H})$ زیرمجموعه‌ای از $\mathrm{U}^n$ است. اکنون، به‌صورتِ استقرائی، به ازای هر فرمولِ $A$ و $\mathbf{H}\epsilon\mathbf{K}$، یک ارزشِ صدق، $\varphi(A, \mathbf{H})$، نسبت به تخصیصِ داده‌شده‌ای از اعضای $\mathrm{U}$ به متغیرهای آزادِ $A$ تعریف می‌کنیم. موردِ متغیر گزاره‌ای واضح است. به‌ازای هر فرمولِ اتمیِ $P^n(x_۱, \ldots, x_n)$، که $P^n$ یک حرفِ محمولیِ $n$موضعی است و $n \geq ۱$، با مفروض بودنِ تخصیصی از اعضای $a_۱$، $\ldots$، $a_n$ از $\mathrm{U}$ به $x_۱$، $\ldots$، $x_n$، تعریف می‌کنیم $\varphi(P^n(x_۱, \ldots, x_n), \mathbf{H}) = \mathbf{T}$ اگر $n$تاییِ $(a_۱, \ldots, a_n)$ عضوی از $\varphi(P^n, \mathbf{H})$ باشد؛ در غیر این صورت، نسبت به تخصیصِ مفروض، $\varphi(P^n(x_۱, \ldots, x_n), \mathbf{H}) = \mathbf{F}$. با داده شدنِ این تخصیص‌ها برای فرمول‌های اتمی، می‌توانیم با استقراء تخصیص‌ها را برای فرمول‌های پیچیده بسازیم. گام‌های استقراء برای اداتِ گزاره‌ایِ $\wedge$، $\sim$، $\Box$ قبلاً داده شده‌اند. فرض کنید یک فرمولِ $A(x, y_۱, \ldots, y_n)$ داریم، که $x$ و $y_i$ها تنها متغیرهای آزادِ حاضرند، و یک ارزشِ صدقِ $\varphi(A(x, y_۱, \ldots, y_n), \mathbf{H})$ به ازای هر تخصیصی به متغیرهای آزادِ $A(x, y_۱, \ldots, y_n)$ تعریف شده است. در این صورت $\varphi((x)A(x, y_۱, \ldots, y_n), \mathbf{H})$ را نسبت به تخصیصِ $b_۱$، $\ldots$، $b_n$ به $y_۱$، $\ldots$، $y_n$ (که $b_i$ها اعضای $\mathrm{U}$اند) $\mathbf{T}$ تعریف می‌کنیم، اگر به ازای هر تخصیصِ $a$، $b_۱$، $\ldots$، $b_n$ به‌ترتیب به $x$، $y_۱$، $\ldots$، $y_n$، که $a\epsilon\psi(\mathbf{H})$، $\varphi(A(x, y_۱, \ldots, y_n), \mathbf{H}) = \mathbf{T}$؛ در غیر این صورت $\varphi((x)A(x, y_۱, \ldots, y_n), \mathbf{H})$ را نسبت به تخصیصِ داده‌شده $\mathbf{F}$ تعریف می‌کنیم. توجه کنید که قیدِ $a\epsilon\psi(\mathbf{H})$ یعنی که، در $\mathbf{H}$، فقط روی اشیاء واقعاً موجود در $\mathbf{H}$ سور می‌گذاریم.

برای توضیحِ این معناشناسی، مثال‌های نظریهٔ نقضی به‌دست می‌دهیم برای دو پیشنهادِ آشنا برای قوانینِ تسویرِ وجهی ــ «فرمولِ بارکان[۱]» $(x)\Box A(x) \supset \Box(x)A(x)$ و معکوس‌اش $\Box(x)A(x) \supset (x)\Box A(x)$. برای هر کدام یک ساختارِ مدلیِ $(\mathbf{G}, \mathbf{K}, R)$ در نظر می‌گیریم، که $\mathbf{K} = \{\mathbf{G}, \mathbf{H}\}$، $\mathbf{G} \neq \mathbf{H}$، و $R$ همان حاصل‌ضربِ دکارتیِ $\mathbf{K}^۲$ است. به‌وضوح $R$ انعکاسی، متعدی، و متقارن است، پس ملاحظاتِ ما حتی در S5 هم موضوعیت دارد.

برای فرمولِ بارکان، $(\mathbf{G}, \mathbf{K}, R)$ را با تعریفِ $\psi(\mathbf{G}) = \{a\}$ و $\psi(\mathbf{H}) = \{a, b\}$، که $a$ و $b$ متمایزند، به یک ساختارِ مدلیِ سوری گسترش می‌دهیم. بعد، به ازای یک حرفِ محمولیِ تک‌موضعیِ $P$، یک مدلِ $\varphi$ تعریف می‌کنیم که در آن $\varphi(P, \mathbf{G}) = \{a\}$، $\varphi(P, \mathbf{H}) = \{a\}$. در این صورت به‌وضوح وقتی $a$ به $x$ تخصیص داده شود $\Box P(x)$ درست است؛ و چون $a$ تنها شیء در حوزهٔ $\mathbf{G}$ است، $(x)\Box P(x)$ هم چنین است. اما، $(x)P(x)$ به‌وضوح در $\mathbf{H}$ غلط است (چون وقتی $b$ به $x$ تخصیص داده شود $\varphi(P(x), \mathbf{H}) = \mathbf{F}$)، و لذا $\Box(x)P(x)$ در $\mathbf{G}$ غلط است. پس مثالِ نقضی برای فرمولِ بارکان داریم. توجه کنید که این مثالِ نقض کاملاً مستقل از این است که وقتی $b$ به $x$ تخصیص داده می‌شود به $P(x)$ در $\mathbf{G}$ ارزشِ صدقی تخصیص داده می‌شود یا نه، پس این مثال در نظام‌های هینتیکا و پرایر هم موضوعیت دارد. چنین مثال‌های نقضی را فقط در صورتی می‌توان مردود دانست، و فرمولِ بارکان را به مقامِ سابق بازگرداند، که ساختارهای مدلی این شرط را برآورند که هرگاه $\mathbf{H}R\mathbf{H}'$، $\mathbf{H}, \mathbf{H}'\epsilon\mathbf{K}$، $\psi(\mathbf{H}) \subseteq \varphi(\mathbf{H}')$.

برای معکوسِ فرمولِ بارکان، قرار دهید $\psi(\mathbf{G}) = \{a, b\}$، $\psi(\mathbf{H}) = \{a\}$، که باز هم $a \neq b$. تعریف کنید $\varphi(P, \mathbf{G}) = \{a, b\}$، $\varphi(P, \mathbf{H}) = \{a\}$، که $P$ حرفِ محمولیِ تک‌موضعیِ داده‌شده‌ای است. در این صورت به‌وضوح $(x)P(x)$ در هر دوی $\mathbf{G}$ و $\mathbf{H}$ برقرار است، و لذا $\varphi(\Box(x)P(x), \mathbf{G}) = \mathbf{T}$. اما وقتی $b$ به $x$ تخصیص داده شود، $\varphi(P(x), \mathbf{H}) = \mathbf{F}$ و لذا، وقتی $b$ به $x$ تخصیص داده شود، $\varphi(\Box P(x), \mathbf{G}) = \mathbf{F}$. پس $\varphi((x)\Box P(x), \mathbf{G}) = \mathbf{F}$، و مثالِ نقضِ مطلوب‌مان برای معکوسِ فرمولِ بارکان را به‌دست آورده‌ایم. اما این مثالِ نقض متکی است به اینکه، در $\mathbf{H}$، وقتی $b$ به $x$ تخصیص داده شود $P(x)$ واقعاً غلط است؛ پس ممکن است این نتیجه از بین برود اگر، به‌ازای این تخصیص، $P(x)$ در $\mathbf{H}$ فاقدِ ارزشِ صدق اعلام می‌شد. در این وضعیت، هنوز هم مثالِ نقض‌مان را خواهیم داشت اگر الزام کنیم که حکمِ ضروری در همهٔ جهان‌های ممکن درست باشد ('$L$'ـِ پرایر)، و فقط الزام نکنیم که هرگز غلط نباشد ('$NMN$'ـِ پرایر). با قراردادِ فعلی‌مان، می‌توانیم این مثالِ نقض را با صِرفِ الزامِ این امر از دور خارج کنیم که، به‌ازای هر س.م.س.، $\psi(\mathbf{H}) \subseteq \psi(\mathbf{H}')$ هرگاه $\mathbf{H}R\mathbf{H}'$.

این مثال‌های نقض به مشکلِ خاصی منجر می‌شوند: در S5ـِ سوری، پادمُدل‌هایی هم برای فرمولِ بارکان و هم برای معکوس‌اش عرضه کرده‌ایم. اما به نظر می‌رسد که پرایر نشان داده باشد [۱۲] که فرمولِ بارکان در S5ـِ سوریْ استنتاج‌پذیر است؛ و به نظر می‌آید که معکوس حتی در $M$ـِ سوری هم با این استدلال استنتاج شود:

(A) $(x)A(x) \supset A(y)$ (طبقِ نظریهٔ تسویر)

(B) $\Box((x)A(x) \supset A(y))$ (طبقِ ضروری‌سازی)

(C) $\Box((x)A(x) \supset A(y)) \supset \Box(x)A(x) \supset \Box A(y)$ (اصلِ موضوعِ A2)

(D) $\Box(x)A(x) \supset \Box A(y)$ (از (B) و (C))

(E) $(y)(\Box(x)A(x) \supset \Box A(y))$ (تعمیم روی (D))

(F) $\Box(x)A(x) \supset (y)\Box A(y)$ (طبقِ نظریهٔ تسویر، و (E))

به نظر می‌رسد حکم را با استفاده از اصولی به‌دست آورده‌ایم که همگی باید در نظریهٔ مدل‌ها معتبر باشند. در واقع، خطا در اِعمالِ ضروری‌سازی

---

۱. منسوب به Ruth Barcan Marcus، منطقدانِ امریکایی-م.

بر (A) نهفته است. در فرمولی مثلِ (A)، به متغیرهای آزاد تعبیرِ عمومیت می‌دهیم [۱۳]: وقتی (A) در مقامِ قضیه ادعا می‌شود، ادعای بستارِ عمومیِ معمولی‌اش، یعنی

$$(\mathrm{A}')\qquad (y)((x)A(x) \supset A(y))$$

را خلاصه می‌کند. حالا اگر ضروری‌سازی را بر (A′) اِعمال کنیم به‌دست خواهیم آورد

$$(\mathrm{B}')\qquad \Box(y)((x)A(x) \supset A(y))$$

از طرفِ دیگر، خودِ (B) به‌صورتِ ادعای

$$(\mathrm{B}'')\qquad (y)\Box((x)A(x) \supset A(y))$$

تعبیر می‌شود. برای به‌دست آوردنِ (B″) از (B′)، نیاز به قانونی به شکلِ $\Box(y)C(y) \supset (y)\Box C(y)$ داریم، که درست معکوسِ فرمولِ بارکان است که داریم سعی می‌کنیم اثبات‌اش کنیم. در واقع، به‌سادگی بررسی می‌شود که اگر $A(x)$ را با $P(x)$ جایگزین کنیم (B″) در پادمدلی که در بالا برای معکوسِ فرمولِ بارکان دادیم نادرست می‌شود.

می‌توانیم از این نوع مشکل اجتناب کنیم اگر، به اقتفای کواین [۱۴]، نظریهٔ تسویر را چنان صورت‌بندی کنیم که فقط فرمول‌های بسته را بتوان ادعا کرد. ادعای فرمول‌های حاویِ متغیرهای آزاد در بهترین حالت برای سهولت است؛ ادعای $A(x)$ با متغیرِ آزاد $x$ را می‌توان همواره با ادعای $(x)A(x)$ جایگزین کرد.

اگر $A$ فرمولی حاویِ متغیرهای آزاد باشد، یک بستارِ $A$ را هر فرمولِ بدونِ سوری تعریف می‌کنیم که با قراردادنِ سورهای عمومی و علامت‌های ضرورت، با هر ترتیبی، در جلوی $A$ به‌دست می‌آید. بعد اصولِ موضوعِ $M$ـِ سوری را بستارهای این قالب‌های کلّیْ تعریف می‌کنیم:

(۰) همهٔ همان‌گویی‌های قدرِ صدقی

(۱) $\Box A \supset A$

(۲) $\Box(A \supset B). \supset .\Box A \supset \Box B$

(۳) $A \supset (x)A$، که $x$ در $A$ آزاد نیست

(۴) $(x)(A \supset B). \supset .(x)A \supset (x)B$

(۵) $(y)((x)A(x) \supset A(y))$

قاعدهٔ استنتاجْ تفکیک[1] [یا: وضعِ مقدم] برای استلزامِ مادّی است. ضروری‌سازی را می‌توان به‌صورتِ یک قاعدهٔ استخراج‌شده به‌دست آورد

برای به‌دست آوردنِ توسیع‌های سوریِ S4، S5، نظامِ براؤری، کافی است به این قالب‌های اصولِ موضوعْ همهٔ بستارهای اصلِ موضوعِ تحویلِ مناسب را اضافه کنید.

نظام‌هایی که به‌دست آورده‌ایم این خواص را دارند: توسیعِ سرراستی‌اند از منطق‌های گزاره‌ایِ وجهی، بدونِ جرح و تعدیلاتِ Qی پرایر؛ بر خلافِ نحوهٔ عرضهٔ هینتیکا، قاعدهٔ جایگزینیْ بدونِ محدودیت برقرار است؛ و با این حال نه فرمولِ بارکانْ استنتاج‌پذیر است نه معکوسِ آن. به‌علاوه، قوانین نظریهٔ تسویر —جرح و تعدیل‌شده برای پذیرشِ حوزهٔ تهی— برقرارند. قضیهٔ تمامیّتِ معناشناختی‌ای که برای منطقِ گزاره‌ایِ وجهی عرضه کردیم را می‌توان به این نظام‌های جدید گسترش داد.

در نظامِ فعلی اگر بخواهیم می‌توانیم وجود را در مقامِ یک محمول معرفی کنیم. به لحاظِ معناشناختی، وجودْ یک محمولِ تک‌موضعیِ $\mathbf{E}(x)$ است که، به ازای هر مدلِ $\varphi$ روی هر س.م.ی $(\mathbf{G}, \mathbf{K}, R)$، در برابریِ $\varphi(\mathbf{E}, \mathbf{H}) = \psi(\mathbf{H})$ به ازای هر $\mathbf{H} \epsilon \mathbf{K}$ صدق می‌کند. به لحاظِ اصلِ موضوعی، می‌توانیم آن را با اصل گرفتنِ بستارهای فرمول‌های به شکلِ $(x)A(x) \wedge \mathbf{E}(y). \supset .A(y)$ و $(x)\mathbf{E}(x)$ معرفی کنیم. محمولِ $P$ی به‌کار رفته در بالا در مثالِ نقضِ معکوسِ فرمولِ بارکان را اکنون می‌توان به سادگی به‌صورتِ وجود باز شناخت. این امر نشان می‌دهد که وجود چه فرقی با محمولِ همان‌گویانهٔ $A(x) \vee \sim A(x)$ دارد اگرچه $\Box(x)\mathbf{E}(x)$ اثبات‌پذیر است. زیرا اگرچه $(x)\Box(A(x) \vee \sim A(x))$ معتبر است، $(x)\Box\mathbf{E}(x)$ نیست؛ اگرچه این ضروری است که هر چیزی وجود دارد، نتیجه نمی‌شود که هر چیزی خاطیتِ وجودِ ضروری را دارد.

می‌توانیم اینهمانی را به‌صورتِ معناشناختی در نظریهٔ مدل این‌طور معرفی کنیم که $x = y$ را در جهانِ $\mathbf{H}$ وقتی به $x$ و $y$ مقدارِ واحدی تخصیص داده شود درست و در غیر این‌صورت غلط تعریف کنیم؛ در این صورت وجود را می‌توان برحسبِ اینهمانی تعریف کرد، با مقرر کردنِ اینکه $\mathbf{E}(x)$ یعنی $(\exists y)(x = y)$. به دلایلی که در اینجا عرضه نمی‌شود، اگر مفهومِ ساختارِ مدلیِ سوری را پیچیده‌تر می‌کردیم می‌شد نظریهٔ این‌همانیِ وسیع‌تری به‌دست آورد.

با مطالبِ خلاصه و مجملی در موردِ تعبیرهای «اثبات‌پذیری»ی منطق‌های وجهی، که در هر مورد فقط برای حسابِ جمله‌ها عرضه می‌کنیم، بحث را ختم می‌کنیم. خواننده نکتهٔ اصلیِ این مقاله را درک کرده خواهد بود اگر این بخش را حذف کند. تعبیرهای اثبات‌پذیری مبتنی‌اند بر تمایل به الحاقِ یک عملگرِ ضرورت به نظامی صوری، مثلاً حسابِ پئانو، به شیوه‌ای که، به ازای هر فرمولِ $A$ از این نظام، $\Box A$ درست تعبیر شود اگر و تنها اگر $A$ در دستگاه اثبات‌پذیر باشد. استدلال شده است که چنین تعبیرهای «اثبات‌پذیری»ای برای عملگری وجهیْ لازم نیستند و می‌توانند با یک محمولِ اثبات‌پذیری، که بر عددِ گودلِ $A$ حمل می‌شود، جایگزین شوند؛ اما مقالهٔ پروفسور مانتگی‌یو در این مجلّد دست‌کم شکّی در موردِ این دیدگاه ایجاد می‌کند.

بگذارید نظامِ صوریِ $\mathbf{PA}$ی حسابِ پئانو را، آنچنان که در کلی‌نی [۱۵] صورت‌بندی شده، بررسی کنیم. به قواعدِ ساختْ عملگرهای ∧، ∼، و □ را الحاق می‌کنیم (عاطف و ناقضِ الحاق‌شده متمایز از عاطف و ناقضِ نظامِ اولیه‌اند)، که فقط بر فرمول‌های بسته عمل می‌کنند. در نظریهٔ مدلی که در بالا عرضه کردیم، فرمول‌های اتمی را متغیرهای گزاره‌ای، یا حرف‌های محمولی‌ای گرفتیم که متغیرهای فردیِ در پرانتز قرار گرفته‌ای به دنبال‌شان می‌آیند؛ در اینجا فرمول‌های اتمی را صرفاً فرمول‌های خوش‌ساختِ بستهٔ $\mathbf{PA}$ (نه فقط فرمول‌های اتمیِ $\mathbf{PA}$) می‌گیریم. یک ساختِ مدلیِ $(\mathbf{G}, \mathbf{K}, R)$ تعریف می‌کنیم، که $\mathbf{K}$ مجموعهٔ همهٔ مدل‌های شمارای متمایزِ (غیریکریختِ) $\mathbf{PA}$ است، $\mathbf{G}$ مدلِ استانداردِ اعدادِ طبیعی است، و $R$ حاصل‌ضربِ دکارتیِ $\mathbf{K}^2$ است. مدلِ $\varphi$ را با الزامِ این تعریف می‌کنیم که، به ازای هر فرمولِ اتمیِ $P$ و هر $\mathbf{H} \epsilon \mathbf{K}$، $\varphi(P, \mathbf{H})$ مساویِ $\mathbf{T}$ ($\mathbf{F}$) است اگر و تنها اگر $P$ در مدلِ $\mathbf{H}$ درست (غلط) باشد. (به یاد آورید که $P$ فرمولِ درست‌ساختی از $\mathbf{PA}$، و $\mathbf{H}$ یک مدلِ شمارای $\mathbf{PA}$ است.) بعد ارزیابی برای فرمول‌های مرکب را مثلِ قبل می‌سازیم [۱۶]. گفتنِ اینکه $A$ درست است گفتنِ این است که در جهانِ واقعِ $\mathbf{G}$ درست است؛ و، به ازای هر $P$ی اتمی، $\varphi(\Box P, \mathbf{G}) = \mathbf{T}$

1. detachment

اگر و تنها اگر $P$ در $\mathbf{PA}$ اثبات‌پذیر باشد. (توجه کنید که $\varphi(P,\mathbf{G})=\mathbf{T}$ اگر و تنها اگر $P$ به مفهومِ شهودیْ درست باشد.) چون $(\mathbf{G},\mathbf{K},R)$ یک S5ـس.م. است، همهٔ قوانینِ S5 در این تعبیر معتبر خواهند بود؛ و می‌توانیم نشان دهیم که فقط قوانین S5 عموماً معتبر خواهند بود. (مثلاً اگر $P$ فرمولِ تعیین‌ناپذیرِ گودل باشد، $\varphi(\Box P\vee\Box\sim P,\mathbf{G})=\mathbf{F}$، که مثالِ نقضی برای این «قانون» است که $\Box A\vee\Box\sim A$.)

تعبیرِ اثبات‌پذیریِ دیگر این است: باز هم فرمول‌های اتمی را فرمول‌های درست‌ساختِ بستهٔ $\mathbf{PA}$ می‌گیریم، و بعد با استفاده از اداتهای الحاقیِ $\wedge$، $\sim$، و $\Box$ فرمول‌های جدید می‌سازیم. $\mathbf{K}$ را مجموعهٔ همهٔ زوج‌های مرتبِ $(\mathbf{E},\alpha)$ بگیرید، که $\mathbf{E}$ توسیعِ سازگاری از $\mathbf{PA}$ است، و $\alpha$ مدلی (شمارا) از نظامِ $\mathbf{E}$. قرار دهید $\mathbf{G}=(\mathbf{PA},\alpha_0)$، که $\alpha_0$ مدلِ استانداردِ $\mathbf{PA}$ است. می‌گوییم $(\mathbf{E},\alpha)R(\mathbf{E}',\alpha')$، که در آن $(\mathbf{E},\alpha)$ و $(\mathbf{E}',\alpha')$ در $\mathbf{K}$اند، اگر و تنها اگر $\mathbf{E}'$ توسیعی از $\mathbf{E}$ باشد. برای $P$های اتمی، $\varphi(P,(\mathbf{E},\alpha))$ را $(\mathbf{F})\mathbf{T}$ تعریف کنید اگر و تنها اگر $P$ در $\alpha$ درست (غلط) باشد. در این صورت می‌توانیم نشان دهیم که، به ازای $P$های اتمی، $\varphi(\Box P,(\mathbf{E},\alpha))=\mathbf{T}$ اگر و تنها اگر $P$ در $\mathbf{E}$ اثبات‌پذیر باشد؛ به‌ویژه، $\varphi(\Box P,\mathbf{G})=\mathbf{T}$ اگر و تنها اگر $P$ در $\mathbf{PA}$ اثبات‌پذیر باشد. چون $(\mathbf{G},\mathbf{K},R)$ یک S4ـس.م. است، همهٔ قوانینِ S4 برقرارند. اما این‌طور نیست که همهٔ قوانینِ S5 برقرار باشند؛ اگر $P$ فرمولِ تعیین‌ناپذیرِ گودل باشد، $\varphi((\sim\Box P\supset\Box\sim\Box P),\mathbf{G})=\mathbf{F}$. اما بعضی قوانین معتبرند که در S4 اثبات‌پذیر نیستند؛ به‌ویژه، می‌توانیم اثبات کنیم که به ازای هر $A$، $\varphi(\Box\sim\Box(\Diamond A\wedge\Diamond\sim A),\mathbf{G})=\mathbf{T}$، که قضایای S4.1ـِ مک‌کینسی را به‌دست می‌دهد [۱۶]. با جرح و تعدیلاتِ مناسب می‌توان این مشکل را برطرف کرد؛ اما ما در اینجا واردِ این مطلب نمی‌شویم.

می‌توان تعبیراتِ مشابهی برای $M$ و نظامِ براوئری بیان کرد؛ اما، در نظرِ مؤلف، جذابیتِ اینها کمتر از آنهایی است که در بالا عرضه شد. یک ردهٔ دیگر از تعبیراتِ اثبات‌پذیری، توسیع‌های «انعکاسی»یِ $\mathbf{PA}$، را ذکر می‌کنیم. فرض کنید $\mathbf{E}$ نظامی صوری باشد که شاملِ $\mathbf{PA}$ است، و فرمول‌های بسته‌اش از فرمول‌های بستهٔ $\mathbf{PA}$ با استفاده از اداتهای $\&$، $\neg$، و $\Box$ ساخته شده‌اند. (می‌گویم '$\&$' و '$\neg$' تا مشخص کنم که دارم همان عاطف و ناقضِ خودِ $\mathbf{PA}$ را به‌کار می‌گیرم، نه اینکه عاطف و ناقضِ جدیدی معرفی کنم. یادداشتِ [۱۶] را ببینید.) در این صورت $\mathbf{E}$ یک توسیعِ انعکاسیِ $\mathbf{PA}$ خوانده می‌شود اگر و تنها اگر: (۱) توسیعی غیراساسی از $\mathbf{PA}$ باشد؛ (۲) $\Box A$ در $\mathbf{E}$ اثبات‌پذیر باشد اگر و تنها اگر $A$ باشد؛ (۳) ارزیابیِ $\alpha$ای باشد که فرمول‌های بستهٔ $\mathbf{E}$ را به توی مجموعهٔ $\{\mathbf{T},\mathbf{F}\}$ طوری بنگارد که عاطف و ناقض از جدول‌های صدقِ معمول تبعیت کنند، همهٔ فرمول‌های بستهٔ درستِ $\mathbf{PA}$ مقدارِ $\mathbf{T}$ بگیرند، $\alpha(\Box A)=\mathbf{T}$ اگر و تنها اگر $A$ در $\mathbf{E}$ اثبات‌پذیر باشد، و همهٔ قضایای $\mathbf{E}$ مقدارِ $\mathbf{T}$ بگیرند. می‌توان نشان داد که توسیع‌های انعکاسی‌ای از $\mathbf{PA}$ هستند که شاملِ اصولِ موضوعِ S4 یا حتی S4.1اند، اما هیچ توسیعِ انعکاسی‌ای شاملِ S5 نیست.

سرانجام، می‌گوییم که، با استفاده از نگاشتِ معمولِ منطقِ شهودگرایانه به توی S4، می‌توانیم نظریهٔ مدلی برای حسابِ محمولاتِ شهودگرایانه به‌دست آوریم. این نظریهٔ مدل را در اینجا نمی‌آوریم، اما در عوض، فقط در موردِ حسابِ گزاره‌ها، تعبیرِ مفیدِ ویژه‌ای از منطقِ شهودگرایانه را ذکر می‌کنیم که از این نظریهٔ مدلْ حاصل می‌شود. فرض کنید $\mathbf{E}$ توسیعی سازگار از $\mathbf{PA}$ باشد. می‌گوییم یک فرمولِ $P$یِ $\mathbf{PA}$ در $\mathbf{E}$ توجیه‌شده است اگر و تنها اگر در $\mathbf{E}$ اثبات‌پذیر باشد. فرمول‌های خوش‌ساختِ $P$یِ $\mathbf{PA}$ را اتمی می‌گیریم، و فرمول‌ها را از روی آنها با استفاده از اداتهای شهودگرایانهٔ $\wedge$، $\vee$، $\neg$، و $\supset$ می‌سازیم. بعد به‌طورِ استقرائی مقرر می‌کنیم: $A\wedge B$ در $\mathbf{E}$ توجیه‌شده است اگر و تنها اگر $A$ و $B$ باشند؛ $A\vee B$ در $\mathbf{E}$ توجیه‌شده است اگر و تنها اگر $A$ یا $B$ باشند؛ $\neg A$ در $\mathbf{E}$ توجیه‌شده است و اگر و تنها اگر هیچ توسیعِ سازگاری از $\mathbf{E}$ نباشد که $A$ را توجیه کند؛ $A\supset B$ در $\mathbf{E}$ توجیه شده است اگر و تنها اگر هر توسیعِ $\mathbf{E}'$ از $\mathbf{E}$ که $A$ را توجیه کند $B$ را هم توجیه کند.

در این صورت همهٔ نمونه‌های هر قانونِ منطقِ شهودگرایانه در $\mathbf{PA}$ توجیه شده است؛ اما، مثلاً، $A\vee\neg A$ نیست، اگر $A$ فرمولِ تعیین‌ناپذیرِ گودل باشد. در کارِ بعدی، این تعبیر را باز هم گسترش خواهیم داد، و با استفاده از آن نشان می‌دهیم که با استفاده از آن می‌توانیم تعبیری برای نظامِ FCیِ کرایزل برای دنباله‌های انتخابِ مطلقاً آزاد بیابیم [۱۷]. در ضمن، روشن است که در تعبیراتِ اثبات‌پذیریِ S4 و S5 می‌توان $\mathbf{PA}$ را با هر نظامِ قدرِ صدقیِ دیگری (یعنی با هر نظامی که مدل‌هایش هر فرمولِ بسته را به‌صورتِ درست یا غلط معلوم کنند) جایگزین کرد؛ اما این تعبیرِ شهودگرایی بر هر گونه نظامِ صوری‌ای قابلِ اِعمال است.

### افزوده

من دیگر نمی‌توانم بنویسم «هومز وجود ندارد، اما ۱و در اوضاع و احوالِ دیگری وجود می‌داشت». دیگر به نظرم نمی‌آید که نام‌های تخیلی‌ای از قبیلِ «شرلوک هومز» هویّاتِ خاصِ ممکن‌ـولی‌ـناموجودی را نام‌گذاری کنند که در شرایطی وجود می‌داشتند. البته می‌شد شخصی در قرنِ نوزدهم وجود داشته باشد که ماجراجویی‌هایی از آن قبیل که در داستان‌های هومز توصیف شده را انجام داده باشد. هر شخصِ واقعیِ آن دوره (مثلاً داروین) ممکن بود چنان کرده باشد، گرچه بر آنام که هیچ کسی نکرده است؛ یا، به صورتی دیگر، می‌شد شخص (یا اشخاصِ) دیگری، در $\psi(\mathbf{H})$ اما نه در $\psi(\mathbf{G})$، متولد شده باش(ن)د و اَعمالِ هومزواری داشته بوده باشند. اما حق نداریم هیچ چنین هویّتِ خاصی را «شرلوک هومز» بخوانیم. حکمِ «شرلوک هوُمز می‌شد وجود داشته باشد» اکنون به نظرِ من بیجا می‌آید.

این تغییر بر نظرِ من در موردِ وضعِ زبان‌شناختیِ نام‌های تخیلی در زبانِ معمولی اثر می‌گذارد ولی بر مطالبِ مدل‌نگریکِ داخلِ متن اثر نمی‌گذارد، یعنی: (۱) برخی هویّات که واقعاً وجود دارند می‌شد وجود نداشته باشند، و می‌شد هویّاتی باشند غیر از آنهایی که واقعاً وجود دارند، پس لازم نیست $\psi(\mathbf{H})$ به‌ازای همهٔ $\mathbf{H}\epsilon\mathbf{K}$ها ثابت باشد. (۲) با توجه به (۱)، اگر به متغیرِ '$x$' یک هویّتِ $a$ تخصیص داده شود، به‌نحوی که $a\in\psi(\mathbf{H}_2)$، $a\notin\psi(\mathbf{H}_2)$، آیا باید ــ نسبت به این تخصیص به $x$ ــ به $\varphi(P(x),\mathbf{H}_2)$ ارزشی بدهیم؟ به‌کارگرفتنِ نام‌های تخیلی برای روشن ساختنِ این نکاتْ اتفاقی بود. واضح است که نمی‌توانم در موردِ این موضوعِ زبان‌شناختیْ تفصیل دهم، گرچه به برخی مسائلِ فلسفی در موردِ وضعِ «هویّاتِ ممکنِ محقق‌نشده» مربوط است.

۱۱ ژانویهٔ ۱۹۷۱

## یادداشتها

۱. نظریه‌ای که در اینجا داده شده است با نظریه‌های مؤلفانِ زیادی مرتبط است: برای فهرست‌هایی از اینها نگاه کنید به

S. Kripke, 'Semantical Analysis of Modal Logic', *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9 (1963), 67-96.

J. Hintikka, 'Modality and Quantification', *Theoria*, 27 (1961), 119-28.

مؤلفانی که به نظریهٔ حاضر نزدیک‌ترین‌اند به نظر می‌رسد هینتیکا وکانگر [Kanger] باشند. با این حال، تا آنجا که من خبر دارم، این نحوهٔ برخورد با موضوع منحصر به فرد است، اگرچه آشنایی با روش‌های بسیار متفاوتِ هینتیکا و پرایر [Prior] برای آن منابعِ الهام بوده است.

**2**. 'A Completeness Theorem in Modal Logic', *Journal of Symbolic Logic*, 24 (1959), 1-15.

**3**. 'Semantical Analysis of Modal Logic', Ibid., pp. 323-4 (Abstract)

۴. در موردِ اینها نگاه کنید به

'A Completeness Theorem in Modal Logic', *Journal of Symbolic Logic*, 24 (1959), 1-15

و

'Semantical Analysis of Modal Logic', *Zeitschrift für Mathematische Logik und grundlagen der Mathematik*, 9, 67-96.

۵. برای اثبات، نگاه کنید به

'Semantical Analysis...', Zeitschrift...,.

**6**. G. Frege, 'Uber Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik*, 100 (1892), 25-50.

ترجمه‌های انگلیسی در

Geach and Black, *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell, 1952),

و در

Feigel and Sellars (eds.), *Readings in Philosophical Analysis* (New York: Appleton Century Crofts, 1949).

**7**. P. F. Strawson, 'On referring', *Mind*, n.s., 59 (1950), 320-44.

[ترجمهٔ فارسی: پ. ف. استراوسون، «پیرامون اشاره»، ترجمهٔ رضا محمدزاده، ارغنون، سالِ دوم، شمارهٔ ۷ و ۸ (پاییز و زمستانِ ۱۳۷۴)، ۳۲۹-۲۸۹-م.]

**8**. Bertrand Russell, 'On denoting', *Mind*, n.s., 14 (1905), 479-93.

**9**. 'Modality and Quantification'

**10**. A. N. Prior, *Time and Modality* (Oxford: Clarendon Press, 1957, VIII+148 pp).

۱۱. طبیعی است که فرض کنیم هر محمولِ n‌ـ‌تایی باید در هر جهانِ $\mathbf{H}$ در موردِ همهٔ افرادی که در آن جهان وجود ندارند غلط باشد؛ یعنی اینکه مصداقِ هر حرفِ محمولی باید از افرادِ واقعاً موجود تشکیل شود. این کار را می‌توانیم به این صورت انجام دهیم که به‌صورتِ معناشناختی قید کنیم که $\varphi(P^n,\mathbf{H})$ زیرمجموعه‌ای از $[\psi(\mathbf{H})]^n$ باشد؛ برخوردِ معناشناختیِ ذیل از جنبه‌های دیگر بدونِ تغییر کفایت خواهد کرد. مجبوریم به نظامِ اصولِ موضوعِ ذیل همهٔ بسترهای فرمول‌های به شکلِ $P^n(x_1,\ldots,x_n) \wedge (y)A(y). \supset .A(x_i)$، $1 \le i \le n$، را اضافه کنیم. ترجیح داده‌ایم این‌کار را نکنیم چون قاعدهٔ جایگزینی دیگر برقرار نخواهد بود؛ قضایایی در موردِ فرمول‌های اتمی برقرار خواهند شد که وقتی فرمول‌های اتمی با فرمول‌های دلخواه جایگزین شده باشند برقرار نخواهند بود. (این به یک پرسشِ پاتنم [Putnam] و کالمر [Kalmar] پاسخ می‌دهد.

۱۲. نگاه کنید به

'Modality and Quantification in s5', *Journal of Symbolic Logics*, 21 (1956), 60-2.

۱۳. ادعا نشده است که تعبیرِ عمومیت برای قضایای با متغیرهای آزاد یگانه تعبیرِ ممکن است. ممکن است بخواهیم که فرمولِ $A$ اثبات‌پذیر باشد اگر و تنها اگر، به ازای هر $\varphi$، به ازای هر تخصیص به متغیرهای آزادِ $A$، $\varphi(A,\mathbf{G}) = \mathbf{T}$. اما در این صورت $(x)A(x) \supset A(y)$ قضیه نخواهد بود؛ درواقع، در پادمُدلی که در بالا برای فرمولِ بارکان دادیم اگر $b$ به $y$ تخصیص داده شود، $\varphi((x)P(x) \supset P(y),\mathbf{G})] = \mathbf{F}$. بدین ترتیب نظریهٔ تسویر باید به شیوه‌هایی که در

Hintikka, 'Existential Presuppositions and Existential Commitments', *Journal of Philosophy*, 56 (1959), 125-37.

و

H. Leblanc and T. Hailperin, 'Nondesignating Singular Terms', *Philosophical Review*, 68 (1959), 239-45

پیشنهاد شده بازنگری شود. این فرآیند چیزهایی زیادی دارد که قابلِ قبول‌اش می‌سازد، اما ما آن را برنگزیده‌ایم چون می‌خواستیم نشان دهیم که مشکل را بدونِ بازنگری در نظریهٔ تسویر یا منطقِ گزاره‌ایِ وجهی هم می‌توان حل کرد.

**14**. W. Quine, *Mathematical Logic* (Cambridge, Mass: Harvard Univ. Press, 1940; 2nd ed., rev., 1951, XII+346 pp.).

**15**. S. C. Kleene, *Introduction to Metamathematics* (New York: D. van Nostrand, 1952, x+550 pp.)

۱۶. ممکن است اعتراض شود که **PA** نقداً نمادهایی برای عاطف و ناقض دارد، مثلاً '&' و '¬'؛ پس چرا نمادهای جدیدِ '∧' و '∼' را الحاق می‌کنیم؟ جوابْ این است که اگر $P$ و $Q$ فرمول‌های اتمی‌ای باشند، آنگاه $P\&Q$ نیز به مفهومِ فعلی اتمی است، چرا که در **PA** خوش‌ساخت است؛ اما $P \wedge Q$ این‌طور نیست. برای آنکه بتوانیم نظریهٔ قبلی را ــ که در آن ترکیبِ عطفیِ فرمول‌های اتمی اتمی نیست ــ به‌کار گیریم، '∧' را لازم داریم. با این حال، به ازای هر $\mathbf{H}\epsilon\mathbf{K}$ و $P$ و $Q$ی اتمی، $\varphi(P\&Q,\mathbf{H}) = \varphi(P \wedge Q,\mathbf{H})$، و لذا خلطِ & با '∧' در عمل باعثِ ضرری نمی‌شود. نکاتِ مشابهی در موردِ ناقض، و در موردِ تعبیرِ اثبات‌پذیریِ S4 در پاراگرافِ بعدی موضوعیت دارند.

۱۷. نگاه کنید به

J. C. C. Mckinsey, 'On the Syntactical Construction of Systems of Modal Logic', *Journal of Symbolic Logic*, 10 (1945), 83-94.

**18**. G. Kreisel, 'A Remark on Free Choice Sequences and the Topological Completeness Proofs', *Journal of Symbolic Logic*, 23 (1958), 369-88.

ترجمهٔ ک. ل.

# ریاضیات، آمار، و تدریس[•]

جورج کاب، دیوید مور*

ترجمهٔ مجتبی گنجعلی، محمدقاسم وحیدی‌اصل

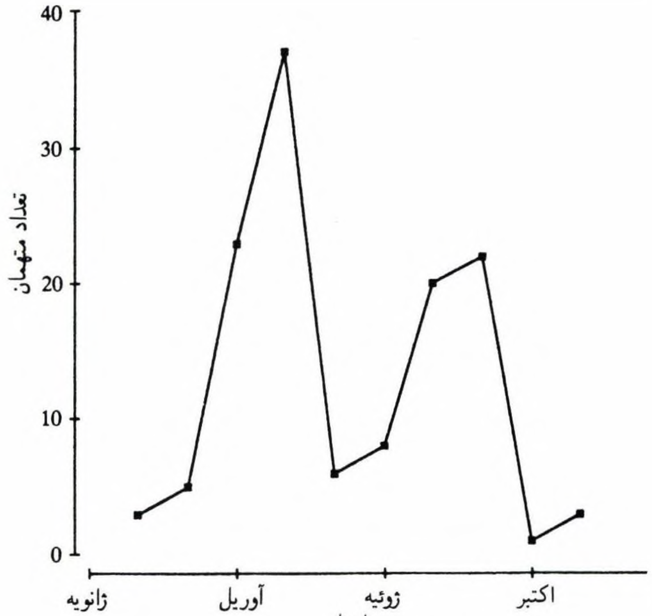
تفکر آماری چه تفاوتی با تفکر ریاضی دارد؟ نقش ریاضیات در آمار چیست؟ اگر از آمار محتوای ریاضی آن را بزدایید، چه جوهرهٔ عقلانی از آن باقی می‌ماند؟

در آنچه در زیر می‌آید، پاسخهایی به این پرسشها ارائه کرده آنها را به دنباله‌ای از مثالها ربط می‌دهیم که چشم‌اندازی از شیوهٔ عمل رایج آماری به‌دست می‌دهند. ضمن این کار، و مخصوصاً در پایان مقاله، به برخی از نتایج بحث برای تدریس آمار اشاره می‌کنیم.

**۱. مقدمه: نظری اجمالی به تفکر آماری.** آمار شاخه‌ای روش‌شناختی از دانش است. وجود آن برای خودش نیست بلکه برای این است که مجموعه‌ای منسجم از ایده‌ها و ابزارها برای کار با داده‌ها به شاخه‌های علمی دیگر ارائه کند. نیاز به چنین شاخه‌ای از دانش ناشی از «همه جا حاضر بودن تغییرپذیری»[1] است. افراد با هم تفاوت دارند. نتایج اندازه‌گیریهای مکرر روی فردی واحد، یکسان نیستند. در برخی موقعیتها، می‌خواهیم موارد غیرعادی را در توده‌ای انبوه از داده‌ها پیدا کنیم. در موقعیتهایی دیگر، تمرکز روی تغییر نتایج اندازه‌گیریهاست، و گاه می‌خواهیم اثرات نظام‌مند را در برابر نوفهٔ زمینه‌ای مربوط به تغییرات فردی آشکار سازیم. آمار ابزارهایی برای کار با داده‌ها به‌دست می‌دهد که همه جا حاضربودن تغییرپذیری در آنها ملحوظ شده است.

**۱.۱ نقش مضمون.** تمرکز بر روی تغییرپذیری، به‌طور طبیعی محتوای خاصی به آمار می‌دهد که آن‌را از خود ریاضیات و دیگر علوم ریاضی مجزا می‌سازد، اما چیزی بیش از محتوای صِرف است که تفکر آماری را از تفکر ریاضی متمایز می‌کند. آمار نیازمند نوع متفاوتی از تفکر است، زیرا داده‌ها فقط اعداد نیستند، آنها اعدادی همراه با مضمون هستند.

1. omnipresence of variability




**شکل۱** تعداد افرادی که در ناحیهٔ اسکس[1] ماساچوست متهم به جادوگری شده‌اند.

مثال ۱، رازِ اندوور[2]. وقتی نمودار دنبالهٔ متناهی

$$(۳, ۵, ۲۳, ۳۷, ۶, ۸, ۲۵, ۲۲, ۱, ۳)$$

رسم شود (شکل ۱) الگوی مشخصی را نشان می‌دهد، اما این اعداد و الگو معنی و اهمیتی ندارند مگر آنکه مضمون آنها را بدانیم. در حقیقت این اعداد تعداد کل ماهانه افرادی هستند که از فوریهٔ ۱۶۹۲ تا چند ماه بعد در

1. Essex 2. Andover

ناحیهٔ اسکس ایالت ماساچوست، رسماً متهم به جادوگری شده‌اند. نمودار شکل ۱ دو موج از اتهام‌زدن‌ها را نشان می‌دهد که با یک نقطهٔ حضیض در تابستان ۱۶۹۲ از هم جدا شده‌اند. الگوی شکل ۱ معنی‌دارتر می‌شود وقتی‌که بدانیم دار زدن اولین جادوگر محکوم شده (بریجت بیشاپ[1]) در ۱۰ ژوئن ۱۶۹۲ صورت گرفته است. تصور اثر هشیارکنندهٔ اولین اعدام در جامعهٔ کوچک دهکدهٔ سیلم[2] (حالا دنورز[3]) سخت نیست. اما دومین موج اتهامات چرا؟ چنین معلوم می‌شود که اتهام‌زدن‌های موج اول علیه ساکنان دهکدهٔ سیلم، شهر سیلم، و هر شش شهر مجاور بجز یکی از آنها صورت گرفته است؛ در موج دوم، اکثریت اتهام‌زدن‌ها علیه ساکنان یکی دیگر از شهرهای مجاور، یعنی اندوور صورت گرفته است. منابع ما [۳ و ۴] توضیح زیادی در مورد آنچه در شهر اندوور اتفاق افتاده نمی‌دهند، ولی این الگو همراه با آنچه از مضمون می‌دانیم حداقل بخشی از ماجرا را حکایت و پرسشهای جالبی را مطرح می‌کند.

هر چند این مثالِ نخست تقریباً هیچ محتوای ریاضی ندارد، ملاحظهٔ تأثیر متقابل بین الگو و مضمون آن، نمونهٔ بخش تفسیری تفکر آماری است. برای مثالی آشناتر از نوعی بسیار متفاوت، آزمون برابری میانگینهای دو توزیع نرمال را در نظر بگیرید.

مثال ۲ الف، مدلی برای مقایسهٔ میانگینهای نرمال. مدل استانداردی متضمن دو مجموعه از متغیرهای تصادفی مستقل همتوزیع (iid) را در نظر بگیرید:

$$X_1, X_2, \cdots X_n \text{ iid } N(\mu_1, \sigma_1) \quad Y_1, Y_2, \cdots, Y_m \text{ iid } N(\mu_2, \sigma_2)$$

نتیجه می‌شود که $\overline{x} = \frac{\sum x_i}{n}$ و $s_1^2 = \frac{\sum (x_i - \overline{x})^2}{n-1}$ آماره‌های بسنده برای $\mu_1$ و $\sigma_1^2$ هستند. نتیجه‌ای مشابه برای $Y$ها برقرار است. به‌طور غیررسمی، آماره‌ای برای یک پارامتر، بسنده است اگر از تمام اطلاعات در بارهٔ آن پارامتر که در نمونه موجود است، استفاده کند. به‌طور رسمیتر، توزیع شرطی داده‌ها، به شرط آمارهٔ بسنده به پارامتر وابسته نیست. قضیهٔ رائوـبلک‌ول[4] تضمین می‌کند که هیچ برآوردگر نااریبی نمی‌تواند واریانسی کوچکتر از برآوردگر نااریب مبتنی بر آمارهٔ بسنده داشته باشد. هر دوی $\overline{x}$ و $s_1^2$ نااریب‌اند: $E(\overline{x}) = \mu_1$ و $E(s_1^2) = \sigma_1^2$. و بالاخره توزیع توأم آنها معلوم است: میانگین نمونه‌ای $\overline{x}$ دارای توزیع نرمال با واریانس $\frac{\sigma_1^2}{n}$ است، و مستقل از آن، $\frac{(n-1)s_1^2}{\sigma_1^2}$ دارای توزیع خی دو با $(n-1)$ درجهٔ آزادی است. حال فرض کنید که می‌خواهیم فرض $H_0 : \mu_1 = \mu_2$ را آزمون کنیم. اگر $\sigma_1^2 = \sigma_2^2$ آنگاه برآوردگری بسنده و نااریب برای واریانس مشترک از طریق ادغام کردن به‌دست می‌آید:

$$s_p^2 = [(n-1)s_1^2 + (m-1)s_2^2]/(n+m-2)$$

اگر $H_0$ درست باشد، آنگاه $\frac{(\overline{x}-\overline{y})}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$ دارای توزیع $t$ی استیودنت با $n+m-2$ درجهٔ آزادی است، و می‌توانیم از مقدار $t$ که از روی داده‌ها محاسبه شده برای آزمون فرض صفر استفاده کنیم. اگر $t$ از ۰ به‌اندازهٔ کافی دور باشد، نتیجه می‌گیریم که $\mu_1 \neq \mu_2$.

1. Bridget Bishop　2. Salem Village　3. Danvers

4. Rao-Blackwell

این مثال از دو لحاظ تفاوت بارزی با مثال اول دارد: محتوای ریاضی و نقش مضمون. مثال ۱ که اساساً هیچ محتوای ریاضی ندارد، جوهرهٔ عقلانی‌اش تقریباً به‌طور کامل از تأثیر متقابل الگو و ماجرا ناشی می‌شود. مثال ۲، که اساساً هیچ محتوای غیرریاضی ندارد، جوهرهٔ عقلانی‌اش را بدون رجوع صریح به مضمون کاربردی به‌دست می‌آورد.

هر چند ریاضیدانان اغلب بر مضمون کاربردی، هم برای ایجاد انگیزه و هم به‌عنوان منبعی از مسائل تحقیقاتی، تکیه دارند، کانون توجه نهایی در تفکر ریاضی، الگوهای مجرد است: مضمون بخشی از جزئیات نامربوطی است که باید روی شعلهٔ تجرید گداخته شوند تا بلور ساختار نابی که قبلاً ناپیدا بوده آشکار شود. **در ریاضیات، مضمون، ساختار را در پردهٔ ابهام نمی‌افکند.** تحلیلگران داده‌ها نیز مانند ریاضیدانان در پی یافتن الگوها هستند، اما در تحلیل داده‌ها در نهایت، اینکه الگوها دارای معنی و ارزشمند باشند، بستگی به این دارد که چگونه تار این الگوها با پود ماجرا به‌هم بافته شوند. **در تحلیل داده‌ها، مضمون است که معنا می‌بخشد.**

این تفاوت، پیامدهای ژرفی در تدریس دارد. برای تدریسِ خوب آمار، فهمیدن نظریهٔ ریاضی کافی نیست؛ حتی کافی نیست که علاوه بر آن، قسمت غیرریاضی آمار را فهمیده باشیم. لازم است مانند معلم ادبیات، ذخیره‌ای آماده از مثالهای واقعی داشته باشیم، و بدانیم چگونه از آنها استفاده کنیم تا دانشجو قدرت داوری نقادانه پیدا کند. در ریاضیات، که در آن مضمون کاربردی از اهمیت بسیارکمتری برخوردار است، مثالهای فی‌البداهه به‌خوبی نیاز را رفع می‌کند و معلمان ریاضی در ابداع فی‌المجلسِ مثالها مهارت به‌دست می‌آورند (به تابعی برای نمایش قاعده زنجیری نیاز دارید؟ مسأله‌ای نیست: فوراً مثالی بسازید). ولی در آمار مثالهای فی‌البداهه به درد نمی‌خورند، زیرا تأثیر متقابل الگو و مضمون را به‌طور مؤثق نشان نمی‌دهند. همان‌طور که برتراند راسل ریاضیات را به‌علت حالت تجریدی خشک و بی‌روحش به مجسمه‌سازی تشبیه می‌کند، شاید بتوان تحلیل‌داده‌ها را که در آن الگو و مضمون جداشدنی نیستند به فن شعر تشبیه کرد. تصور کنید که درسی مقدماتی در عروض تدریس می‌کنید و می‌خواهید مصرع شش رکنی داکتیلی[1] را معرفی کنید. برای این‌کار کافی نیست بگویید «TA ta ta، TA ta ta، TA ta ta ...» بلکه لازم است شاگردان شما داکتیل[2]ها را در یک شعر واقعی بشنوند [۲۵]:

This is the forest primeval. The murmuring pines and the hemlocks.

[این جنگل روزگاران آغازین است. کاجها و صنوبران در نجوا][3]. همین‌طور معلم آمار لازم است دنیای داده‌ها را بشناسد. اگر، برای مثال، وقتی نمودارهای

1. dactylic hexameter

۲. dactyl، (در شعر) کلمه‌ای دارای سه هجا که دو هجای آن کوتاه و یک هجای آن بلند باشد.ـم.

۳. مصرع آغازین شعر حماسی Evangeline از هنری وودزورث (Wadsworth).

برای خواننده شاید بهتر باشد مثالی از اوزان شعر فارسی بیاوریم:

وقتی می‌خواهید «بحر متقارب مثمن سالم» را در درس عروض به شاگردان معرفی کنید، کافی نیست بگویید «فعولن فعولن فعولن فعولن» بلکه شاگردان شما باید این وزن را در یک شعر واقعی احساس کنند:

درخت تو گر بار دانش بگیرد　　به زیر آوری چرخ نیلوفری را

توزیع داده‌ها را تدریس می‌کنید از داده‌های مربوط به مدت زمان بین دو فورانِ اولد فیت‌فول[1] [۳۰] و طول مدت سلطنت پادشاهان و ملکه‌های انگلیس [۱۳] استفاده کنید، دانشجویان شما می‌توانند چیزی بیش از خود روشها فرا گیرند. شکل دو مدی زمانهای بین دو فوران، دو نوع فوران را به ذهن متبادر می‌کند و توزیع سلطنت پادشاهان چاولگی به سمت مقادیر بالا را که در زمانهای انتظار معمول است، نشان می‌دهد.

نقشهای متفاوت مضمون در ریاضیات و آمار، مخصوصاً به‌صورتی که در دو مثال افراطی نخستین نشان داده شد، ممکن است مؤید استنتاج نادرستی به‌نظر آید که در این گفتهٔ پولاک دیده می‌شود: «آماردانان زیادی اکنون مدعی‌اند که آمار چیزی کاملاً جدا از ریاضیات است، به‌طوری که درسهای آماری نیازی به هیچ نوع آمادگی در ریاضیات ندارند.» هر چند اقناع و اثبات در آمار از نوع ریاضی نیست ([۲۲] و [۲۴] را ببینید)، همهٔ درسهای آماری نیازمند قدری، و برخی از آنها نیازمند مقدار زیادی آمادگی در ریاضیات‌اند. نظریه‌های ریاضی بغرنج، داربست بخشهایی از آمارند، و مطالعهٔ این نظریه‌ها بخشی از تعلیماتی است که آماردانان می‌بینند. اما، گرچه آمار بدون ریاضیات نمی‌تواند شکوفا شود، عکس آن نادرست است. اینکه آمار بخشی ضروری از تعلیمات ریاضیدانان نیست، به‌طور تلویحی از این گفتهٔ احتمال‌دان برجسته، دیوید آلدوس [۱] برمی‌آید که وی «به کاربردهای احتمال در همهٔ حوزه‌های علمی بجز آمار علاقه‌مند است».

بنابراین، به نقش ریاضیات در علم آمار چیست؟ پاسخ باید با نگاهی نظام‌مندتر به منطق تحلیل داده‌ها آغاز شود.

**۲.۱ مرور کلی نموداری بر تحلیل آماری.** درسی به سبک قدیم که پایبند به عرضهٔ کاربردها باشد، شاید مثال دوم را با یک تمرین کاربردی به پایان برساند. داده‌های زیر از [۲۵] گرفته شده‌اند هر چند خود تمرینْ ساختگی است؛ کل بررسی در [۲۱] توصیف شده است.

مثال ۲ ب. کلسیم و فشار خون. آیا افزایش مقدار کلسیم در رژیم غذایی ما، فشار خون را کاهش می‌دهد؟ اعداد زیر مقدار کاهش فشار خون سیستولیک را بعد از ۱۲ هفته برای ۲۱ فرد انسانی در اختیار ما قرار می‌دهد. ۱۰ فرد در گروه ۱ به‌مدت ۱۲ هفته یک مکمل کلسیم را مصرف کردند؛ ۱۱ فرد در گروه ۲ از یک دارونما[2] استفاده کردند. این فرض را که کلسیم تأثیری روی فشار خون نداشته است آزمون کنید.

گروه ۱ (کلسیم): ۷، ۴-، ۱۸، ۱۷، ۳-، ۵-، ۱، ۱۰، ۱۱، ۲-

گروه ۲ (دارونما): ۱-، ۱۲، ۱-، ۳-، ۳، ۵-، ۵، ۲، ۱۱-، ۱-، ۳-

این تمرین، اگر آن‌را به اختصار توصیف کنیم، کاریکاتوری بیش نیست، کاریکاتوری که این دیدگاه خطا را موجب می‌شود که به‌محض اینکه استنتاجات ریاضی از یک مدل تکمیل شدند، کاربردها اساساً چیزی نیستند جز محاسباتی

الگوها $\xrightarrow{(۲)}$ داده‌ها $\xrightarrow{(۱)}$ طرح

تعبیر $\xrightarrow{(۵)}$ نتایج $\xrightarrow{(۴)}$ روشها $\xrightarrow{(۳)}$ مدلها

**شکل ۲** نمایش نموداری مراحل تولید و تحلیل داده‌ها

معمولی. برای چشم‌اندازی واقعیتر، شکل ۲ را که نموداری از مراحل تحلیل آماری است، در نظر گیرید. قبل از بررسی جزئیات این طرح خام توجه به دو نکته ضروری است.

۱. این طرح خلاصه‌وار که پیشروی اکید از چپ به راست را القا می‌کند، بیش از اندازه ساده‌گرایانه است. در عالم واقع، فرایند تحلیل داده‌ها نه خطی و نه یکسویه است. درگذارهای مختلف بین مراحل، گاه دو مرحلهٔ متوالی و گاه حتی بیش از دو مرحله، بر یکدیگر تأثیر متقابل دارند. مثلاً هر چند انتخاب طرح برای تولید داده‌ها، ساختار داده‌های حاصل را تعیین می‌کند، اما داشتن اطلاعاتی براساس داده‌هایی که از قبل در دست است، می‌تواند به شکل‌گیری طرح کمک کند، همان‌گونه که اطلاع از اندازهٔ تغییرات از یک مورد به مورد دیگر به تصمیم‌گیری در بارهٔ اینکه چند مورد آزمایش لازم است کمک می‌کند. همچنین داده‌ها ممکن است مدلی را القا کنند، اما این مدل به روشهایی منجر شود که ما را دوباره به سراغ داده‌ها بفرستند تا تخلفات احتمالی از مفروضات مدل را امتحان کنیم. شاید، همان‌طور که خواهیم دید، مرحلهٔ نهاییِ تفسیر نتایج که از همه مهمتر است به‌طور حیاتی به مرحلهٔ اول که نوع طرحِ مورد استفاده در تولید داده‌هاست وابسته باشد.

۲. ارائهٔ این ترتیب سردستی و محدود از مراحل برای القای این نظر نیست که مباحثی که در یک درس آمار مقدماتی تدریس می‌شوند باید همین ترتیب را دنبال کنند. به دلایلی که بعداً می‌آوریم، توصیهٔ ما آغاز کردن چنین درسی با روشهای کاوش و توصیف داده‌ها، سپس بازگشت به تولید داده‌ها و از آنجا رفتن به‌سوی استنباط رسمی است.

با فرض رعایت شدن این هشدارها، این روند نما می‌تواند چارچوب مفیدی برای بازبینی نقش ریاضیات در آمار باشد و نیز عناصر جوهرهٔ غیرریاضی آمار را به اختصار بنمایاند. در این زمینه به چهار نکتهٔ واضح اشاره می‌کنیم:

۱. طرح، کاوش، و تفسیر، اجزای مرکزی تفکر آماری هستند. هر سهٔ این اجزا به‌شدت به مضمون وابسته‌اند، ولی در سطح مقدماتی متضمن ریاضیات بسیار کمی هستند. نظریهٔ (عمدتاً غیرریاضی) طرح آزمایشها چندین دهه عمر دارد و بسیار رشد کرده است. نظریهٔ کاوش جدیدتر است، و در حال حاضر هنوز جنبهٔ ابتدایی دارد، گرچه ابزارهای مبتنی بر کامپیوترِ کاوش پیشرفته‌اند؛ نظریهٔ تفسیر را در بهترین حالت می‌توان نامنسجم توصیف کرد.

۲. درس کلاسیک آمار ریاضی چنان به‌خوبی با گذار (۳) از مدلها به روشها متناظر است که «از مدلها به روشها» شاید کم‌وبیش بتواند عنوان یک درس باشد. مضمون در اینجا چندان ذی‌ربط نیست زیرا مدلها مانند مدل مثال ۲ الف به‌طور مجرد مطرح می‌شوند و معمولاً در نتیجه‌گیریها

---

۱. Old Faithful، آبفشانی در پارک طبیعی هوستون آمریکا، در سال ۱۹۷۴ میلادی، ۳۰۳۴ زمان بین دو فوران آن اندازه‌گیری و معلوم شد که میانگین مدت زمان بین دو فوران ۶۶٫۵ دقیقه، کوتاهترین زمان ۳۶، و طولانیترین آن ۱۰۲ دقیقه بوده است. از ۵۰۰۰۰ فوران مشاهده و ثبت‌شده در ۱۰۵ سال گذشته، میانگین زمان بین فورانها همیشه بین ۶۰ و ۶۷ دقیقه، کمترین زمان ۳۳ دقیقه، و بیشترین آن ۱۰۲ دقیقه بوده است.ـم.

2. placebo

از یکی از اصول بهینه‌سازی (کمترین مربعات، ماکسیمم درستنمایی) برای استنتاج رسمی روش استفاده می‌شود.

۳. گذار (۴) از روشها به نتایج، کانون توجه درسهای سبک قدیمی است که مانند کتابهای آشپزی تنظیم می‌شوند و در آنها هر روش در مجموعه‌ای از فرمولها خلاصه می‌شود. مضمون در اینجا نیز بی‌ربط است، از این نظر که می‌توانید الگوریتمهای محاسباتی را فرا گیرید، و درواقع آنها را بهتر بیاموزید، اگر در هیچ‌گونه دغدغهٔ خاطر در مورد اینکه این روشها به چه درد می‌خورند، به خود راه ندهید. با این حال، در برخی درسها تلاش شده است این قرص بزرگ گلوگیر را با لعابی نازک از مضمون بدلی آسانتر آمادهٔ بلع کنند. خوشبختانه، کامپیوتر این‌گونه درسها را در زباله‌دانی تاریخ برنامه‌های درسی می‌افکند.

۴. شاید طنزآمیز باشد که گذارهای (۳) و (۴) که اغلب مرکز توجه درسهای مقدماتی بوده‌اند، دقیقاً دو موردی هستند که از نظر ذهنی سرراست‌تر و مکانیکی‌تر از بقیه‌اند (با توجه به درک محدود فعلی و نظریهٔ کمتر پیشرفتهٔ گذارهای دیگر) و اغلب کمترین امکان را برای داوری و خلاقیت عرضه می‌کنند.

برای بسط این نکات با تفصیل بیشتر، به مثال کلسیم و فشار خون برمی‌گردیم. در آنچه خواهد آمد، مراحل شکل (۲) را زیر سه عنوان کلیتر ترکیب می‌کنیم: تولید داده‌ها، تحلیل داده‌ها، و استنباط.

## ۲. محتوای آمار

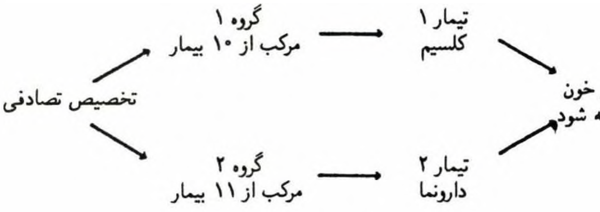
**۱.۲ تولید داده‌ها.** مدل استاندارد در مثال ۲ الف نقصی بسیار جدی دارد؛ این مدل تمایزی بین داده‌های مشاهده‌ای (مثلاً داده‌های یک بررسی نمونه‌ای) و داده‌های یک آزمایش مقایسه‌ای تصادفی‌شده قائل نمی‌شود. این تمایز بین مشاهده و آزمایش، یکی از مهمترین تمایزات در آمار است. پژوهشگران اغلب می‌خواهند به نتایج علّی برسند، مثلاً اینکه کلسیم علت کاهش در فشار خون است. آزمایشها اغلب اجازهٔ نتیجه‌گیریهای علّی را می‌دهند، در حالی‌که مطالعات مشاهده‌ای تقریباً همواره مسائل علیّت را حل‌وفصل نشده و در معرض مناقشه باقی می‌گذارند. با این‌حال مدلهای ریاضی نظریهٔ آمار برای داده‌های مشاهده‌ای و آزمایشی یکسان‌اند.

مطالعهٔ کلسیم درواقع یک آزمایش بود:

مثال ۲ پ، طرح بررسی کلسیم [۲۱]. بازبینی نمونه‌ای بزرگ از اشخاص رابطه‌ای بین مصرف کلسیم و فشار خون را آشکار کرد. این رابطه برای مردان سیاهپوست قویتر از دیگران بود. به همین علت پژوهشگران آزمایشی را به اجرا درآوردند.

آزمودنیهای آزمایش، ۲۱ مرد سیاهپوست سالم بودند. یک گروه ۱۰ نفره از این مردان که به‌طور تصادفی انتخاب شده بودند یک مکمل کلسیم مصرف کردند. گروه شاهد مرکب از ۱۱ مرد، یک قرص دارونما را که شبیه مکمل کلسیم بود، دریافت کردند. آزمایش از دو طرف کور[۱] بود.

۱. double_blind. در این آزمایش نه آزمودنیها و نه پزشکان همکار طرح خبر ندارند که چه کسی دارو و چه کسی دارونما مصرف کرده است.ـم.




**شکل۳** ساده‌ترین آزمایش مقایسه‌ای تصادفی‌شده

آیا می‌توانیم نتیجه بگیریم که کلسیم علت کاهش در فشار خون است؟ چنین استنباطی، که یک اختلاف مشاهده شده را می‌توان به‌معنای ظاهریش پذیرفت، بر سه پایه استوار است که دو تا از آنها بر تولید داده‌ها متکی هستند:

(۱) استدلالی ــ که تنها برای نمونه‌های تصادفی و آزمایشهای تصادفی‌شده جنبهٔ عادی و سرراست دارد ــ مبتنی بر اینکه مدلی احتمالاتی در مورد داده‌ها قابلِ اِعمال است.

(۲) استدلالی ــ مبتنی بر احتمال و نسبتاً سرراست ــ مبنی بر اینکه اختلاف مشاهده شده «واقعی» است، یعنی بزرگتر از آن است که انتساب آن به تغییرات شانسی موجه باشد.

(۳) استدلالی ــ که بجز در حالت آزمایشهای تصادفی‌شده اغلب پرزحمت، و امکان لغزش در آن زیاد است ــ مبنی بر اینکه اختلاف مشاهده شده ناشی از تأثیر یک عامل اختلاطی مجزا از عاملِ موردِنظر نیست.

آزمون $t$ در مثال ۲ الف مانند همهٔ آزمونهای آماری و بازه‌های اطمینان آماری تنها به استدلال دوم می‌پردازد: «اگر فرض کنیم که مدل شانسی خاصی قابل به‌کارگیری است، چقدر محتمل است که اختلاف قابل مشاهده‌ای به این بزرگی به‌دست آوریم؟» دو استدلال دیگر به طرح وابسته هستند.

آزمایش بالینی در بارهٔ تأثیر کلسیم بر فشار خون یک آزمایش مقایسه‌ای تصادفی‌شده بود. شکل ۳ این طرح را برحسب نکات عمدهٔ آن ارائه می‌کند. حُسن عمدهٔ تخصیص تصادفی آزمودنیها، آن است که استدلالهای (۱) و (۳) را به‌صورت سرراست و مکانیکی درمی‌آورد. بنابراین مسألهٔ استنتاج علت را به امتحان برازنده بودن یک مدل، و سپس در صورت مناسب بودن برازش، به انجام محاسباتی سرراست تقلیل می‌دهد. تخصیص تصادفی آزمودنیها اریبی را در تشکیل گروههای تیمار و دارونما از بین می‌برد و گروههایی را به‌وجود می‌آورد که تنها از طریق تغییرات شانسی، پیش از اعمال تیمارها، با هم تفاوت دارند. این طرح مقایسه‌ای به یاد ما می‌آورد که با همهٔ آزمودنیها دقیقاً رفتار یکسانی بجز از نظر محتوای قرصهایی که مصرف می‌کنند به عمل می‌آید. بنابراین اگر اختلافهایی در میانگین کاهش فشار خون، بزرگتر از آنچه می‌توان انتظار داشت که ناشی از شانس باشد، مشاهده کنیم، می‌توانیم اطمینان داشته باشیم که کلسیم بوده که تأثیری را که مشاهده می‌کنیم باعث شده است.

از ابزارهای عمدهٔ دیگر تولید داده‌ها، بررسی نمونه است که در آن، نمونه‌ای را برای تولید اطلاعاتی در بارهٔ جامعه‌ای بزرگتر انتخاب و بازبینی می‌کنند. مثالهای جالب در این زمینه فراوان‌اند: نظرسنجیهای سالم و ناسالم، گردایهٔ داده‌های اجتماعی و اقتصادی دولت، منابع داده‌های دانشگاهی همچون مرکز پژوهش آرای ملی[۱] در دانشگاه شیکاگو از این جمله‌اند. طرحهای

1. National Opinion Research Center

آماری برای نمونه‌گیری با اصرار بر اینکه شانس غیرشخصی باید نمونه را انتخاب کند، شروع می‌شود. ایدهٔ اصلی طرحهای آماری برای تولید داده‌ها، خواه از طریق نمونه‌گیری یا انجام آزمایش، استفادهٔ عمدی از شانس است. استفادهٔ صریح از سازوکارهای شانسی برخی از منابع اصلی اریبی را از بین می‌برد. این‌کار همچنین تضمین می‌کند که مدلهای احتمالی ساده، فرایند تولید داده‌های ما را توصیف می‌کنند، و بنابراین روشهای استنباط استاندارد قابل به‌کارگیری‌اند. با این حال، همان‌طور که مثال زیر نشان می‌دهد، برخلاف آزمایشهای تصادفی‌شده، مطالعات مشاهده‌ای به‌صورتی سرراست تن به استنباط علیّت نمی‌دهند. شرح این بررسی نخست توسط بِشت[1] و واکر[2] به‌عنوان یک مثال در [۱۲] آمده است؛ شرحی که ما از آن می‌آوریم به تأسی از [۲۶] است.

مثال ۳، سیگارکشیدن و تندرستی. در یکی از نخستین مطالعات مشاهده‌ای در بارهٔ سیگارکشیدن و تندرستی، نرخ مرگ‌ومیر برای سه گروه از مردان مقایسه شده است. این نرخها برحسب مرگ‌ومیر در هر سال برای هر ۱۰۰۰ مرد عبارت بودند از:

غیرسیگاریها ۲۰٫۲، سیگاریها ۲۰٫۵،
مصرف‌کنندگان سیگار برگ و پیپ ۳۵٫۵

برای آزمودن اینکه اختلافهای مشاهده شده ناشی از شانس هستند یا نه، می‌توانیم از مدلی مشابه با مدل مثال ۲ الف استفاده کنیم. اندازه‌های نمونه آن‌قدر بزرگ‌اند که می‌توانیم به‌آسانی تغییرات شانسی را به‌عنوان توضیحی برای اختلافهای مشاهده شده رد کنیم و به این نتیجه‌گیری ظاهری برسیم که سیگار کشیدن با مخاطرهٔ کمی همراه است، اما پیپ یا سیگار برگ یا هر دو کاملاً خطرناک‌اند. درواقع، این نتیجه‌گیری معتبر می‌بود هرگاه این داده‌ها از یک آزمایش تصادفی‌شدهٔ کنترل‌شدهٔ دو طرف کور، مانند مطالعهٔ کلسیم به‌دست آمده بودند. ولی این فرض در اینجا مطرح نیست. از آنجا که این بررسی مشاهده‌ای است، باید در مورد عاملهای دیگری که مرتبط با عادت سیگارکشیدن هستند و ممکن است علتِ این اختلافِ مشاهده شده باشند پرس‌وجو کنیم. در اینجا، سن یکی از چنین عوامل اصلی است: مصرف‌کنندگان سیگار برگ و پیپ معمولاً پیرتر از سیگاریها هستند و خطر مرگ با سن افزایش می‌یابد. در این مطالعه، متوسط سن برای سه گروه عبارت بوده‌اند از:

غیرسیگاریها ۵۴٫۹ سال، سیگاریها ۵۰٫۵ سال،
مصرف‌کنندگان سیگار برگ و پیپ ۶۵٫۹ سال

تنها بعد از تصحیح نرخهای مرگ با توجه به اختلاف سن است که اعدادی مطابق با آنچه انتظار داریم، به‌دست می‌آوریم:

غیرسیگاریها ۲۰٫۳، سیگاریها ۲۸٫۳،
مصرف‌کنندگان سیگار برگ و پیپ ۲۱٫۲

به نظر ما دو مثال اخیر تواماً متضمن دو درس مهم برای ریاضیدانانی است که آمار تدریس می‌کنند: نخست اینکه نتایج حاصل از یک مطالعه

1. Best 2. Walker

به‌طور قطع به چگونگی تولید داده‌ها وابسته است و دوم اینکه در مدلهای ریاضی استاندارد، تولید داده‌ها نادیده گرفته می‌شود.

ارائهٔ ایده‌های آماری برای تولید داده‌ها به‌منظور پاسخگویی به پرسشهای مشخص، از مؤثرترین خدمات آمار به معرفت بشری است. تولید داده‌ها با طرح بد، شایعترین نقیصهٔ جدّی در مطالعات آماری است. تولیدِ خوش‌طرح داده‌ها اجازه می‌دهد که روشهای استاندارد تحلیل را به‌کار برده و به نتایج روشنی برسیم. آماردانان حرفه‌ای به خاطر تخصصشان در طرح مطالعه دستمزد دریافت می‌کنند؛ اگر مطالعه خوب طرح‌ریزی شود (و هیچ حادثهٔ بد غیرمنتظره‌ای رخ ندهد)، برای تحلیل نیازی به افراد حرفه‌ای نیست. به عبارت دیگر، طرح تولید داده‌ها واقعاً مهم است. اگر صرفاً بگویید «فرض کنید $X_1$ تا $X_n$ مشاهده‌هایی مستقل و همتوزیع هستند»، شما آمار تدریس نمی‌کنید.

**۲.۲ تحلیل داده‌ها: کاوش و توصیف.** تحلیل داده‌ها شکل نوین «آمار توصیفی» است که با ابزارهای توصیفی استادانه‌تر و بسیار بیشتر و به‌ویژه با نظریه‌ای عمدتاً متعلق به جان توکی[1] از متخصصان آزمایشگاه بل و پرینستن تجهیز شده است. این نظریه اینک به نام *تحلیل کاوشی داده‌ها* یا EDA[2] مشهور است. هدف EDA، نظیر کاوشگری که وارد سرزمینهای ناشناخته می‌شود، آن است که ببیند داده‌های در دست چه می‌گویند. ما این مسأله را که آیا این داده‌ها نمایندهٔ جهانی بزرگترند موقتاً (اما نه برای همیشه) کنار می‌گذاریم.

جدول ۱ تمایزات بین EDA و استنباط استاندارد را به اختصار و در حد مقدماتی [۲۵] ارائه می‌کند:

جدول ۱ مقایسهٔ تحلیل کاوشی داده‌ها با استنباط رسمی مبتنی بر احتمال.

| تحلیل کاوشی داده‌ها | استنباط آماری |
|---|---|
| هدفْ کاوش نامحدود داده‌ها و جستجو برای یافتن الگوهای جالب است | هدفْ پاسخ‌دادن به پرسشهای مشخصی است که قبل از تولید داده‌ها عنوان شده‌اند. |
| نتایج تنها برای افراد و شرایطی که در مورد آنها داده‌هایی در دست داریم، صادق‌اند. | نتایج در مورد گروه بزرگتری از افراد یا رده‌ای گسترده‌تر از شرایط صادق‌اند. |
| نتایجْ غیررسمی و مبتنی بر آنچه در داده‌ها می‌بینیم، هستند. | نتایج به پشتوانهٔ حکمی که حاکی از اطمینان ما به آنهاست، رسمی‌اند. |

در عمل، تحلیل کاوشی پیش‌نیازی برای استنباط رسمی است. اغلب داده‌های واقعی شامل موارد غیرمنتظره‌ای هستند که برخی از آنها ممکن است استنباط برنامه‌ریزی‌شده‌ای را نامعتبر یا ما را مجبور به اصلاح آن گرداند. این یکی از دلایلی است که به‌کارگذاشتن داده‌ها از طریق یک شیوهٔ استنباط پیچیده (و بنابراین، مکانیکی‌شده) قبل از کاوش دقیق در آنها نشانه‌ای از بی‌تجربگی در آمار است. رفت و برگشت بین داده‌ها و مدلها با ابزارهای تشخیصی پیشرفته ادامه می‌یابد به‌طوری این ابزارها امکان نقد مدلهایی خاص را با استفاده از داده‌ها می‌دهند. این ابزارها روح EDA را با نتایجی از تحلیل ریاضیِ پیامدهای مدلها درهم می‌آمیزند.

1. John Tukey 2. Exploratory Data Analysis

| دارونما | | کلسیم |
|---|---|---|
| 1 | −1 | |
| 5 | −0 | 5 |
| 33111 | −0 | 234 |
| 3 | 0 | 1 |
| 5 | 0 | 7 |
| 2 | 1 | 01 |
| | 1 | 78 |

**شکل۴** نمودار ساقه‌ای موازی کاهش فشار خون سیستولیک برای دوگروه از مردان

همان‌گونه که قبلاً دیده‌ایم، مدل مثال ۲ الف از آن‌رو که بین مشاهده و آزمایش تمایزی قائل نمی‌شود، ناقص است. این مدل همچنین نظیر اغلب مدلهای ریاضی آرمانی برای پدیده‌های واقعی، غیرواقعی است. جمله‌ای به جورج باکس[1] آماردان منسوب است که می‌گوید «همهٔ مدلها نادرست‌اند، ولی بعضی از آنها مفیدند». استفاده‌کننده از روشهای استنباطِ مبتنی بر این مدل باید به‌دقت در بارهٔ کفایت این مدل برای وضعیت مورد بحث و داده‌ها کاوش کند. آیا نقیصه‌هایی در تولید داده‌ها (خواه نمونه‌ای باشد خواه آزمایشی) وجود داشته‌اند که استنباط را بی‌معنا کنند؟ آیا داده‌ها که یقیناً مشاهدات مستقلی روی یک توزیع کاملاً نرمال نیستند، به‌اندازهٔ کافی نرمال‌اند که اجازهٔ استفاده از روشهای استاندارد را بدهند؟ این سؤال با بررسی کاوشی خودِ داده‌ها همراه با اطلاع از اینکه تحلیل برنامه‌ریزی شده چقدر در مقابل انحراف از فرضهای مدل «استوار» است، پاسخ داده می‌شود.

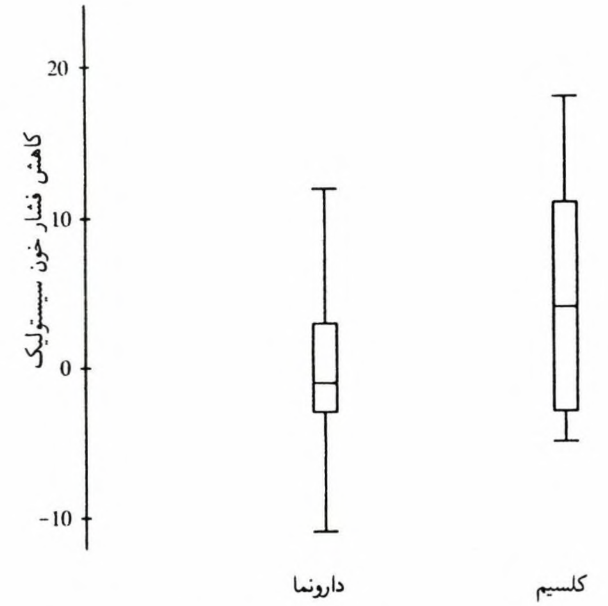
مثال ۲ د، کاوش مقدماتی داده‌های کلسیم. تحلیل ممکن است با این طرح کلی ساده آغاز شود: نمودار، شکل، مرکز و پراکندگی.

نمودار. یک نمودار ساقه‌ای هر مقدارِ داده‌ها را به یک ساقه و یک برگ تفکیک کرده، سپس برگها را روی ساقه‌های مشترک جور می‌کند. شکل ۴ یک نمودار ساقه‌ای پشت به‌پشت مفید برای مقایسهٔ دوگروه را نشان می‌دهد.

شکل. توزیع گروه دارونما تک‌مدی و متقارن است. با این حال گروه تیمار تصور ضعیفی از دومدی بودن را به‌وجود می‌آورد که امکان وجود دو نوع آزمودنی را مطرح می‌کند. آیا ممکن است کسانی باشند که به کلسیم واکنش نشان دهند وکسانی‌که به آن واکنش نشان ندهند؟ براساس این داده‌ها به‌هیچ عنوان چیزی نمی‌توان گفت، ولی این امکان ارزش توجه را دارد.

مرکز و پراکندگی. نموداری مفید برای مقایسهٔ مرکزها، پراکندگیها و تقارنها، نمودار جعبه‌ای است (شکل ۵). هر جعبه مکان چارکها و میانهٔ توزیع را مشخص می‌کند؛ «جاروبکها» از چارکها تا کرانگین‌ترین نقاط در محدودهٔ ۱٫۵ برابر دامنهٔ بین‌چارکی از نزدیکترین چارک، امتداد می‌یابند، و نقاط دورتر از میانه به‌طور مجزا نشان داده شده‌اند. ما در اینجا اختلافی بین میانه‌ها ملاحظه می‌کنیم، اما متوجه اختلاف بارزتری در پراکندگیها نیز می‌شویم، اختلافی که در فرض برابری واریانسها که برای توجیه برآورد ادغام‌شده در مثال ۲ الف به‌کار رفت ایجاد شبهه می‌کند.

1. George Box

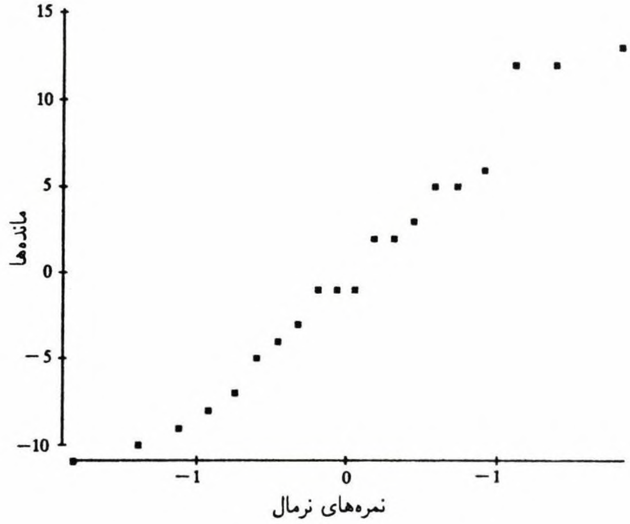



**شکل۵** نمودار جعبه‌ای موازی برای کاهش فشار خون سیستولیک در دوگروه از مردان

نمودار چندکی نرمال. پیش از استفاده از آزمون $t$ برای مقایسهٔ میانگینها، به صواب نزدیکتر است که این پرسش را مطرح کنیم. آیا داده‌ها دلیلی در اختیار ما می‌گذارند که مدل نرمال مثال ۲ الف را زیر سؤال ببریم؟ در اینجا از هر مشاهدهٔ میانگین گروه را کم می‌کنیم تا مانده‌ها را به‌دست آوریم، سپس مانده‌های مرتب شده را در مقابل چندکهای متناظر یک توزیع نرمال رسم می‌کنیم؛ (نگاه کنید به شکل ۶). عرضهای نقاط، ۲۱ ماندهٔ مرتب‌اند که خط حقیقی را به ۲۲ زیربازه تقسیم می‌کنند. طولهای نقاط ۲۱ مقداری هستند که خط حقیقی را به ۲۲ پاره‌خط که تحت مدل نرمال متساوی‌الاحتمال هستند تقسیم می‌کنند. اگر داده‌ها از تنها یک توزیع نرمال آمده باشند، می‌توانیم انتظار داشته باشیم که نقاط نزدیک یک خط قرار گیرند.

برای داده‌های کلسیم الگو تقریباً خطی است، هر چند جهش عمودی قبل از سه نقطهٔ منتهی‌الیه سمت راست، مانده‌های مشاهده‌شده‌ای را نشان می‌دهد که از مقادیر پیشگویی شده توسط مدل نرمال بزرگتر هستند، الگویی که با پراکندگیهای نابرابر در نمودار جعبه‌ای سازگار است.

در آموزشی که ساختار ریاضی دارد و معطوف به چگونگی استنتاج روشها از مدلهاست، در بسیاری موارد فقط کلیترین هشدارها در بارهٔ واقعیتهای عملی مطرح می‌شود. آمار در عمل به گفتگویی بین مدلها و داده‌ها شباهت دارد. مدلها برای فرایندی که داده‌ها را تولید می‌کند به راستی نقش اصلی را در استنباط آماری ایفا می‌کنند. بنابراین کاوش ریاضی خواص و پیامدهای مدلها مهم است (همان‌طور که در اقتصاد و فیزیک مهم است). اما داده‌ها همچنین می‌توانند مدلهای پیشنهادی را نقد یا حتی باطل بودن آنها را آشکار نمایند. در مثالهای کلسیم، تحلیل کاوشی به ما هشدار می‌دهد که خیلی زیاد به فرض برابری واریانسها متکی نباشیم، و از آزمون $t$ی اصلاح‌شده‌ای استفاده کنیم که واریانسهای متمایزی برای دو گروه برآورد می‌کند. می‌توان اظهارنظر باکس را به این صورت تعبیر کرد که آمار صرفاً ریاضیات نیست: فضاهای ریاضی درست‌اند؛ روشهای آماری اگر با مهارت به‌کار روند، گاهی مؤثرند.



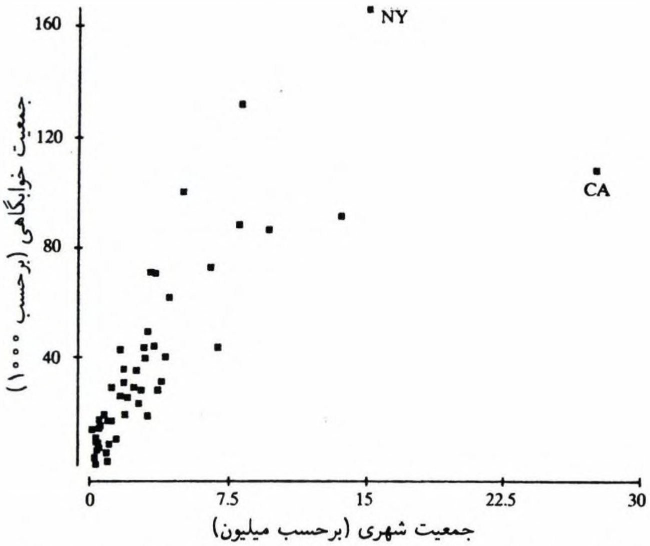

**شکل۶** نمودار چندکی نرمال برای دادههای فشار خون

امکانات گستردهٔ محاسبات ارزان، بهویژه کارهای گرافیکی، با تمایل به اینکه «ببینیم دادهها چه میگویند» درآمیخته و به تولید انبوهی از ابزارهای جدید منتهی شده است: پیشاپیش همه، نمودارهای ساقهای و نمودارهای جعبهای مثال ۲ ب را داریم، اما همچنین هموارکنندههای نمودار پراکندگی آزاد-مدل، الگوریتمهای رگرسیونی مقاوم، ایدههای هوشمندانهٔ نمایش دادههای با بعد بالا برروی پردههای دوبعدی، و بسیاری ابزارهای پیشرفتهتر تشخیصی برای وضعیتهای خاص در دست است. نرمافزارهای آماری متعارف بسیاری از اینها را انجام میدهند. کتابهای [۷] و [۹] به قلم دانشمندانی از آزمایشگاههای بل که از توکی تأثیر پذیرفتهاند، بسیاری از مطالب گرافیکی پایه را ارائه میکنند. بستههای نرمافزاری S و S_PLUS که در آزمایشگاههای بل ابداع شدهاند، بیشترکارهای گرافیکی جدید را انجام میدهند، و نیز چندین رده از مدلهای جدید را به اجرا درمیآورند. برای بحث مفصل موضوع اخیر، ر. ک [۸].

هر چند ممکن است این وسوسه در مبتدیان بهوجود آید که تحلیل دادهها را صرفاً مجموعهای از ابزارها و روشهای زیرکانه تلقی کنند، ولی ارزش این ابزارها در استفاده از آنها بهروشی نظاممند و مطابق با استراتژیهایی است که برای سازماندهیِ بازبینی دادهها تنظیم شدهاند.

۱. از ساده به پیچیده پیش بروید: ابتدا هر متغیر را جداگانه بررسی کنید، سپس رابطهٔ آنها با یکدیگر را ملاحظه کنید.
۲. از سلسلهمراتبی از ابزارها استفاده کنید: ابتدا نمودار دادهها را رسم کنید، سپس توصیفهای عددی مناسبی از جنبههای مشخصی از دادهها را انتخاب کنید، بعد از آن در صورت اقتضا مدل ریاضی جمعوجوری برای الگوی کلی دادهها برگزینید.
۳. هم به الگوی کلی و هم به هرگونه انحراف چشمگیر از آن الگو توجه کنید.

بخشی از نظریهٔ یکپارچهساز (اما غیرریاضی) EDA این است که این اصول در وضعیتهای مختلف بهکار روند. اگر دادههایی از فقط یک متغیرکمّی در دست داشته باشیم، میتوانیم توزیع را با یک نمودار ساقهای نمایش دهیم، توجه کنیم که اگر تقریباً متقارن است، میانگین و انحراف معیار را بهعنوان خلاصههای عددی محاسبه کنیم، و از یک نمودار چندکی نرمال استفاده کنیم تا ببینیم که آیا توزیع نرمالْ مدل جمعوجور مناسبی برای الگوی کلی است یا خیر. اما اگر دو متغیرکمّی در دست داشته باشیم، نمودار پراکنش را رسم میکنیم، سو و قوت پیوند خطی را بهکمک همبستگی اندازه میگیریم، و در صورت اقتضا از یک خط راست برازش یافته بهعنوان مدل برای الگوی کلی استفاده میکنیم. بنابراین دستورالعمل یکمتغیرهٔ «نمودار، شکل، مرکز، پراکندگی» در چارچوب دادههای دومتغیره بهصورت «نمودار، شکل، سو، قوت» درمیآید.

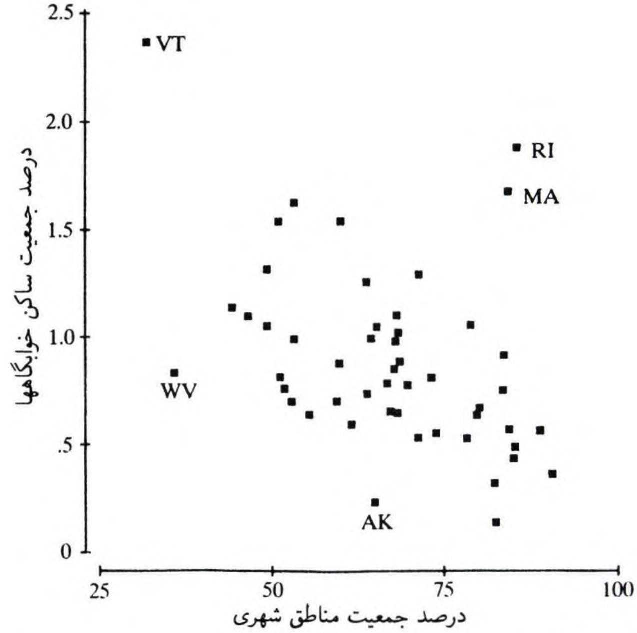



**شکل۷** نمودار پراکنش نقطهای جمعیت خوابگاهی در مقابل جمعیت شهری برای ۵۰ ایالت در آمریکا

در اینجا، مانند هر جای دیگر، تحلیل صرفاً جستجویی برای یافتن الگوها نیست، بلکه جستجویی برای یافتن الگوهای بامعنی است. همانطورکه مثال زیر نمایش میدهد، بهترین برازش لزوما مفیدترین نیست.

مثال ۳، خوابگاهها و شهرها. هر نقطه در شکل ۷ یکی از ۵۰ ایالت کشور امریکا را نشان میدهد که در آن مختص افقی برابر است با جمعیت شهری هر ایالت، و مختص عمودی برابر است با تعداد دانشجویان مقیم خوابگاه در آن ایالت. شکل ۷ چند جنبهٔ بارز دارد. برای مثال، نمودارْ پرهایشکل است و نقاط بسیاری در سمت چپ پایین یکجا جمع شدهاند: اغلب ایالتها جمعیت شهری نسبتاً کوچکی (حدود دو میلیون) و نیز جمعیت خوابگاهی نسبتاً کوچکی (زیر ۵۰۰۰۰) دارند، تنها تعداد کمی از ایالتها جمعیت شهری یا جمعیت خوابگاهی بسیار بزرگی دارند، و تغییر از ایالتی به ایالت دیگر (فضای بیشتر بین نقاط) برای ایالتهای با مقادیر بزرگتر، بیشتر است. الگوی پیوند بین دو متغیر مثبت و قوی است: جمعیتهای شهری کوچکتر با جمعیتهای خوابگاهی کمتر، و جمعیتهای شهری بزرگتر با جمعیتهای خوابگاهی بیشتر همراه است، و برای همهٔ ایالتها بجز تعداد کمی از آنها، دانستن اندازهٔ جمعیت شهری به ما اجازهٔ پیشگویی جمعیت خوابگاهی آنرا در حدود دامنهٔ نسبتاً کوچکی میدهد.

علیرغم برازش خوب بین تصویر و ماجرا، در این تحلیل یک مشخصهٔ بسیار مهم نادیده گرفته شده است. اگر آنچه را که این الگو در ظاهر نشان میدهد بپذیریم که ایالتهای با جمعیت شهری بزرگ، جمعیتهای خوابگاهی




**شکل۸** نمودار پراکنش داده‌های خوابگاهها و شهرها پس از تصحیح نسبت به جمعیت «متغیر پنهان»

بزرگی نیز دارند، ممکن است به این نتیجه‌گیری وسوسه شویم که شهرها دانشگاهها را بیشتر جذب کرده‌اند. هر چند نمونه‌های فراوانی دال بر تأیید این امر به ذهن می‌آید، این تفسیر ساده‌لوحانه اشتباه است: هر دو متغیر ما اندازه‌های غیرمستقیمی از اندازهٔ جمعیت ایالتها هستند، بنابراین چندان شگفت‌آور نیست که این دو مقدار، پیوند قوی مثبتی را نشان می‌دهند. برای آشکارکردن رابطهٔ بامعناتری، مجبوریم که «نسبت به متغیر پنهان تصحیح انجام دهیم»: جمعیت شهری را به جمعیت کل تقسیم کنید تا درصد شهری را به‌دست آورید، جمعیت خوابگاهی را بر جمعیت کل تقسیم کنید تا درصدی را که در خوابگاهها زندگی می‌کنند، به‌دست آورید، و نمودار نتایج را رسم کنید (شکل ۸).

حال این رابطه ضعیفتر، اما آنچه بیان می‌کند جالبتر است. سوی پیوند برعکس شده است: ایالتهای روستایی ــ آنهاکه درصدکمتری از ساکنانشان در نواحی شهری زندگی می‌کنند ــ درصد بیشتری از ساکنانشان در خوابگاههای دانشگاهی زندگی می‌کنند. این با قدری تأمل معقول به‌نظر می‌آید. به پولمن[1] در ایالت واشنگتن، یا ایمس[2] در ایالت آیوا، نورمن[3] در ایالت اکلاهما، لاورنس[4] در ایالت کانزاس بیندیشید.[5] ایالتهای روستایی ممکن است از نظر تعداد مطلق، کالجها و دانشگاههای کمتری داشته باشند، اما دانشجویان این ایالتها درصد بیشتری از کل جمعیت ایالت را تشکیل می‌دهند، و نیز احتمال بیشتری دارد که در خوابگاهها زندگی کنند.

**۳.۲ استنباط رسمی: استدلال در مقابل شانس.** استنباط آماریْ روشهایی آماری برای استنتاج از داده‌ها در بارهٔ جامعه یا فرایندی که داده‌ها از آن استخراج شده‌اند، به‌دست می‌دهد. در اینجا قائل شدن تمایز بین آماره‌های نمونه‌ای و پارامترهای جامعه ضرورت پیدا می‌کند (در تحلیل داده‌ها چنین نبود. مقادیر واقعی پارامترها بر ما نامعلوم‌اند. ما آماره‌ها را در دست داریم، اما اگر تولید داده‌ها را تکرار کنیم، آنها مقادیر مختلفی به خود می‌گیرند. در استنباط باید این تغییرپذیری در نمونه به‌حساب آورده شود.

احتمالْ نوعی از تغییرپذیری، یعنی تغییرپذیری شانسی در پدیده‌های تصادفی، را توصیف می‌کند. وقتی یک سازوکار تصادفی به‌طور صریح برای تولید داده‌ها به‌کار می‌رود، احتمالْ تغییراتی را توصیف می‌کند که انتظار داریم در نمونه‌های تکراری از یک جامعه یا آزمایشهای تکراری در شرایطی یکسان، مشاهده کنیم. یعنی به این پرسش پاسخ می‌دهد که «چه اتفاقی می‌افتد اگر این کار را چندین بار انجام دهیم؟»

استنباط آماری استاندارد مبتنی بر احتمال است. این استنباطْ نتایجی را از داده‌ها همراه با اشاره‌ای به میزان اطمینان ما به این نتایج، ارائه می‌دهد. حکم مربوط به اطمینان براساس این سؤال است که «چه اتفاقی می‌افتد اگر این روش استنباط را چندین بار به‌کار بریم؟» این دقیقاً از نوع سؤالاتی است که احتمالْ می‌تواند پاسخ دهد، و به همین دلیل است که آن را مطرح می‌کنیم. این اشاره به میزان اطمینان ما به روشهای خود، که به زبان احتمال بیان می‌شود، چیزی است که استنباط رسمی را از نتایج غیررسمی مبتنی بر مثلاً، تحلیل کاوشی داده‌ها، متمایز می‌سازد.

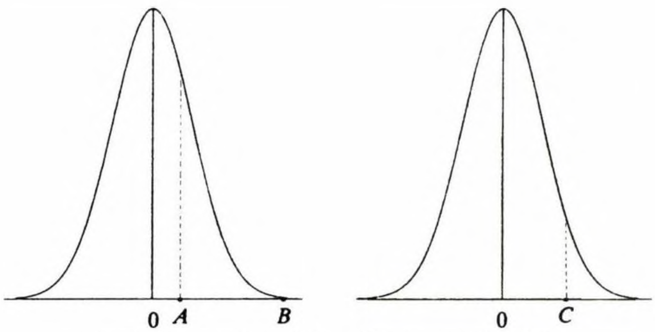
هر شیوهٔ استنباطیِ خاص با یک آماره یا شاید چندین آماره، که از داده‌های نمونه‌ای محاسبه می‌شوند، شروع می‌شود. توزیع نمونه‌گیری، توزیع احتمالی است که توصیف می‌کند این آماره چگونه تغییر خواهد کرد هرگاه نمونه‌های بسیاری از یک جامعه استخراج شوند. در آمار مقدماتی دو نوع شیوهٔ استنباطی، بازه‌های اطمینان و آزمونهای معنی‌داری را ارائه می‌کنیم. بازهٔ اطمینانْ پارامتر نامعلومی را برآورد می‌کند. هدف از آزمون معنی‌داریْ ارزیابی شواهد حضور عامل مورد بررسی در جامعه است.

هر بازهٔ اطمینان مرکب از دستورالعملی است برای برآوردکردن پارامتری نامعلوم با استفاده از داده‌های نمونه‌ای، معمولاً به شکل «برآورد $\pm$ حاشیهٔ خطا»، و یک سطح اطمینان که احتمالِ آن است که دستورالعمل واقعاً بازه‌ای تولید کند که مقدار واقعی پارامتر را دربرداشته باشد. یعنی، سطح اطمینان به این سؤال پاسخ می‌دهد که «اگر از این روش چندین بار استفاده کنم، چند بار جواب درستی به من می‌دهد؟»

آزمون معنی‌داری با این فرض آغاز می‌شود که عامل مؤثری که در جستجوی آن هستیم در جامعه حضور ندارد. طی آن سؤال می‌شود که «در چنین صورتی، آیا نتیجهٔ نمونه‌ای تعجب‌آور است یا خیر؟» یک احتمال ($p$-مقدار) می‌گوید که نتیجهٔ نمونه‌ای چقدر تعجب‌آور است. نتیجه‌ای که در صورت عدم حضور عاملی که در جستجوی آن هستیم به‌ندرت به‌دست می‌آید گواه خوبی است بر اینکه آن عامل درواقع حضور دارد. شکل ۹ این استدلال را در مثال پزشکی ما نشان می‌دهد. خم نرمال در آن شکل، معرفِ توزیع نمونه‌ای اختلاف $\bar{y}-\bar{x}$ بین کاهشها در میانگین فشار خون درگروههای کلسیم و دارونما برای حالتی است که اختلافی بین دو میانگین جامعه وجود ندارد. این توزیع، که تغییرپذیریِ صرفاً شانسی را نشان می‌دهد دارای میانگین ۰ است. برآمدهای بزرگتر از ۰ از آزمایشهایی می‌آیند که در آنها کلسیم فشار خون را بیشتر از دارونما کاهش می‌دهد. اگر نتیجهٔ A را مشاهده

1. Pullman  2. Ames  3. Norman  4. Lawrence

۵. شهرهای کوچک دانشگاهی در آمریکا.ـم.




**شکل۹** ایدهٔ معنی‌داری آماری: آیا این مشاهده تعجب‌آور است؟

کنیم، تعجب نمی‌کنیم زیرا پیشامدی که با صفر این‌قدر فاصله دارد، اغلب شانسی رخ می‌دهد. این نتیجه هیچ‌گونه گواه معتبری بر این ادعا که کلسیم بر دارونما برتری دارد در اختیار نمی‌گذارد. اگر نتیجهٔ B را مشاهده کنیم، آزمایش نتیجه‌ای چنان قاطع به‌بار آورده است که تقریباً هرگز به‌طور شانسی نمی‌تواند رخ دهد. در این صورت شاهدی قوی داریم که میانگین کلسیم از میانگین دارونما بیشتر است. $p$-مقدار (احتمال دُم راست) برای نقطهٔ A برابر است با ۰٫۲۴ و برای نقطهٔ B برابر است ۰٫۰۰۰۵. این احتمالها میزان تعجب‌آوربودن مشاهده‌ای به این بزرگی را وقتی‌که عامل مؤثری در جامعه وجود ندارد به‌صورت کمّی درمی‌آورند. در مورد داده‌های واقعی چطور؟ نقطهٔ C مقدار مشاهدهٔ شده $\overline{x} - \overline{y} =$ ۵٫۲۷۳ را نشان می‌دهد. $p$-مقدار متناظر ۰٫۰۵۵ است. کلسیم در ۵٫۵٪ از تعداد زیادی آزمایش، حداقل به این مقدار بر دارونما صرفاً به‌علت تغییرات شانسی برتری دارد. آزمایش شواهدی به‌دست می‌دهد که کلسیم مؤثر است اما این شواهد چندان قوی نیستند. نکته‌ای برای آنهایی که نگران جزئیات‌اند: در این محاسبات $p$-مقدار، تغییرات میانگین نمونه‌ای معلوم فرض شده است. در عمل، ما باید انحراف معیارها را از روی داده‌ها برآورد کنیم. آزمون حاصل $p$-مقدار بزرگتری دارد: $p =$ ۰٫۰۷۲.

**۳. تدریس:** در بحث تدریس، می‌توانیم توجه خود را بر محتوا متمرکز کنیم، یعنی چیزی که می‌خواهیم دانشجویان فرا گیرند، یا بر فنِ آموزش، یعنی کاری که می‌کنیم تا به آنها در فراگرفتن کمک کنیم. البته این دو مبحث مرتبط هستند. به‌ویژه تغییرات در فن آموزش اغلب با تغییر بعضی از اولویتها در نوع چیزهایی که می‌خواهیم دانشجویانمان یاد بگیرند به پیش برده می‌شود. با این حال مناسبتر است که محتوا و فن آموزش را جداگانه مورد بحث قرار دهیم. این بخش هماهنگ با قسمتهای دیگر مقاله، محتوا را مورد توجه قرار می‌دهد، و به‌ویژه، شامل وجهی از گفتگو بین آماردانان و ریاضیدانانی است که به تدریس آمار هم می‌پردازند.

**۱.۳ آمار باید به عنوان آمار تدریس شود.** آماردانان متقاعد شده‌اند که آمار با وجود آنکه وابسته به ریاضیات است، از شاخه‌های ریاضیات نیست. مانند اقتصاد و فیزیک، آمار استفاده‌ای گسترده و اساسی از ریاضیات می‌کند، اما قلمرو خاص خود را برای کاوش و مفاهیم اساسی خود را برای هدایت این کاوش دارد. با توجه به این اعتقادات، به‌طور طبیعی ترجیح می‌دهیم که آمار مقدماتی به عنوان آمار تدریس شود. انجمن آمریکایی آمار[1] (ASA) و جامعهٔ ریاضی آمریکا[2] (MAA) کمیته‌ای مشترک برای بحث در مورد برنامهٔ درسی آمار مقدماتی تشکیل داده‌اند. توصیه‌های این گروه نمودی از این نظر است که آموزش آمار باید بر ایده‌های آماری متمرکز شود. گزیده‌ای از این توصیه‌ها به نقل از [۱۰] در زیر می‌آید؛ بحثی مفصلتر در [۱۱] آمده است.

> تقریباً هر درسی در آمار با تأکید بیشتر بر داده‌ها و مفاهیم و توجه کمتر به نظریه و آوردن دستورالعملهای کمتر، بهتر می‌شود. تا سرحد امکان، به محاسبات و کارهای گرافیکی باید حالت مکانیکی و سرراست داده شود. هدف اصلی هر درس مقدماتی باید کمک به دانشجویان در یادگیری اصول تفکر آماری باشد [ که مشتمل‌اند بر ] نیاز به داده‌ها، اهمیت تولید داده‌ها، همه جا حاضربودن تغییرپذیری، کمّی‌سازی و توضیح تغییرپذیری.

توصیه‌های کمیتهٔ ASA/MAA منعکس‌کنندهٔ تغییراتی در حوزهٔ آمار در طی دو سه دههٔ پیشین است. آمار دانشگاهی، برخلاف ریاضیات به حیطه‌ای گسترده‌تر که کاروزری حرفه‌ای غیردانشگاهی باشد، متصل است. فناوری محاسباتی، کاروزری آماری را کاملاً تغییر داده است. پژوهشگران دانشگاهی، تا حدی بر اثر تقاضاهای کاروزری و تا حدی به علت تواناییهای فناوری نوین، سلیقه‌های خود را در پژوهش تغییر داده‌اند. روشهای خودگردان[3]، هموارسازی ناپارامتری داده‌ها، تشخیصیات[4] رگرسیونی، و رده‌هایی عامتر از مدلها که مستلزم برازشهای تکراری‌اند، ازجمله ثمرات جدید توجه مجدد به تحلیل داده‌ها و استنباط علمی‌اند. افرون و تیبشیرانی [۱۴] بخشی از این کار را برای غیرمتخصصان توصیف کرده‌اند.

**۲.۳ نه ریاضیات نه جادوگری.** تأکید زیاد بر استنباطِ مبتنی بر احتمال، یکی از نشانه‌های افراط در ریاضی در درس آمار مقدماتی است، و در عین حال، اینکه مدرسانِ برخوردار از تربیت و تفکر ریاضی دوست ندارند از ارائهٔ آمار براساس نظریه دست بکشند، مبنای قابل احترامی دارد: برای اجتناب از معرفی آمار به‌عنوان جادوگری است. مسلماً تدریس آمار مقدماتی به‌صورت جادوگری کار رایجی است. استفاده‌کنندهٔ آمار از جهات بسیاری شبیه شاگرد جادوگر است. سحر و جادو تأثیری خودبه‌خودی دارد که پایان‌نامه‌ها را قابل پذیرش و مطالعات را قابل چاپ می‌گرداند. منظور این نیست که ما بفهمیم سحر چگونه کار می‌کند ـــ این در حیطهٔ کار خود جادوگر است. سحر باید به‌طور دقیق دستورالعمل را دنبال کند تا فاجعه‌ای پیش نیاید ـــ کاوش و انعطاف‌پذیری هم، مانند فهمیدن، برای شاگردان ممنوع است. خوشبختانه جادوگران نرم‌افزاری تدارک دیده‌اند که تقلید دقیق از سحرهای پذیرفته شده را به‌صورت ماشینی و سرراست درآورده است.

خطر[تلقی] آمار به‌عنوان جادو، واقعی است. اما دفاع مناسبْ عقب‌نشینی به‌سوی عرضهٔ ریاضی مطلب نیست که برای تشریح موضوع، ناکافی است و برای دانشجویان اغلب غیرقابل درک است. درکِ ریاضی تنها نوع درک نیست. حتی در اغلب موضوعات علمی که از ریاضیات استفاده می‌کنند و در

1. American Statistical Association
2. Mathematical Association of America
3. Bootstrap
4. diagnostics

آنها درک پدیدهٔ هدف و مفاهیم اصلی موضوع اولویت دارند، درک ریاضی، سودمندترین نوع آن هم نیست. باید تلاش کنیم تا چارچوبی عقلانی عرضه کنیم که مجموعهٔ ابزارهایی را که آماردانان به‌کار می‌برند قابل معنا گرداند و مشوق کاربرد انعطاف‌پذیر آنها در حل مسائل باشد. دانشجویان، ریاضیات را وقتی درک می‌کنند که قدرت تجرید، استنتاج و بیان نمادین را ارج گذارند، و بتوانند به‌صورتی انعطاف‌پذیر از ابزارها و استراتژیهای ریاضی در مواجهه با مسائل گوناگون استفاده کنند. استدلال بر مبنای داده‌های تجربی ناحتمی، روشی عقلانی است که همان‌قدر قدرتمند و نافذ است. چگونه می‌توانیم به بهترین وجه دانشجویان خود را راهنمایی کنیم که این روش عقلانی را درک کنند، ارج گذارند، و شروع به جذب آن کنند؟

**۳.۳ با تحلیل کاوشی داده‌ها شروع کنید.** هر چند ترتیبی که با توجه به شکل ۲ به ذهن می‌آید، شروع از تولید داده‌هاست، تجربه عکس آن را می‌گوید. تحلیل کاوشی داده‌ها آغاز خوبی است زیرا ملموستر است؛ در آن نیازی به تمایز قائل شدن بین جامعه و نمونه نیست و نیازی نیست که وجوه تصادفی‌سازی که در مقابل اریبی حفاظ ایجاد می‌کنند، مورد بحث قرار گیرند. روشهای پایه‌ای از نظر مفهومی و الگوریتمی ساده‌اند، و داده‌ها در دست‌اند ــ اعداد واقعی بر روی صفحهٔ کاغذ، نه اشباح داده‌های آینده، که در طرح و آزمایش با آنها سروکار داریم. به‌علاوه، ایجاد انگیزه مطرح نیست. دانشجویان تحلیل کاوشی را دوست دارند و درمی‌یابند که می‌توانند آن‌را انجام دهند که این خود مزیتیِ اساسی در تدریس موضوعی است که اکثر دانشجویان از آن وحشت دارند. درگیرکردن دانشجویان در تعبیر نتایج از همان اوان کار، قبل از آنکه ایده‌های دشوارتر بر سر راهشان ظاهر شده توجه آنها را به خود جلب کنند، می‌تواند به ایجاد عادات خوبی در آنها کمک کند که وقتی به مبحث استنباط می‌رسید بهرهٔ خود را خواهد داد. سرانجام، شروع با تحلیل داده‌ها زمینه را بر این طرح و استنباط آماده می‌کند. تجربهٔ کار با توزیعهای داده‌ها، دانشجویان را با همه جا حاضربودن تغییرپذیری و وجود بالقوهٔ اریبی آشنا می‌کند که این دو جنبه، دلایل اصلی نیاز ما به طرح دقیق است. اگر طرح را قبل از تحلیل داده‌ها تدریس کنید، برای دانشجو مشکلتر خواهد بود بفهمد که چرا طرح اهمیت دارد. تجربهٔ کار با توزیعهای داده‌ها همچنین بهترین راهِ آماده شدن برای رویارویی با مفهوم دشوار توزیع نمونه‌گیری است.

کوشش کرده‌ایم این موضوع را مطرح کنیم که مجموعه‌ای منسجم (هر چند غیرریاضی) از ایده‌ها و ابزارهای وابسته برای کاوش در داده‌ها وجود دارد. دانشجویان لازم است با نوشتن شرحهای منسجمی در بارهٔ داده‌ها، استفاده از این ایده‌ها و ابزارها را تمرین کنند. برای کمک به آنها هم خلاصه‌ای از آنچه را که باید بنویسند و هم مثالهایی که می‌توانند به‌عنوان مدل به‌کار روند، فراهم آورده‌ایم. برای مثال، شکل ۱۰، طرحی برای توصیف تنها یک متغیر کمّی است.

دنبال کردن این طرح هم نیاز به اطلاع در بارهٔ ابزارهایی دارد که در آن ذکرشده است و هم نیاز به قدرت تشخیص برای انتخاب از بین شقوق و تفسیر نتایج. قدرت تشخیص از تجربهٔ کار با داده‌ها به‌دست می‌آید. دانشجویان در ابتدا نمی‌توانند آن‌طور که کلمات و معادلات را می‌فهمند، از نمودارها

الف. داده‌ها را توصیف کنید
- تعداد مشاهدات
- ماهیت متغیر
- چگونه اندازه‌گیری شده است
- واحدهای اندازه‌گیری

ب. نمودار داده‌ها را رسم کنید، با انتخاب از میان
- نمودار نقطه‌ای
- نمودار ساقه‌ای
- بافت نگار

پ. الگوی کلی را توصیف کنید
- شکل
  - شکل مشخصی وجود ندارد؟
  - چاوله یا متقارن؟
  - قله‌های یگانه یا چندگانه؟
- مرکز و پراکندگی؛ با انتخاب از میان
  - خلاصهٔ پنج عددی
  - میانگین و انحراف معیار
- آیا نرمال بودن یک مدل مناسب است (نمودار چندکی نرمال)؟

ت. انحرافات آشکار از الگوی کلی را وارسی کنید
- دورافتاده‌ها
- رخنه‌ها یا خوشه‌ها

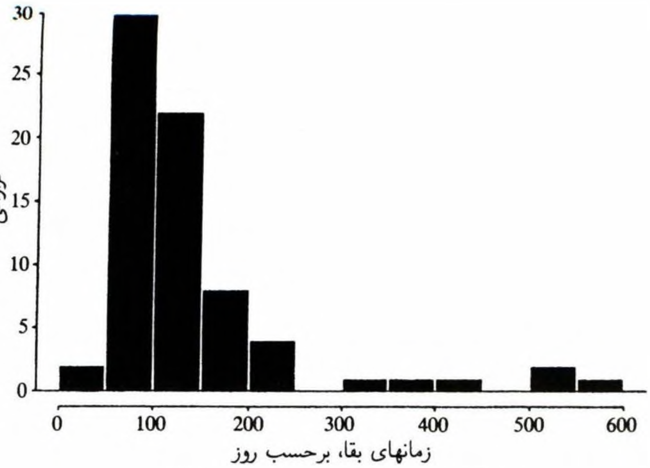
خ. به زبان زمینهٔ مسأله، یافته‌هایتان در پ و ت را تعبیر کنید. توصیفهای موجهی برای یافته‌های خود پیشنهاد کنید.

**شکل ۱۰** طرح توصیف داده‌ها برای یک متغیر کمی

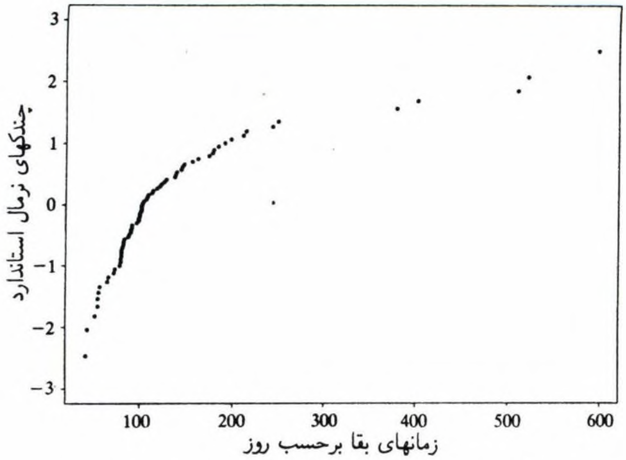
سردرآورند. در زیر مثالی از یک تحلیل بنیادی از یک متغیر ارائه می‌شود. توصیف روابط بین چندین متغیر مستلزم ابزارهای استادانه‌تر و قدرت تشخیص بیشتری است.

> در مطالعه‌ای در بارهٔ مقاومت در مقابل ابتلا به بیماری [۲]، پژوهشگرانْ باسیل سل را به ۷۲ خوکچهٔ هندی تزریق و مدت زمان زنده ماندن آنها را بعد از ابتلا به بیماری برحسب روز اندازه‌گیری کرده‌اند. هم بافت‌نگار (شکل ۱۱) و هم نمودار چندکی نرمال (شکل ۱۲) نشان می‌دهند که توزیع زمانهای بقا شدیداً چاوله به راست است. دادهٔ دورافتاده‌ای در بین نیست گرچه بعضی به مراتب بیشتر از میانگین زنده مانده‌اند، که به نظر یکی از مشخصه‌های توزیع کلی می‌آید و نه حاکی از، مثلاً، خطا در اندازه‌گیری یا ثبت این موردها.

چاولگی شدید حاکی از آن است که خلاصهٔ پنج عددی (مینیمم = ۴۳ روز،




**شکل۱۱** بافت‌نگار زمانهای بقای خوکچهٔ هندی




**شکل۱۲** نمودار چندکی نرمال برای زمانهای بقای خوکچهٔ هندی

چارک اول = ۵ر۸۲ روز، میانه = ۵ر۱۰۲ روز، چارک سوم = ۵ر۱۵۱ روز، و ماکسیمم = ۵۹۸ روز) خلاصهٔ عددی بهتری نسبت به میانگین و انحراف معیار ($\overline{x}$ = ۸ر۱۴۱ و $s$ = ۲ر۱۰۹ روز) است. تغییرات بسیار بزرگی در زمانهای بقا بین موردها وجود دارد ــ برای مثال، چارک سوم تقریباً ۱۵۰٪ میانه است و بزرگترین ۶ مشاهده بیشتر از دو برابر میانه هستند. بدون اطلاعات بیشتر، نمی‌توانیم زمان بقای مورد مبتلایی را دقیقاً پیشگویی کنیم. به‌علاوه شیوه‌های $t$ی استاندارد را نباید برای استنباط در بارهٔ زمانهای بقا به‌کار برد. در استنباط می‌توان یک توزیع غیرنرمال را به‌عنوان مدل به‌کار برد یا به دنبال تبدیل به مقیاسی بود که بیشتر نرمال باشد.

هر چند بیشتر دانشجویانی که وارد نخستین درس آمار می‌شوند، انتظار یک درس تشریفاتی توخالی را دارند، EDA به آنها این شگفتی دلپسند را عرضه می‌کند که روشهایی موجودند که در یافتن معنا به‌کار می‌آیند. این شگفتی چنان به مذاق آنها خوش می‌آید که خطر افتادن از آن طرف بام را دارد. بعضی از دانشجویان ممکن است به سمت این اعتقادِ دلخوش‌کننده رانده شوند که هر داستانی در بارهٔ داده‌ها که تطابق آن با الگوها قابل‌قبول به نظر رسد باید درست باشد. حالا زمان برای ترکیب مناسبی از طرح و تردیدگری مناسب است.

**۴.۳ طرح را به‌عنوان پلی بین تحلیل داده‌ها و استنباط تدریس کنید.** بحث مقدماتی در طرح تولید داده‌ها به‌طور طبیعی بین تحلیل کاوشی و استنباط جا دارد: طرح بی‌نقص چیزی است که استنباط را ممکن می‌سازد. منتظر ماندن برای معرفی توزیعهای احتمال تا زمانی که مواد پایه‌ای طرح عرضه شده باشد چند مزیت دارد. از یک نظر، این ترتیب به روشن شدن این مطلب کمک می‌کند که توجیه مدلهای احتمال باید براساس تصادفی بودن فرایند تولید داده‌ها باشد، و بنابراین حفاظی در مقابل پذیرشِ بدون تأملِ مدلهای احتمال است. از جنبهٔ دیگر، یادگیری مطالبی در بارهٔ تولید داده‌ها، دانشجویان را پیش از مواجهه با توزیع نمونه‌گیری که فی‌نفسه مشکل است، با مفاهیم ضروری مانند جمعیت و نمونه، پارامتر و آماره آشنا می‌کند.

مهمترین نکته برای دانشجویان این است که بفهمند چرا آزمایشهای مقایسه‌ای تصادفی، «محک‌زرین» برای بررسی شواهد علیت است. یک منبع غنی از حکایتهای هشداردهندهٔ مربوط به زندگی واقعی کتاب [۶] است که توسط دو پزشک یعنی بانکر و بارنس و یک آماردان یعنی موستلر ویرایش شده است و مثالهای تکان‌دهنده‌ای از شیوه‌های درمان پزشکی در بر دارد که پیش از آنکه علم پزشکیْ آزمایشهای مقایسه‌ای تصادفی‌شده را بپذیرد، به‌صورت استاندارد درآمده بودند اما پس از اینکه در معرض آزمون صحیح قرار گرفتند، بی‌ارزش بودن آنها آشکار شد.

البته جنبهٔ آماری طرح آزمایشها و بررسیهای نمونه‌ای چیزی بیش از این دربردارد. طرحهایی که درعمل به‌کار می‌روند اغلب کاملاً پیچیده‌اند، و باید بین کارآیی از یک طرف و نیاز به اطلاعاتی با دقت متفاوت در بارهٔ بسیاری از عاملها و اثر متقابل آنها از طرف دیگر، توازن برقرار شود. طرحهای ساده ــ آزمایشهای تصادفی که دو یا چند تیمار را مقایسه می‌کنند، یا نمونه‌های تصادفی ساده از یک یا چندین جامعه ــ روشنگر مهمترین ایده‌ها و مؤیّدِ استنباطی هستند که در نخستین درس آمار مقدماتی تدریس می‌شود. شما باید در بارهٔ این طرحها صحبت کنید، اما نیازی نیست که از این فراتر روید. برخی مطالب مهم دیگر، برای مثال، شیوه‌های طرح و آزمون کردن پرسشهای بررسی[1] و شیوه‌های آموزش دادن و نظارت بر مصاحبه‌کننده‌ها، معمولاً در درسهای آماری ارائه نمی‌شوند. دانشجویان آمار باید آگاه باشند که این مهارتهای عملی مهم هستند، و جریان تولید داده‌ها، حتی وقتی‌که با یک طرح بی‌نقص آماری شروع می‌کنیم، ممکن است دچار کجروی شود. اینکه در اینجا چقدر وقت صرف کنید، موضوعی است که به صلاحدید شما در مورد نیازهای مخاطبان بستگی دارد.

**۵.۳ استنباط: دو مانع در برابر فهمیدن.** بخش ۳.۲ مختصراً نحوهٔ کار استنباط را توصیف کرده است. از آنجا که جزئیات در عمل به‌صورت سرراست و مکانیکی درآمده‌اند، دوست داریم که دانشجویان بیشتر کوشش خود را صرف درک ایده‌ها کنند. درک این ایده‌ها آسان نیست. اولین مانعْ مفهوم توزیع نمونه‌گیری است. موردی ساده همچون استفاده از نسبت $\hat{p}$ی یک نمونه از کارگران بیکار برای برآورد نسبت $p$ی کارگران بیکار در کل جامعه را انتخاب کنید. مثالهای فیزیکی (نمونه‌گیری مهره‌ها از یک جعبه)، شبیه‌سازی‌های

1. survey questions

کامپیوتری و آزمایشهای فکری همه به انتقال این ایده که به‌ازای نمونه‌های بسیار، مقادیر بسیار $\hat{p}$ داریم کمک می‌کنند. دائم بپرسید «چه اتفاقی خواهد افتاد اگر این کار را به دفعات بسیار انجام دهم؟» این پرسش کلید منطق استنباط آماری استاندارد است.

به محض اینکه ایدهٔ توزیع نمونه‌گیری دریافته شد، ابزارهای تحلیل داده‌ها به ما کمک می‌کنند تا گامهای بعدی را برداریم. در مواجهه با هر توزیع، در بارهٔ شکل، مرکز، و پراکندگی سؤال می‌کنیم. شکل توزیع نمونه‌گیری $\hat{p}$ تقریباً نرمال است. میانگین آن برابر با نسبت جامعه‌ایِ نامعلوم، $p$، است. این مطلب حاکی از آن است که $\hat{p}$ به‌عنوان برآوردگر $p$، اریبی یا خطای سیستماتیک ندارد. دقت برآوردگر با پراکندگی توزیع نمونه‌گیری توصیف می‌شود که آن را به یمن نرمال بودن برحسب انحراف معیار آن اندازه‌گیری می‌کنیم.

دومین مانع عمدهٔ استدلال، آزمونهای معنی‌داری است. هر چند ایدهٔ اصلی («آیا این برآمد تعجب‌آور است؟») بغرنج نیست، جزئیات موضوع مرعوب‌کننده است. هیچ راه گریزی از فرضهای صفر و مقابل و آزمونهای یکطرفه در مقابل دوطرفه وجود ندارد. منطق آزمون کردن که با چنین عبارتی شروع می‌شود: «موقتاً فرض کنید عامل مؤثری که در جستجوی آن هستیم، حضور ندارد»، سرراست نیست. دوست داریم که اکثر دانشجویان ما ایدهٔ توزیع نمونه‌گیری را بفهمند؛ می‌دانیم که بسیاری از آنها استدلال آزمونهای معنی‌داری را نخواهند فهمید. پس قدری کوتاه می‌آییم ولی پافشاری می‌کنیم که باید بتوانند معنای $p$ـمقدار تولیدشده به‌وسیلهٔ نرم‌افزار یا گزارش‌شده در یک مجلهٔ علمی را بیان کنند. این بخشی از توصیهٔ کلی ماست که اصرار داریم دانشجویان شرحهای خیلی مختصری از یافته‌های آماری بنویسند. «در این بررسی، دو روش تدریس روخوانی به دانش‌آموزان کلاس سوم مقایسه می‌شود. یک آزمون $t$ی دو نمونه‌ای که میانگین نمرات دو گروه تیمار در یک آزمون روخوانی استاندارد را با هم مقایسه می‌کند، $p$ـمقداری برابر ۰٫۰۱۹ داشته است، یعنی در این بررسی، اثری به آن بزرگی مشاهده شده است که تنها حدود ۲٪ مواقع ممکن است به‌طور شانسی رخ دهد. این گواه بسیار قوی حاکی از آن است که روش جدید منجر به نمرهٔ میانگین بیشتری در مقایسه با روش استاندارد می‌شود.»

اما دو نکتهٔ پایانی در بارهٔ استنباط: نخست اینکه، درک مفهومیِ ایده‌ها تقریباً تصویری است، یعنی مبتنی است بر کشیدنِ شکلِ توزیع نمونه‌گیری و استفاده از تاکتیکهایی که در تحلیل داده‌ها آموخته می‌شوند. حتی مقدار بسیار زیادی ریاضیات صوری نمی‌تواند جایگزین این دید تصویری شود، و هیچ مقدار استنتاج ریاضی نمی‌تواند به دانشجویان ما کمک کنند که به این دید دست یابند. ریاضیات در دانستن حقایق جنبهٔ اساسی دارد، اما این امر به معنای آن نیست که ما باید ریاضیات را به دانشجویانمان تحمیل کنیم.

دوم اینکه می‌خواهیم آموخته‌های دانشجویان خیلی بیشتر از تصویر کلی مطلب و چند دستورالعمل برای اجرای آن در زمینه‌های مشخص باشد. ذیلاً چند نکتهٔ دیگر، هم عملی و هم نظری، تقریباً به ترتیب اهمیت ذکر می‌شوند. اینکه تا کجای فهرست باید پیش روید، بستگی به مخاطبان شما دارد.

- مطالعهٔ شیوه‌های مشخصی از استنباط، خصوصیاتی را آشکار می‌کند که بین همهٔ آنها مشترک است و تمام دانشجویان باید آنها را بفهمند. به‌دست آوردن اطمینان بیشتر براساس داده‌هایی واحد، به بهای حاشیهٔ خطای بزرگتر صورت می‌گیرد. حتی اثراتی آن‌قدر کوچک که عملاً بی‌اهمیت هستند، از نظر آماری بسیار بااهمیت‌اند هرگاه آزمون معنی‌داری را بر مبنای نمونه‌ای بسیار بزرگ انجام دهیم.
- بسیاری چیزها ممکن است نادرست درآیند به‌طوری که در ارزش استنباط ایجاد تردید شود. مقایسهٔ آزمودنیهایی که داوطلب مصرف کلسیم می‌شوند در مقابل دیگرانی که چنین نمی‌کنند، چیز زیادی در بارهٔ اثرات کلسیم به ما نمی‌گوید، زیرا آنها که داوطلب مصرف کلسیم می‌شوند، ممکن است عموماً از افرادِ مراقب سلامتی خود باشند. یک دادهٔ بسیار دورافتاده می‌تواند نتایج آزمایش پزشکی ما را به هر سو منحرف کند و باز استنباط را بی‌اعتبار نماید. جریان تولید داده‌ها را بازبینی کنید. نمودار داده‌ها را رسم کنید. پس از آن، اگر شد، به سراغ استنباط بروید.
- خود شیوه‌های استنباط به ما نمی‌گویند که چیزی اشتباه شده است. برای مثال، حاشیهٔ خطا در یک بازهٔ اطمینان، تنها تغییرات شانسی در نمونه‌گیری تصادفی را در بردارد. همان‌طور که روزنامهٔ نیویورک تایمز معمولاً در کادری همراه با نتایج نظرسنجی خود ذکر می‌کند، «علاوه بر خطای نمونه‌گیری، مشکلات عملی اجرای هرگونه سنجش افکار عمومی ممکن است منابع دیگری از خطا را وارد نظرسنجی کند.»
- شیوه‌های رایج استنباط درواقع بر مدلهای ریاضی مبتنی هستند، نظیر استنباطی که در مثال پزشکی ما دیده می‌شود: $X_1$، $X_2$، $\cdots$، $X_n$ مستقل و دارای توزیع یکسان $N(\mu_1, \sigma_1)$، و $Y_1$، $Y_2$، $\cdots$، $Y_m$ مستقل و دارای توزیع یکسان $N(\mu_2, \sigma_2)$ هستند. این مدل دقیقاً درست نیست، آیا مفید است؟ در واقع شیوه‌های $t$ی دو نمونه‌ای که وقتی می‌خواهیم $\mu_1$ و $\mu_2$ را مقایسه کنیم از این مدل ناشی می‌شوند، در مقابل غیرنرمال بودن استوارند، در نتیجه این مدل عملاً به شیوه‌های مفید منجر می‌شود. اما آمارهٔ نسبت واریانسهای $F$ برای مقایسهٔ $\sigma_1$ و $\sigma_2$ به غیرنرمال بودن بسیار حساس است تا آن حد که از ارزش عملی چندانی برخوردار نیست. حتی مبتدیان لازم است از وجود چنین مشکلاتی آگاه شوند.
- اغلب در شرایطی می‌خواهیم استنباط را انجام دهیم که داده‌هایمان از یک نمونهٔ تصادفی یا یک آزمایش مقایسه‌ای تصادفی نیامده‌اند. برای مثال به اندازه‌گیری روی قطعه‌های متوالی که در یک خط مونتاژ در حرکت‌اند، فکر کنید. استنباط براساس یک مدل احتمال برای فرایندی که داده‌هایمان را تولید می‌کنند، قابل توجیه است، و درستی مدل را می‌توان تا حدی از خود داده‌ها ارزیابی کرد. تولید تصادفی داده‌ها نمونهٔ عالی و امن‌ترین زمینه برای استنباط است، اما تنها زمینهٔ مجاز نیست.
- استنباط استقراییِ از داده‌ها از نظر مفهومی پیچیده است. شگفت‌آور نیست که راههای دیگری برای اندیشیدن در بارهٔ آن وجود دارد. نظریهٔ استاندارد آماری تمایل به این برداشت از استنباط دارد که هدف از آن تصمیم‌گیری است. برای مثال، آزمون باید بین فرضهای صفر و مقابل تصمیم بگیرد. این کار فوراً به خطاهای نوع I و II و غیره منجر می‌شود. رهیافت تصمیم‌گیری به زحمت با منطق این سؤال: «آیا این برآمد تعجب‌آور است؟» که برحسب $p$ـمقدارها ابراز می‌شود، سازگاری دارد. به نظر ما ارزیابی قدرت شواهد، هدف متداولتری است تا اتخاذ یک تصمیم، اما همه با این نظر موافق

نیستند. مکتب فکری بیزی با ارائهٔ توصیف صریحی از اطلاع پیشین موجود در هر وضعیت آماری و با ترکیب اطلاع پیشین با داده‌ها برای رسیدن به یک تصمیم، پا را فراتر از این می‌گذارد. تقریباً همهٔ آماردانان فکر می‌کنند که این ایده، گاهی خوب است. پیروان مکتب بیز فکر می‌کنند که همهٔ مسائل آماری را می‌توان با این الگو وفق داد. این موضع اقلیتی از آماردانان است (که سخت به آن باور دارند). کار مشکلی در پیش رو داریم.

**۶.۳ در مورد احتمال چطور؟** احتمال بخشی اساسی از هر برنامهٔ آموزش ریاضیات است. احتمال حوزه‌ای پرظرافت و قدرتمند از ریاضیات است که تأثیرات متقابل آن با دیگر حوزه‌های ریاضیات، به کل این موضوع غنا می‌بخشد. احتمال همچنین برای مطالعهٔ جدی ریاضیات کاربردی و مدلبندی ریاضی جنبهٔ اساسی دارد. قلمرو تعیّن‌گرایی در پدیده‌های طبیعی و اجتماعی محدود است، بنابراین توصیف ریاضی رفتار تصادفی باید نقش عمده‌ای در توصیف جهان ایفا کند. خواه مذاق ریاضی ما مایل به نابگرایی باشد و خواه به مدلسازی، احتمال هر دو را ارضا می‌کند. با این حال در اینجا، بحث ما در بارهٔ آمار مقدماتی است و نه ریاضیات.

از دیدگاه منطق قیاسی که بخش زیادی از آموزش آمار را در گذشته شکل داده است، احتمال اساسیتر از آمار است: احتمالْ مدلهای شانسی را که تغییرپذیری در داده‌های مشاهده شده را توصیف می‌کنند، به‌دست می‌دهد. ولی از نظر فهم بهتر موضوع، ما معتقدیم که آمار اساسیتر از احتمال است: در حالی که تغییرپذیری داده‌ها را مستقیماً می‌توان دریافت، مدلهای شانسی را تنها بعد از آنکه آنها را در ذهنمان ساختیم می‌توان دریافت. در دنیای ریاضی آرمانی افلاطونی، می‌توانیم با یک جوجهٔ احتمالاتی شروع کنیم و از منطق قیاسی استفاده کنیم تا یک تخم‌مرغ آماری بگذارد، اما در دنیای درهم‌وبرهم علم تجربی، باید از تخم‌مرغ به‌عنوان دادهٔ مشاهده شده شروع کنیم و یک جوجهٔ احتمالاتی پیشین را به‌عنوان یک استنباط بسازیم. در درس آمار مقدماتی تنها ارزش جوجه در آن است که توضیح دهد تخم‌مرغها از کجا آمده‌اند. حداقل در این چارچوب، قدری نامنصفانه به نظر می‌رسد که از دانشجویان مبتدی بخواهیم مطالبی در بارهٔ تخم‌گذار یاد بگیرند پیش از اینکه با خود تخم‌مرغ آشنا شوند. با قدری مبالغه می‌توان گفت این کار شبیه آن است که مطالعهٔ شیمی با مکانیک کوانتومی آغاز شود.

بنابراین، جای احتمال در نخستین درس آمار کجا باید باشد؟ موضع ما استاندارد نیست، ولی بتدریج هوادارانی می‌یابد: نخستین درسها در آمار اساساً نباید به‌هیچ عنوان نظریهٔ احتمال رسمی را در برداشته باشند.

چرا؟ اولاً به این دلیل که احتمالِ غیررسمی برای درک مفهومی استنباط کافی است. هر چند ساختار نظری استنباط آماریِ استاندارد مبتنی بر احتمال است، نقش احتمال محدود به پاسخ‌دادن به این سؤال است که «چه اتفاقی می‌افتد اگر این روش را دفعات بسیار زیادی به‌کار ببریم؟» پاسخ را توزیع نمونه‌گیری یک‌آماره به‌دست می‌دهد که الگوی تغییرپذیری برآمدهایِ، مثلاً، نمونه‌های تصادفی بسیاری را از جامعهٔ یکسان، ثبت می‌کند. اگر توافق کنیم که استخراج عملی این توزیعها بهتر است به درسهای پیشرفته‌تر موکول شود، می‌توان آنها را به‌عنوان توزیعهایی در نظر گرفت که با استفاده از ابزارهای تحلیل داده‌ها بدون ادوات احتمال رسمی به‌دست آمده‌اند. قواعد مربوط به $P(A\cup B)$ چیز بسیار کمی به یک درس آمار می‌افزاید.

دومین دلیل برای اجتناب از احتمال رسمی آن است که احتمال از نظر مفهومی مشکلترین موضوع در ریاضیات مقدماتی است. تاریخ ایده‌های احتمالاتی [۱۶] و [۲۷] را ببینید) مجذوب‌کننده اما کمی ترس‌آور است. این موضوع مدتها برای ذهنهایی قویتر از ذهن ما فوق‌العاده گیج‌کننده بوده است. روانشناسانی که اولینِ آنها تورسکی[1] و همکارانش بوده‌اند، ثابت کرده‌اند که این سردرگمی حتی در میان آنهایی که می‌توانند اصول احتمال رسمی را از بر بگویند و می‌توانند تمرینات کتابهای درسی را انجام دهند، ادامه دارد. احساس شهودی ما از رفتار تصادفی شدیداً و به‌طور سیستماتیک، معیوب است، برای مثال [۲۸] و مجموعهٔ [۱۹] را ببینید. بدتر اینکه متخصصان آموزش ریاضی هیچ راه مؤثری برای تصحیح این شهود معیوب ما نیافته‌اند. گارفیلد و الگرن [۱۵] مروری بر یک پژوهش در این زمینه را با این اظهارنظر به پایان می‌رسانند که «تدرس مفهوم احتمال هنوز هم کار بسیار دشواری به‌نظر می‌آید، و مملو از ابهام و تصورات غلط است.» آنها پیشنهاد می‌کنند بررسی شود که «چگونه ایده‌های مفید استنباط آماری را می‌توان مستقل از احتمالِ درست از نظر فنی، تدریس کرد». به اعتقاد ما تمرکز بر ایدهٔ توزیع نمونه‌گیری اجازهٔ این کار را، حداقل در سطحی مناسب برای مبتدیان، به ما می‌دهد.

البته مفاهیم استنباط آماری، اگر هم مبتنی بر توزیعهای نمونه‌گیری عرضه شوند، دشوارند. باید توجه خود و توجه دانشجویان خود را ــ که بردباری محدودی نسبت به ایده‌های مشکل دارند ــ به ایده‌های اساسی آمار محدود کنیم. ما اعضای هیأت علمی تصور می‌کنیم که احتمال رسمی این ایده‌ها را روشن می‌کند. اما این امر تقریباً در مورد هیچ‌یک از دانشجویان ما صادق نیست.

**۷.۳ در مورد دانشجویان رشتهٔ ریاضی چطور؟** دانشجویان رشتهٔ ریاضی به‌طور سنتی آمار را به‌عنوان دومین درس از یک سلسله درس دو ترمی که به نظریهٔ احتمال و آمار اختصاص دارد می‌گذرانند. امیدواریم واضح باشد که ما سیاحتی در آماره‌های بسنده، نااریبی، برآوردگرهای ماکسیمم درستنمایی، و قضیهٔ نیمن پیرسون[2] را راه نویدبخشی برای کمک به دانشجویان در فهمیدن ایده‌های اساسی آمار تلقی نمی‌کنیم. از طرف دیگر، دانشجویان رشتهٔ ریاضی حتماً باید بخشی از ساختارهای استنباط آماری را ببینند. پس چه باید کرد؟

ما ترجیح می‌دهیم که پیش از مطرح کردن نظریه، ایده‌ها و روشهای آماری و کاربردهای آنها با محور قراردادن داده‌ها به دانشجو معرفی شود. یعنی دانشجویان ریاضی لزوماً از این اصل که نخستین آشنایی با آمار نباید براساس احتمال رسمی باشد، مستثنی نیستند. اگر دانشجویان زمینه‌های محاسباتی قوی دارند، یک درس داده‌محور می‌تواند با سرعت کافی به سمت معرفی آمار واقعاً مفید و کاربردهای جدی پیش رود. نیاز به نظریه را می‌توان زمانی‌که با مسائل عملی مواجه می‌شویم، روشن کرد، و نظریه وقتی‌که وضعیت آن در عمل روشن باشد، بهتر معنا می‌یابد. با این حال در بسیاری مؤسسات، محدودیتها یا تردید اعضای هیأت علمی این مسیر را دشوار می‌سازد. در برخی دیگر از مؤسسات، هماهنگی کمی بین درسهای «کاربسته» و نظری وجود دارد، به‌طوری که درواقع دروس نظری مبتنی بر درسهای کاربردی نیستند.

1. Tversky 2. Neyman Pearson

بنابراین لازم است مجدداً این موضوع را مورد توجه قرار دهیم که یک درس یک‌ترمی در معرفی آمار برای دانشجویان ریاضی و دیگر دانشجویانی که پایهٔ محاسباتی قوی دارند چگونه باید باشد. راحت‌ترین کار این است که این درس پس از درسی در احتمال بیاید و معمولاً هم چنین است. در اینجا با مانعی دیگر مواجه هستیم: وجداناً نمی‌توانیم سلسله درس استاندارد احتمال-آمار را در هر دو ترم طوری برنامه‌ریزی کنیم که آشنایی با آمار به‌طور بهینه انجام گیرد. احتمال نه تنها به منظور آماده‌سازی برای نظریهٔ آماری بلکه فی‌نفسه هم مهم است. هر چقدر که یک گروه در برنامهٔ درسی رشتهٔ خود تأکید بیشتری بر کاربردها و مدلسازی داشته باشد، درس احتمال باید نقش اساسیتری در این تأکید ایفا کند. مقدمه‌ای بر احتمال که بر مدلسازی تأکید کند و شبیه‌سازی و محاسبات عددی را داشته باشد، مطمئناً زمینه را برای آمار آماده می‌کند، اما نسبت به انتقال ایده‌های صرفاً آماری به ترم احتمال در تردید هستیم. اصلاح احتمال و اصلاح آمار دو مقولهٔ متمایزند.

هدف ما باید یک درس یکپارچهٔ آمار باشد که به نوبت از درون تحلیل داده‌ها، تولید داده‌ها، و استنباط گذر کرده بر اصول سازمان‌دهندهٔ هر یک از آنها تأکید کند. حتماً باید از ظرفیتهای ریاضی دانشجویان سود برده آنها را تقویت کنیم. گرچه تحلیل داده‌ها و تولید داده‌ها نظریهٔ یکپارچه‌کننده‌ای ندارند، تحلیل ریاضی حتی می‌تواند به تحلیل داده‌ها وضوح بخشد. در اینجا چند مثال را ذکر می‌کنیم.

- الف: خواص بهینگی شاخصهای گرایش به مرکز را برای $n$ مشاهده در نظر گیرید. میانگین، میانگین توان دوم خطا را مینیمم می‌کند؛ میانه، میانگین قدر مطلق خطا را مینیمم می‌کند (و لزومی ندارد یکتا باشد)؛ نیم دامنه، ماکسیمم مطلق (یا توان دوم) خطا را مینیمم می‌کند؛ سعی کنید میانهٔ مطلق خطا را برای $n = ۳$ مینیمم کنید و رفتار ناخوشایند شاخص حاصل را بررسی کنید.
- ب: دانشجویان ضمن مطالعهٔ احتمال به نابرابری چبیشف برمی‌خورند. سپس ممکن است همچنین با نابرابری جالب $|\mu - m| \leq \sigma$ که میانگین، میانه، و انحراف معیار هر توزیع را با هم مرتبط می‌کند [۲۹] روبه‌رو شوند. داده‌های یک نمونه‌ای را به کمک توزیع تجربی (احتمال $\frac{1}{n}$ در هر نقطهٔ مشاهده شده) برای استخراج نتایجی در بارهٔ اینکه میانگین و میانهٔ نمونه چقدر از هم فاصله دارند، توصیف کنید.
- پ: خط رگرسیونی کمترین توانهای دوم، مشابه میانگین $\overline{x}$ برای پیش‌بینی $y$ از روی $x$ است. آن را به‌دست آورید. سپس، اگر خواستید، به‌کمک نرم‌افزار در بارهٔ مشابه‌های سایر شاخصهای ذکر شده در الف کاوش کنید.

در مورد تولید داده‌ها به‌راحتی می‌توان محاسباتی احتمالاتی انجام داد که نشان دهند چقدر محتمل است که تخصیصهای تصادفی به طرق مشخصی نامتوازن باشند؛ مزایای نمونه‌های بزرگ به‌زودی روشن می‌شود.

بسیار خوب. می‌توانیم درس مقدّماتی متوازنی در آمار به دانشجویان ارائه دهیم که در آن از معلومات ریاضی آنها استفاده شود. پیامد گریزناپذیر آن این است که کمتر در استنباط وقت صرف کنیم. باید تصمیم بگیریم چه چیزی را حفظ و چه چیزی را حذف کنیم. در این زمینه هنوز اجماعی حاصل نشده است، زیرا علی‌رغم غرولندهای فراوان، اصلاح سلسله دروس احتمال و آمارِ رشتهٔ ریاضی هنوز شروع نشده است. فکر کردن در بارهٔ چنین اصلاحی نقطه‌ای مناسب برای پایان دادن به بحث آمار، ریاضیات، و تدریس است. این تکلیفی است که باید در خانه انجام دهید: یک درس آمارِ یک ترمیِ بهتر برای دانشجویان رشتهٔ ریاضی طراحی کنید.

## مراجع


1. Aldous. David (1994), Triangulating the circle. at random, *Amer. Math. Monthly* **101**, 223-233. The remark appears in the biographical note accompanying the paper.
2. Bjerkedal, T. (1960), Acquistion of resistance in guinea pigs infected with different doses of virulent tubercle bacilli, *American Journal of Hygiene* **72**, 130-148.
3. Boyer, Paul and Stephen Nissenbaum (1972). *Salem Village Witchcraft.* Belmont, CA: Wadsworth Publishing Co.
4. Boyer, Paul and Stephen Nissenbaum (1974). *Salem Possessed.* Cambridge, MA: Harvard University Press.
5. Bullock, James O. (1994), Literacy in the language of mathematics, *Amer. Math. Monthly* **101**, 735-743.
6. Bunker, John P., Benjamin A. Barnes. and Frederick Mosteller (eds.) (1977), *Costs. Risks and Benefits of Surgery.* New York: Oxford University Press.
7. Chambers, John M., William S. Cleveland, Beat Kleiner, and Paul A. Tukey (1983). *Graphical Methods for Data Analysis.* Belmont, CA: Wadsworth.
8. Chambers, John M. and Trevor J. Hastie (1992). *Statistical Model in* S. Pacific Grove, CA: Wadsworth.
9. Cleveland, William S. and Mary E. McGill (eds.) (1988), *Dynamic Graphics for Statistics.* Belmont, CA: Wadsworth.
10. Cobb, George W. (1991), Teaching statistics: more data, less lecturing, *Amstat News.* December 1991, pp. 1, 4.
11. Cobb, George W. (1992), Teaching statistics, in L. A. Steen (ed.) *Heeding the Call for Change: Suggestions for Curricular Action,* MAA Notes 22. Washington, DC: Mathematical Association of America.
12. Cochran, W. G. (1968), The effectiveness of adjustment by subclassification in removing bias in observational studies, *Biometrics* **24**, 205-213.
13. Crystal, David (ed.) (1994), *The Cambridge Factfinder.* Cambridge: Cambridge University Press. pp. 174-175.
14. Efron, Bradley and Rob Tibshirani (1991), Statistical data analysis in the computer age. *Science* **253**, 390-395.
15. Garfield, Joan and Andrew Ahlgren (1988), Difficulties in learning basic concepts in probability and statistics: implications for research. *Journal for Research in Mathematics Education* **19**, 44-63.

16. Gigerenzer, G., Z. Swijtink. T. Porter, L. Daston, J. Beatty, and L. Krüger (1989) *The Empire of Chance.* Cambridge: Cambridge University Press.

17. Hoaglin, D. C. (1992), Diagnostics. in D. C. Hoaglin and D. S. Moore (eds.), *Perspectives on Contemporary Statistics*, MAA Notes 21. Washington, DC: Mathematical Association of America, pp. 123-144.

18. Hoaglin, David C. and David S. Moore (eds.) (1992). *Perspectives on Contemporary Statistics*, MAA Notes 21. Washington, DC: Mathematical Association of America.

19. Kapadia, R. and M. Borovcnik (eds.) (1991), *Chance Encounters: Probability in Education.* Dordrecht: Kluwer.

20. Longfellow, Henry Wadsworth (1847). *Evangeline*, Introduction, 1.1.

21. Lyle, Roseann M. et al. (1987), Blood pressure and metabolic effects of calcium supplementation in normotensive white and black men, *Journal of the American Medical Association* **257**, 1772-1776. Dr. Lyle provided the data in the example.

22. Moore, David S. (1988), Should mathematicians teach statistics (with discussion). *College Math. Journal* **19**, 3-7.

23. Moore, David S. (1992), What is statistics? in David C. Hoaglin and David S. Moore (eds.), *Perspectives on Contemporary Statistics*, MAA Notes 21. Washington, DC: Mathematical Association of America, pp. 1-18.

24. Moore, David S. (1992), Teaching statistics as a respectable subject, in Florence Gordon and Sheldon Gordon (eds.), *Statistics for the Twenty-First Century*, MAA Notes 26. Washington, DC: Mathematical Association of America.

25. Moore, David S. (1995), *The Basic Practice of Statistics.* New York: W. H. Freeman.

26. Rosenbaum, Paul R. (1995), *Observational Studies.* New York: Springer-Verlag, p. 60.

27. Stigler, S. M. (1986), *The History of Statistics: The Measurement of Uncertainty Before 1900.* Cambridge, Mass: Belknap.

28. Tversky, Amos and Daniel Kahneman (1983), Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review* **90**, 293-315.

29. Waston, G. S. (1994), letter to the editor, *The American Statistician* **48**, p. 269. This is the last in a sequence of comments on this inequality, and contains references to the earlier contributions.

30. Weisberg, Sanford (1985), *Applied Linear Regression*, 2nd edition. New York: John Wiley and Sons, p. 230.

******

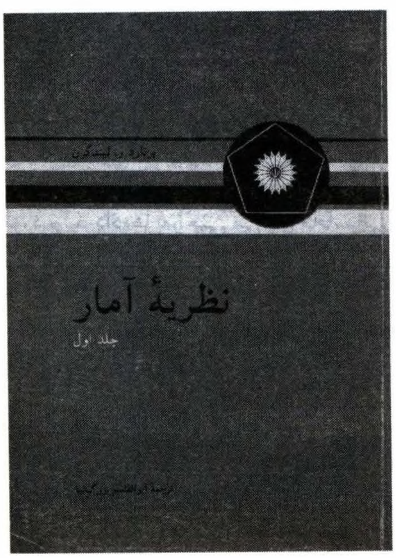
• George W. Cobb and David S. Moore, "Mathematics, statistics, and teaching", *Amer. Math. Monthly*, (9) **104** (1997) 801-823.

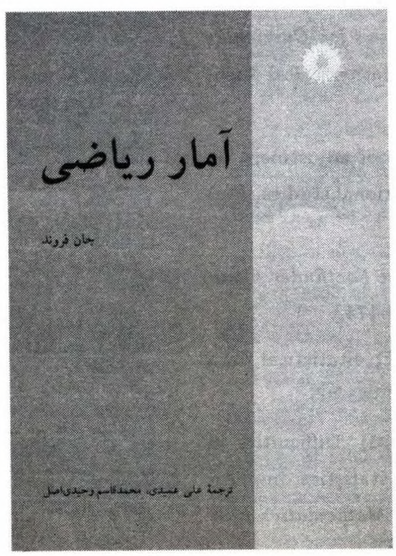
* جورج کاب، کالج مانت هالیوک، آمریکا

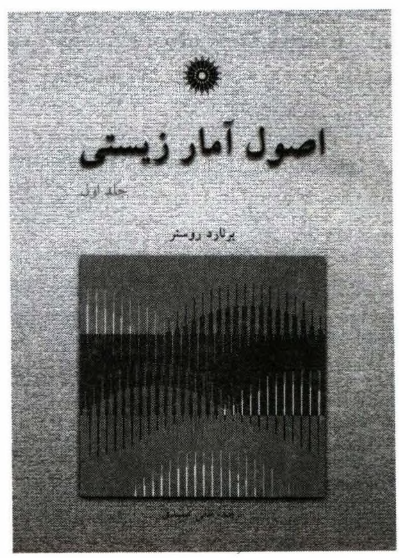
gcobb@mtholyoke.edu

* دیوید مور، دانشگاه پُردو، آمریکا

dsm@stat.purdue.edu



# نسبیت خاص به انضمام شتاب•

گری هلزر*
ترجمهٔ جهانشاه میرزابیگی

## ۱. مقدمه

ریک نوروود در شمارهٔ فوریهٔ ۱۹۹۲ مانتلی خاطرنشان کرد که در دایره‌ای که با سرعت کافی حول محور خود می‌چرخد انقباض لورنتس‌ـفیتزجرالد، مقدار $\pi$ (خارج‌قسمت محیط بر قطر دایره) را به سه کاهش می‌دهد [۱۳، صفحهٔ ۱۱۱]. پروفسور داریو کاستلانوس، در نامه‌ای به ویراستار در شمارهٔ ژانویهٔ ۱۹۹۳ این مجله اظهار داشت که، برعکس، نسبت محیط به قطر افزایش می‌یابد. نه کاهش، زیرا سنجهٔ اندازه‌گیری محیط منقبض می‌شود [۲، صفحهٔ ۶۹]. وی اضافه می‌کند که هندسهٔ اقلیدسی در دستگاه مختصات چرخان صادق نیست. نوروود در شمارهٔ ژوئن‌ـژوئیهٔ ۱۹۹۳، این ناسازگاری را به یک ابهام نسبت می‌دهد: آیا یک خط‌کش چرخان برای اندازه‌گیری دایرهٔ ثابت به‌کار می‌رود یا یک خط‌کش ثابت برای اندازه‌گیری دایرهٔ چرخان [۱۴، صفحهٔ ۵۷۷]؟ وی در ادامه می‌گوید که این تحلیل نادرست است زیرا تحلیل دستگاههای شتابدار مستلزم نظریهٔ نسبیت عام است نه نسبیت خاص. سپس سؤالی را مطرح می‌کند که نمی‌تواند به آن پاسخ دهد. فرض کنید یک قطار به‌قدری دراز است که تمامی یک مسیر دایره‌ای را اشغال می‌کند؛ یعنی سر و ته قطار در مسیر به هم می‌رسد. در سرعتهای خیلی زیاد طول قطارِ کوتاه شده مسیر بلند را اشغال می‌کند. چگونه چنین چیزی ممکن است؟

این بگو مگو گویای این واقعیت است که ابزار ریاضی لازم برای مطالعهٔ نسبیت مقدماتی در میان ریاضیدانان، آن‌طور که باید، شناخته شده نیست. درواقع، برای پاسخ‌دادن به سؤالهایی از نوع سؤال نوروود، که در بالا مطرح شد، نظریهٔ نسبیت عام لازم نیست، بلکه فقط نظریهٔ خمینه‌های (نیمه)ریمانی در حضور اجرام گرانشی لازم است. مطالعهٔ نسبیت خاص بدون در نظر گرفتن شتاب به این دلیل نیست که استفاده از شتاب در این بحث قدغن است، بلکه صرفاً به این دلیل است که مؤلف نمی‌خواهد حسابان را پیش‌نیاز آن فرض کند. البته، این هیچ مانعی در درس حسابان محسوب نمی‌شود.

کمتر از یک دههٔ دیگر نظریهٔ نسبیت خاص صدساله می‌شود؛ با وجود این، هنوز این ایدهٔ پراهمیت جایگاهی در بخشهای کاربردی درسهای متعارف ما پیدا نکرده است. دلیل این امر شاید تا حدی این باشد که ریاضیدانان، برخلاف فیزیکدانان، هنوز فکر می‌کنند که نظریهٔ نسبیت خاص را باید با استفاده از تبدیلات لورنتس مطالعه کرد [۲۰، صفحهٔ ۹۵]. رهیافت مبتنی بر تبدیلات لورنتس با کاربرد حسابان ناسازگار است و برای پاسخ دادن به سؤالهایی از نوع سؤال نوروود مناسب نیست.

در اینجا به جای تبدیلات لورنتس از رهیافت هندسی استفاده می‌کنیم و برای هندسهٔ اقلیدسی، مکانیک نیوتنی و مکانیک نسبیتی رهیافت واحدی به‌کار می‌بریم. برای سهولت، این بررسی به یک بُعد فضا و یک بعد زمان (یعنی، حرکت در امتداد یک خط راست) محدود شده است. این محدودیت تحلیل مستقیم سؤال نوروود را، که مستلزم دو بعد فضا و یک بعد زمان است، منتفی می‌سازد اما در عین حال، موضوعات لازم برای تحلیل شهودی را به قدر کافی روشن می‌کند.

در چندین مورد به‌عنوان مطالبی برمی‌خوریم که می‌توان آنها را جداگانه برای یک مجموعه مسائل، متمم کاربردی یک درس، یا یک بحث کوتاه در درس معادلات دیفرانسیل یا حسابان به‌کار برد.

## ۲. مختصری در بارهٔ خمها در صفحه‌های اقلیدسی، گالیله‌ای و مینکوفسکیایی

این بخش به بحث کوتاهی در هندسهٔ مورد نیاز می‌پردازد و شامل اصلاحات اندکی در نظریهٔ اقلیدسی خمهاست.

### ۱.۲ انحنا در صفحه‌های اقلیدسی، گالیله‌ای و مینکوفسکیایی

$\mathbb{R}^2$ را با صورت دوخطی زیر در نظر می‌گیریم

$$\langle v,w\rangle = v_1w_1+\epsilon v_2w_2 \qquad (1)$$

که در آن $\epsilon$ می‌تواند ۱، ۰، یا ۱− باشد. فاصلهٔ بین دو نقطهٔ $p$ و $q$ با رابطهٔ زیر تعریف می‌شود

$$\| p-q \| = |\langle p-q, p-q\rangle|^{1/2}$$

اگر $\epsilon=0$ و $p$ و $q$ مختصهٔ اول یکسان داشته باشند، فاصلهٔ (علامت‌دار) را تفاضل مختصه‌های دوم تعریف می‌کنیم؛ برای آشنایی با روش کلی، مرجع [۲۱] را ببینید. به‌ازای $\epsilon=1$، صفحهٔ اقلیدسی، به‌ازای $\epsilon=0$، گالیله‌ای و به‌ازای $\epsilon=-1$، مینکوفسکیایی است. در مرجعهای [۲۱] و [۲۲] این سه صفحه را هندسه‌های مسطحهٔ کیلی‌ـکلاین با شاخص سهموی فاصله می‌نامند. $\mathbb{R}^2$ با صورت دوخطی (۱) را به $\mathbb{P}_\epsilon$ نشان می‌دهیم.

طول قوس خم مشتق‌پذیر $\sigma: I\to \mathbb{P}_\epsilon$ از $p=\sigma(a)$ تا $q=\sigma(b)$ را با انتگرال زیر تعریف می‌کنیم

$$L(a,b)=\int_a^b \|\sigma'(u)\|\,du$$

خم مشتق‌پذیر $\sigma: I\to\mathbb{P}_\epsilon$ را، اگر به‌ازای تمامی $u$ها در $I$، $\|\sigma'(u)\|\neq 0$، خم منظم می‌نامیم. طبق معمول، طول قوس هر خم منظم، مستقل از پارامتری‌سازی است و این خمها را می‌توان نسبت به طول قوس پارامتری کرد. برای پارامتر طول قوس $s$ و به‌ازای تمامی $s$ها داریم $\|\sigma'(s)\|=1$.

خم منظم در صفحهٔ اقلیدسی، خمی است با یک بردار مماس که هرگز صفر نمی‌شود. در دو مورد دیگر ویژگی منظم بودن محدودکننده‌تر است. به‌ازای $\epsilon=-1$، بردارهای مماس باید ناصفر و ناموازی با هر یک از بردارهای $(1,\pm1)$ باشند. به‌عنوان مثال، اگر مماس بر یک خم منظم در یک نقطه، بین $(1,-1)$ و $(1,1)$ قرار گیرد، آنگاه تمامی بردارهای مماس بر خم بین این بردارها قرار می‌گیرند. مخصوصاً خمهای بسته نمی‌توانند منظم باشند. اگر $\epsilon=0$، بردارهای مماس بر یک خم منظم باید ناصفر و ناموازی با $(0,1)$ باشند؛ و در اینجا هم خمهای بسته نمی‌توانند منظم باشند.

اگر $\langle v,w\rangle=0$، بردارهای $v$ و $w$ در $\mathbb{P}_\epsilon$ متعامد نامیده و به‌صورت $v\perp w$ نوشته می‌شوند. بردارهای خودْمتعامد را پایدار[1] می‌نامند. به‌ازای $\epsilon=1$، فقط بردار صفر پایدار است. به‌ازای $\epsilon=0$، بردارهای قائم و صفر، و به‌ازای $\epsilon=-1$، بردارهای صفر و بردارهای موازی با $(1,\pm1)$ پایدارند.

تبدیل خطی $J:\mathbb{P}_\epsilon\to\mathbb{P}_\epsilon$ با ماتریس زیر، که آن نیز با $J$ نشان داده می‌شود، هر بردار $v$ را به بردار متعامد $Jv$ می‌برد:

$$J=\begin{pmatrix}0&-\epsilon\\1&0\end{pmatrix}$$

1. isotropic

اگر $v$ پایدار نباشد و $w$ بر $v$ عمود باشد، آنگاه به‌ازای یک عدد حقیقی مانند $k$، به‌راحتی می‌توان درستی رابطهٔ $w=kJv$ را تحقیق کرد.

اکنون فرض کنید که خم منظم $\sigma(s)$ نسبت به طول قوس پارامتری می‌شود. در این صورت، $\pm1=\langle\sigma'(s),\sigma'(s)\rangle$. چون صورتهای دوخطی از قاعدهٔ حاصلضربی زیر پیروی می‌کنند

$$\langle f(s),g(s)\rangle' = \langle f'(s),g(s)\rangle+\langle f(s),g'(s)\rangle$$

نتیجه می‌گیریم $\sigma''\perp\sigma'$. اگر قرار دهیم $\mathrm{T}(s)=\sigma'(s)$ و $\mathrm{N}(s)=J\mathrm{T}(s)$، نتیجه می‌شود که تابعی مانند $\kappa(s)$ وجود دارد که در رابطهٔ زیر صدق می‌کند

$$\mathrm{T}'(s)=\kappa(s)\mathrm{N}(s)$$

تابع $\kappa(s)$ را انحنای $\sigma(s)$ می‌نامند. بردار $\mathrm{T}(s)$، بردار مماس یکه و بردار $\mathrm{N}(s)$، بردار قائم یکه است.

چون $J\mathrm{N}(s)=J^2\mathrm{T}(s)=-\epsilon\mathrm{T}(s)$، مستقیماً به فرمولهای فرنه می‌رسیم:

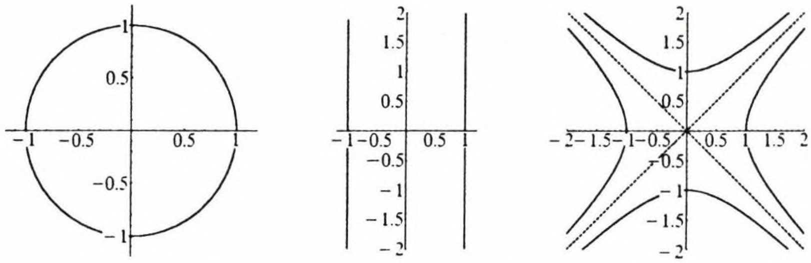
$$\begin{aligned}\mathrm{T}'(s)&=\kappa(s)\mathrm{N}(s)\\ \mathrm{N}'(s)&=-\epsilon\kappa(s)\mathrm{T}(s)\end{aligned}$$

این، برای حالت اقلیدسی ($\epsilon=1$)، تعریف معمولی انحناست. حالتهای گالیله‌ای و مینکوفسکیایی را می‌توان، به ترتیب، برای هندسی کردن مکانیک نیوتنی و نسبیتی به‌کار برد.

**مکانیک نیوتنی.** فرض کنید محل یک ذرهٔ متحرک بر روی یک خط با رابطهٔ $x=f(t)$ داده شود که در آن $t$ زمان است. جهان‌ـخط نیوتنی ذره، خمی در $\mathbb{P}_0$ با ضابطهٔ $\sigma(t)=(t,f(t))$ است. این همان نمایش پارامتری طول قوس است زیرا $\mathrm{T}(t)=\sigma'(t)=(1,f'(t))$ یک بردار یکه در $\mathbb{P}_0$ است. بردار قائم واحد عبارت است از $J\mathrm{T}(t)=(0,1)$. چون $\mathrm{T}'(t)=(0,f''(t))=f''(t)\mathrm{N}(t)$، تابع $\kappa(t)=f''(t)$ تابع انحناست. از این‌رو، در این هندسی‌سازی مکانیک نیوتنی، انحنا شتاب است.

اگر $\sigma(t)=(x(t),y(t))$ هر خم پارامتری‌شده توسط طول قوس در $\mathbb{P}_0$ باشد، آنگاه معادلهٔ $|x'(t)|=\|\sigma'(t)\|=1$ نشان می‌دهد که به‌ازای ثابتی مانند $a$ داریم $x(t)=a\pm t$. برای استفاده از این بحث در مکانیک، توجه خود را به خمهایی به‌صورت $\sigma(t)=(a+t,f(t))$، یعنی به ذراتی که در زمان به جلو پیش می‌روند، محدود می‌کنیم. در این صورت به‌راحتی نتیجه می‌گیریم که طول قوس از $\sigma(t_1)$ تا $\sigma(t_2)$ همان زمان سپری‌شدهٔ $t_2-t_1$ است.

**مکانیک نسبیتی.** در بحثهای نسبیت غالباً چارچوبهای مرجع مختلفی معرفی می‌شود. ما در این مقاله مختصات مفروض $\mathbb{P}_{-1}$ را چارچوب مرجع آزمایشگاه می‌نامیم.




**شکل ۱** دایره‌های یکه، به ترتیب، به‌ازای $\epsilon = 1$، $\epsilon = 0$ و $\epsilon = -1$

باز هم فرض کنید محل یک ذرهٔ متحرک بر روی یک خط با رابطهٔ $x = f(t)$ داده می‌شود که در آن $t$ زمان در چارچوب مرجع آزمایشگاه است. جهان‌ـخط نسبیتی ذره خمی در $\mathbb{P}_{-1}$ با ضابطهٔ $\sigma(t) = (f(t), t)$ است. این نمایش پارامتریِ طول قوس نیست زیرا $\langle \sigma'(t), \sigma'(t)\rangle = [f'(t)]^2 - 1$ در حالت کلی ثابت نیست. قرار می‌دهیم $v = f'(t)$. اگر $\sigma$ منظم باشد، آنگاه $|v|$ نمی‌تواند برابر ۱ باشد. چون ما می‌خواهیم $v = 0$، در هر خم منظم باید داشته باشیم $|v| < 1$. یکاهای به‌کار رفته در این مورد یکاهای هندسی هستند که در آنها زمان و فاصله برحسب یکاهای یکسانی اندازه‌گیری می‌شود و سرعت نور برابر ۱ است؛ مرجعهای [۱۵]، [۲۰]، یا [۱۰] را ببینید.

طول قوس این خم را ویژه‌زمان می‌گویند. ویژه‌زمان مدت زمانی است که ساعت همراه ذره ثبت می‌کند (مقایسه کنید با حالت نیوتنی). فرض کنید $\tau$ ویژه‌زمان باشد. طبق تعریف طول قوس:

$$\frac{d\tau}{dt} = \sqrt{1 - v^2}$$

چون $\sqrt{1 - v^2} < 1$، ساعت همراه ذره در مقایسه با ساعت واقع در چارچوب مرجع کندتر کار می‌کند. این نتیجه به پارادکس دوقلو یا پارادکس ساعت منجر می‌شود.

تعبیر فیزیکی انحنای $\kappa(t)$ شتابی است که ناظر سوار بر ذره می‌گیرد. ما در اینجا $\kappa$ را ویژه‌شتاب می‌نامیم. مسأله، حرکت در امتداد یک خط مستقیم است، بنابراین می‌توانیم شخصی را در نظر بگیریم که سوار بر یک آسانسور بسته است و از یک شتاب‌سنج به‌عنوان وسیلهٔ اندازه‌گیری استفاده می‌کند. ویژه‌شتاب با شتاب مشاهده شده در چارچوب مرجع آزمایشگاه از طریق فرمول زیر مربوط می‌شود

$$\kappa = \frac{f''(t)}{(1 - v^2)^{3/2}}$$

که در آن حرکت با رابطهٔ $x = f(t)$ مشخص می‌شود. این رابطه، نوع مینکوفسکیایی فرمول استاندارد حسابان است (و به همان طریق به‌دست می‌آید).

هنگامی که $\langle \sigma'(t), \sigma'(t)\rangle$ مثبت، صفر، یا منفی است جهان‌ـخط را، به ترتیب، فضاگونه، نورگونه، یا زمان‌گونه می‌نامند. جهان‌ـخط ناظرها زمان‌گونه است. برای ناظرهای لخت، $\kappa(\tau) \equiv 0$. چون برای ناظرهای لخت $\mathrm{T}(\tau)$ ثابت است، جهان‌ـخط آنها یک خط مستقیم است. نقاط $\mathbb{P}_{-1}$ را رویداد می‌نامند و مختصات معین آزمایشگاه در $\mathbb{P}_{-1}$ را، آن‌طورکه توسط یک ناظر لخت با جهان‌ـخط $\sigma(t) = (0, t)$ اندازه‌گیری می‌شوند، مختصات فضا و زمان هر رویداد فرض می‌کنند.

## ۲.۲ دایره‌ها، دورانها، و تابعهای مثلثاتی

دایره مجموعه‌ای از نقاط با فاصلهٔ معین از یک نقطهٔ ثابت، به نام مرکز، است. دایرهٔ یکه در $\mathbb{P}_\epsilon$ عبارت است از مجموعهٔ نقاط $p$ با ضابطهٔ $\|p\| = 1$. دایره‌های یکه در صفحه‌های اقلیدسی، گالیله‌ای، و مینکوفسکیایی در شکل ۱ نشان داده شده است.

دایرهٔ یکهٔ گالیله‌ای دو شاخه دارد که عبارت‌اند از خطهای قائم $x = \pm 1$، و هر نقطه بر محور $y$ یک مرکز است. دایرهٔ یکهٔ مینکوفسکیایی چهارشاخه دارد که متشکل‌اند از یک زوج هذلولی قائم مزدوج با معادله‌های $x^2 - y^2 = \pm 1$.

مجانبهای هذلولیهای دایرهٔ یکهٔ مینکوفسکیایی نیز در شکل ۱ رسم شده‌اند. این مجانبها خمهای نورگونه هستند. بخش $P_{-1}$ میان مجانبها و شامل محور $y$ (محور زمان مکانیک نسبیتی) مخروط نور نامیده می‌شود. بخش شاملِ قسمت مثبت محور $y$ مخروط نور آینده و بخش شاملِ قسمت محور $y$ مخروط نور گذشته است. بردارهای مماس بر خمهای زمان‌گونه (به سمت آینده)، در صورت انتقال به مبدأ، در مخروط نور (آینده) قرار می‌گیرند. از این‌رو، می‌توان دو شاخهٔ دایرهٔ یکه‌ای را که محور $x$ را قطع می‌کنند به‌صورت خمهای زمان‌گونه و دوشاخه‌ای را که محور $y$ را قطع می‌کنند به صورت خمهای فضاگونه پارامتری کرد.

برای یافتن نمایش پارامتری طول قوسِ دایرهٔ یکه $[\sigma(\phi) = (x(\phi), y(\phi))]$، با معادلهٔ زیر شروع می‌کنیم

$$x^2 + \epsilon y^2 = \pm 1$$

با مشتق‌گیری داریم $xx' + \epsilon yy' = 0$ که نشان می‌دهد $\sigma \perp \sigma'$ و در نتیجه $\sigma' = kJ\sigma$. برای نمایش پارامتری طول قوس، $k = \pm 1$. با قراردادن $-\phi$ به‌جای $\phi$ می‌توان جوابهای با $k = -1$ را از جوابهای با $k = 1$ به‌دست آورد. به این ترتیب به معادلهٔ دیفرانسیل ماتریسی $\sigma' = J\sigma$ می‌رسیم که جواب آن، برای تمامی $\phi$های حقیقی، عبارت است از

$$\sigma(\phi) = e^{J\phi}\sigma(0)$$

تبدیل خطی با ماتریس $e^{J\phi}$ دوران به اندازهٔ زاویهٔ $\phi$ نامیده می‌شود. زاویهٔ $\phi$ طول قوس در امتداد دایرهٔ یکه از $\sigma(0)$ تا $\sigma(\phi)$ است. این طول قوس اندازهٔ رادیانی نامیده می‌شود. چون $\sigma(0)$ اختیاری است به نتیجهٔ زیر می‌رسیم.

اگر $u$ و $v$ بردارهای یکه‌ای باشند که ابتدای آنها در مبدأست و روی یک شاخه از دایرهٔ یکه در $\mathbb{P}_\epsilon$ قرار داشته باشند، آنگاه دورانی وجود دارد که $u$ را به $v$ می‌برد.

البته، به‌ازای $\epsilon = 1$ تعداد زیادی از این دورانها وجود دارد، و مقادیر مختلف $\phi$ در ضرایب $2\pi$ با هم فرق می‌کنند. به‌ازای $\epsilon = 0$ و $\epsilon = -1$، $\phi$ منحصر به فرد است.

اکنون فرض کنید $\sigma$ خم منظمی است که نسبت به طول قوس $s$ پارامتری می‌شود. در این صورت $\mathrm{T}(s)$، اگر به‌گونه‌ای انتقال یابد که ابتدای آن در مبدأ باشد، روی دایرهٔ یکه قرار می‌گیرد. فرض کنید $\phi(s)$ زاویهٔ از $\mathrm{T}(0)$ تا $\mathrm{T}(s)$ باشد. در این صورت $\mathrm{T}(s) = e^{J\phi(s)}\mathrm{T}(0)$، و داریم

$$\begin{aligned}\kappa(s)\mathrm{N}(s) = \mathrm{T}'(s) &= \phi'(s)Je^{J\phi(s)}\mathrm{T}(0)\\ &= \phi'(s)J\mathrm{T}(s) = \phi'(s)\mathrm{N}(s)\end{aligned}$$

در این صورت $\kappa(s) = \phi'(s)$ و بنابراین $\kappa(s)$ آهنگ تغییر نسبت به طول قوس زاویه‌ای است که بردار مماس یکه با یک خط ثابت می‌سازد.

در حالت اقلیدسی، این یک نتیجهٔ متعارف است. برای حالت نیوتنی با ضابطهٔ $\sigma(t) = (t, f'(t))$ واضح است که $\phi(t) = f'(t)$ سرعت است. در مکانیک نسبیتی، رابطهٔ $\phi'(\tau) = \kappa(\tau)$ نشان می‌دهد که مشتق $\phi$ نسبت به ویژه‌زمان، ویژه‌شتاب است. از این‌رو، می‌توان زاویهٔ $\phi$ را ویژه‌سرعت نامید، که گاهی پارامتر سرعت نامیده می‌شود. از بحث مثلثاتی زیر نتیجه می‌شود که اگر جهان-خط در چارچوب مرجع آزمایشگاه با $x = f(t)$ مشخص داده شود، آنگاه $f' = \tanh\phi$.

امتحان مستقیم رابطهٔ زیر با استفاده از تعریف تابع نمایی ماتریسی $e^A = \sum_0^\infty A^n/n!$ مشکل نیست

$$e^{J\phi} = \begin{pmatrix} \cos_\epsilon\phi & -\epsilon\sin_\epsilon\phi \\ \sin_\epsilon\phi & \cos_\epsilon\phi \end{pmatrix}$$

که در آن

$$\cos_\epsilon\phi = \sum_{n=0}^{\infty}\frac{(-\epsilon)^n\phi^{2n}}{(2n)!}, \qquad \sin_\epsilon\phi = \sum_{n=0}^{\infty}\frac{(-\epsilon)^n\phi^{2n+1}}{(2n+1)!}$$

به‌ازای $\epsilon = 1$، این توابع همان توابع معمولی سینوس و کسینوس هستند، به‌ازای $\epsilon = -1$، کسینوس و سینوس هیپربولیک، و به‌ازای $\epsilon = 0$ فقط $\cos_0\phi \equiv 1$ و $\sin_0\phi = \phi$ هستند. در تمامی این موارد:

$$\cos_\epsilon^2\phi + \epsilon\sin_\epsilon^2\phi = 1$$

و

$$\partial_\phi\cos_\epsilon\phi = -\epsilon\sin_\epsilon\phi, \qquad \partial_\phi\sin_\epsilon\phi = \cos_\epsilon\phi$$

با برابر قراردادن درایه‌های متناظر معادلهٔ ماتریسی:

$$e^{J(\phi+\psi)} = e^{J\phi}e^{J\psi}$$

فرمولهای مجموع به‌دست می‌آید

$$\cos_\epsilon(\phi+\psi) = \cos_\epsilon\phi\cos_\epsilon\psi - \epsilon\sin_\epsilon\phi\sin_\epsilon\psi$$

$$\sin_\epsilon(\phi+\psi) = \sin_\epsilon\phi\cos_\epsilon\psi + \cos_\epsilon\phi\sin_\epsilon\psi$$

برای ملاحظه بحثی در مثلثات این صفحه‌ها، مرجع [۲۱]، و برای مثلثات در $\mathbb{P}_{-1}$، مرجع [۱] را ببینید.

### ۳.۲ زاویه‌ها، نابرابریها، و طول قوس

دورانها صورت دوخطی (۱) را حفظ می‌کنند. یعنی، برای هر زاویهٔ $\phi$ و هر دو بردار $v$ و $w$ داریم

$$\langle e^{J\phi}v, e^{J\phi}w\rangle = \langle v, w\rangle$$

برای درک این نکته، ابتدا نشان دهید که $e^{J\phi}$ یک ماتریس متعامد است یعنی $(e^{J\phi})^{\mathrm{T}}Be^{J\phi} = B$ که در آن

$$B = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$$

ماتریس صورت دوخطی است. راهنمایی: از $(e^{J\phi})^{\mathrm{T}}Be^{J\phi}$ نسبت به $\phi$ مشتق بگیرید و توجه کنید که $J^{\mathrm{T}}B + BJ = 0$. چون وقتی بردارها به‌صورت ماتریسهای ستونی تعبیر شوند داریم $\langle v, w\rangle = v^{\mathrm{T}}Bw$، نتیجه به‌دست می‌آید.

گوییم دو بردار ناپایدار $v$ و $w$، از یک نوع‌اند اگر بردارهای یکهٔ $v_1 = v/||v||$ و $w_1 = w/||w||$ روی یک شاخه از دایرهٔ یکه باشند. یک زاویهٔ (سودار) از $v$ تا $w$ برابر عدد $\phi$ است به‌گونه‌ای که $w_1 = e^{J\phi}v_1$. در این صورت، $v_1 = e^{J\psi_1}u$ و $w_1 = e^{J\psi_2}u$ که در آن، بسته به شاخهٔ به‌خصوص دایرهٔ یکه، $u = (\pm 1, 0)$ یا $u = (0, \pm 1)$. قرار می‌دهیم $\phi = \psi_2 - \psi_1$. در این صورت $\phi$ زاویه‌ای از $v$ تا $w$ است و، با تعبیر بردارها به‌عنوان ماتریسهای تک‌ستونی داریم

$$\begin{aligned}\langle v_1, w_1\rangle = \langle e^{J\psi_1}u, e^{J\psi_2}u\rangle &= \langle u, e^{J\phi}u\rangle\\ &= u^{\mathrm{T}}Be^{J\phi}u = \pm\cos_\epsilon\phi\end{aligned}$$

که در آن علامت منفی فقط وقتی ظاهر می‌شود که $\epsilon = -1$ و $v$ و $w$ بردارهای زمان‌گونه باشند. از این‌رو، برای بردارهای از یک نوع داریم

$$\langle v, w\rangle = ||v||\,||w||\langle v_1, w_1\rangle = \pm||v||\,||w||\cos_\epsilon\phi$$

که در آن $\phi$ زاویهٔ از $v$ تا $w$ است و علامت منفی فقط وقتی ظاهر می‌شود که $\epsilon = -1$ و بردارها زمان‌گونه باشند. از این بحث مستقیماً به نابرابریهای زیر می‌رسیم.

**نابرابری شوارتس.** فرض کنید $v$ و $w$ بردارهایی از یک نوع باشند. در این صورت اگر $\epsilon = 1$ داریم $|\langle v, w\rangle| \le ||v||\,||w||$ و اگر $\epsilon = 0$ داریم $|\langle v, w\rangle| = ||v||\,||w||$، و سرانجام اگر $\epsilon = -1$ نتیجه می‌شود $|\langle v, w\rangle| \ge ||v||\,||w||$.

با جرح و تعدیلی در استدلال متعارف در حالت اقلیدسی، به سادگی می‌توان به نابرابری زیر رسید.

**نابرابری مثلثی.** فرض کنید $v$ و $w$ بردارهایی از یک نوع‌اند. در این صورت اگر ۱ $= \epsilon$ داریم $||v+w|| \leq ||v|| + ||w||$ و اگر ۰ $= \epsilon$ یا ۱ $- = \epsilon$، به ترتیب داریم $||v+w|| = ||v|| + ||w||$ و $||v+w|| \geq ||v|| + ||w||$.

با استفاده از نابرابری مثلثی می‌توان نشان داد که

> در میان خمهای منظم از نقطهٔ $p$ تا نقطهٔ $q$ در صفحهٔ اقلیدسی، خط مستقیم دارای کمترین طول قوس است، طول قوس تمامی خمهای منظم از نقطهٔ $p$ تا نقطهٔ $q$ در صفحهٔ گالیله‌ای یکسان است، در میان خمهای منظم از نقطهٔ $p$ تا نقطهٔ $q$ در صفحهٔ مینکوفسکیایی، خط مستقیم دارای بزرگترین طول قوس است.

در مکانیک نسبیتی، طول قوس یک خم زمان‌گونه از رویداد $p$ تا رویداد $q$ برابر است با زمان ثبت شده توسط ساعتی که در امتداد خم از $p$ به $q$ می‌رود. اینکه دو ساعت که در امتداد دو جهان-خط متفاوت از $p$ به $q$ می‌روند دو زمان سپری‌شدهٔ متفاوت را ثبت می‌کنند همان صورت عام پاردوکس دوقلوست که در بخش ۱.۳ توصیف شد.

## ۴.۲ قضیهٔ بنیادی خمهای مسطح

همنهشتی مستقیم دورانی است که به دنبال آن یک انتقال می‌آید:

$$\mathscr{T}(\nu) = a + e^{J\phi}\nu$$

این همنهشتیها فاصله‌های میان نقاط $\mathbb{P}_\epsilon$ را حفظ می‌کنند و یک زیرگروه از گروه طولپاییها تشکیل می‌دهند. این تبدیلها را تبدیلهای مستقیم گالیله‌ای به‌ازای ۰ $= \epsilon$ و تبدیلهای مستقیم لورنتسی به‌ازای ۱ $- = \epsilon$ می‌نامیم.

اگر $\sigma(s)$ خمی با پارامتر طول قوس $s$ و $\mathscr{T}$ یک همنهشتی مستقیم باشد، خم جدید زیر را تعریف می‌کنیم

$$\rho(s) = \mathscr{T}(\sigma(s)) = a + e^{J\phi}\sigma(s)$$

در این صورت $s$ نیز پارامتر طول قوس برای $\rho$ است و، چون $J$ و $e^{J\phi}$ تعویض‌پذیرند، تابع انحنای $\rho$ همان تابع انحنای $s$ است. اگر $\mathrm{T}(s)$ و $\mathrm{N}(s)$ بردارهای مماس و قائم یکهٔ $\sigma$ باشند، آنگاه $e^{J\phi}\mathrm{T}(s)$ و $e^{J\phi}\mathrm{N}(s)$ بردارهای مماس و قائم $\rho$ هستند.

برعکس، دو خم که با طول قوس پارامتری شده‌اند و تابع انحنای یکسان دارند، باید مستقیماً همنهشت باشند، به شرط آنکه بردارهای مماس یکهٔ آنها روی یک شاخه از دایرهٔ یکه قرار گرفته باشد. این نتیجه را قضیهٔ بنیادی خمهای مسطح می‌نامیم.

نظریهٔ بنیادی خمهای مسطح از یکتایی جوابهای معادله‌های دیفرانسیل معمولی نتیجه می‌شود. مسألهٔ مقدار آغازی را می‌توان به‌طور فشرده به‌صورت دستگاهی از دو معادلهٔ دیفرانسیل برداری با شرایط آغازی نوشت:

$$\sigma'(s) = \mathrm{T}(s), \sigma(\circ) = \sigma_\circ$$

$$\mathrm{T}'(s) = \kappa(s)J_\epsilon\mathrm{T}(s), \mathrm{T}(\circ) = \mathrm{T}_\circ$$

که در آنها $\sigma_\circ$ هر نقطه در صفحه است و $\mathrm{T}_\circ$ روی دایرهٔ یکه قرار دارد. توجه کنید که برای هر جواب داریم

$$\langle \mathrm{T}, \mathrm{T}\rangle' = ۲\kappa\langle J\mathrm{T}, \mathrm{T}\rangle = \circ$$

و طول $\mathrm{T}$ ثابت است. در نتیجه $s$ یک پارامتر طول قوس است. در حالت اقلیدسی، داریم $\mathrm{T}_\circ = (\cos\phi_\circ, \sin\phi_\circ)$؛ در مکانیک نیوتنی، $\mathrm{T}_\circ = (۱, \phi_\circ)$ و در مکانیک نسبیتی، $\mathrm{T}_\circ = (\sin\mathrm{h}\phi_\circ, \cos\mathrm{h}\phi_\circ)$.

اکنون فرض کنید که $\sigma(s)$ و $\tau(s)$ دو خم پارامتری‌شده توسط طول قوس‌اند و $\kappa(s)$، تابع انحنای آنها، یکسان است. همچنین فرض کنید $\mathrm{T}_\sigma$ و $\mathrm{T}_\tau$، به‌ترتیب، بردارهای مماس یکهٔ این دو خم‌اند. اگر $\mathrm{T}_\sigma(\circ)$ و $\mathrm{T}_\tau(\circ)$ روی یک شاخه از دایرهٔ یکه باشند آنگاه زاویهٔ $\phi$ای وجود دارد که در رابطهٔ $\mathrm{T}_\sigma(\circ) = e^{J\phi}\mathrm{T}_\tau(\circ)$ صدق می‌کند. خم جدید $\rho(s)$ را با ضابطهٔ زیر تعریف می‌کنیم

$$\rho(s) = (\sigma(\circ) - e^{J\phi}\tau(\circ)) + e^{J\phi}\tau(s)$$

چون $\rho$ تصویر $\tau$ تحت یک همنهشتی مستقیم است، $\rho$ نیز دارای تابع انحنای $\kappa(s)$ است. چون $\rho(\circ) = \sigma(\circ)$ و $\mathrm{T}_\rho(\circ) = \mathrm{T}_\sigma(\circ)$، هم $\sigma$ و هم $\rho$ جوابهای یک مسألهٔ مقدار آغازی هستند و بنابراین برابرند. این نتیجه، قضیهٔ بنیادی خمهای مسطح را اثبات می‌کند.

به‌ازای ۱ $= \epsilon$ و برای ذراتی که به‌ازای ۰ $= \epsilon$ و ۱ $- = \epsilon$ در زمان به طرف جلو حرکت می‌کنند، عبارت مربوط به $e^{J\phi}$ برحسب توابع مثلثاتی را، در صورت معلوم بودن $\kappa$، می‌توان برای بیان این خمها برحسب انتگرال به‌کار برد. این روابط، در صفحهٔ اقلیدسی، عبارت‌اند از

$$\phi(s) = \int_\circ^s k(u)du + \phi_\circ, x(s) = \int_\circ^s \cos\phi(u)du + x_\circ, y(s) = \int_\circ^s \sin\phi(u)du + y_\circ$$

برای یک ذرهٔ نسبیتی ($\tau$، ویژه‌زمان) عبارت‌اند از

$$\phi(\tau) = \int_\circ^\tau k(u)du + \phi_\circ, x(\tau) = \int_\circ^\tau \sinh\phi(u)\,du + x_\circ, t(\tau) = \int_\circ^\tau \cosh\phi(u)du + t_\circ$$

و برای یک ذرهٔ نیوتنی عبارت‌اند از

$$x(t) = \int_\circ^t \phi(u)du + x_\circ \quad \text{و} \quad \phi(t) = \int_\circ^t k(u)du + \phi_\circ$$

این فرمولها هنگامی که انتگرال‌گیری انجام‌پذیر باشد مفیدند. دستگاه اصلی معادلات دیفرانسیل برای محاسبات عددی ارجح است زیرا در این صورت نیاز به استفاده از توابع متعالی وجود ندارد. برای آشنایی با یک تابع متمتیکا که جوابها را به دستگاه معادلات دیفرانسیل برمی‌گرداند مرجع [۹] را ببینید.

## ۳. نقاط شتابدار

مطالعهٔ ذرهٔ نقطه‌ای نسبیتی که در یک بعد در حرکت است فقط مستلزم تعریف طول قوس است، یعنی زمانی که توسط ساعتِ روی ذره ثبت می‌شود، و دستگاه فرنهٔ معادلات دیفرانسیل. در اینجا اثر عمدهٔ اتساع زمان است که براساس آن ساعتهای متحرک در مقایسه با ساعتهای ساکن واقع در چارچوب آزمایشگاه کُند کار می‌کنند. برای توضیح این مطلب دو مثال می‌آوریم.

### ۱.۳ پارادُکس ساعت

پارادُکس ساعت، یا پارادُکس دوقلو، شامل یک زوج دوقلوی یکسان است که یکی به مسافرت میان ستاره‌ای می‌رود و دیگری در منزل می‌ماند. وقتی دوقلوی مسافر برمی‌گردد، سن آنها دیگر یکسان نیست. سن دوقلوی مانده در منزل از سن دوقلوی مسافر بیشتر است. این مسأله را در درس حسابان هنگام بحث طول قوس می‌توان حل کرد. فرمول طول قوس و فرمول ویژه‌زمان فقط در یک علامت با هم فرق می‌کنند. فرمول طول قوس عبارت است از

$$s = \int_a^b \sqrt{۱ + (dx/dt)^۲}\,dt \qquad (۲)$$

و فرمول ویژه‌زمان چنین است

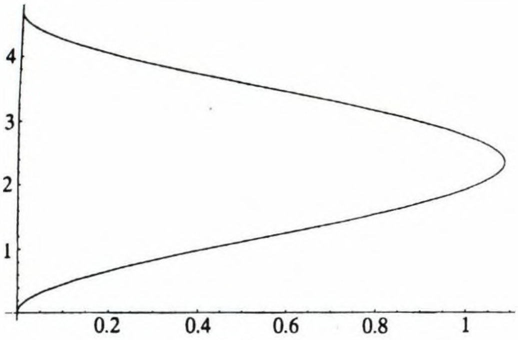
$$\tau = \int_a^b \sqrt{۱ - (dx/dt)^۲}\,dt \qquad (۳)$$

برای اجتناب از یکاهای هندسی، به جای $\nu = dx/dt$ قرار می‌دهیم $\nu/c$ که در آن $c$ سرعت نور در دستگاه یکاهای به‌کار رفته است.

تحلیل این موضوع، در صورت استفاده از تبدیلات لورنتس، نادقیق است زیرا هنگام برگشتن دوقلوی مسافر شتاب به‌طور اجتناب‌ناپذیری تغییر می‌کند. اما، با استفاده از روابط (۲) و (۳) می‌توان صرفاً مسیر فضا‌ـزمان، یعنی رابطهٔ $\sigma(t) = (x(t), t)$، را برای مسافر انتخاب و طول قوس مینکوفسکیایی را محاسبه کرد. هر مسیر انتخاب‌شده‌ای این اثر را نشان می‌دهد. با استفاده از روشهای عددی می‌توان بر مشکلات فنی انتگرال‌گیری غلبه کرد.

اگر دستگاه معادلات دیفرانسیل فرنه را به‌کار ببریم می‌توانیم شتاب را مشخص کنیم. فرض کنید خلبان سفینهٔ دوقلوی مسافر موشکها را به‌گونه‌ای روشن می‌کند که شتاب ثابتی (آن‌طور که در سفینه اندازه‌گیری می‌شود) برای مدت یک‌سال به سفینه بدهد. بعد از یک سال، سفینه ۱۸۰ درجه می‌چرخد و برای مدت یک‌سال بعد با شتاب کندشونده پیش می‌رود تا متوقف شود. روشن کردن موشکها برای سال بعد سفینه را به نیمه راه زمین برمی‌گرداند. در این نقطه دوباره سفینه ۱۸۰ درجه می‌چرخد و سرانجام به زمین برمی‌گردد. جهان‌ـخط در $\mathbb{P}_{-۱}$ در شکل ۲ رسم شده است. زمین در $x = ۰$ واقع است و مسافرت در $t = ۰$ شروع می‌شود. شتاب برابر است با ۱ سال نوری بر سال بر سال، که تقریباً برابر شتاب حاصل از گرانش بر زمین است. مسافرت، به زمان سفینه، چهار سال طول می‌کشد. جهان‌ـخطِ دوقلوی مسافر در شکل ۲ نشان داده شده است. جهان‌ـخط دوقلوی مانده در منزل، محور قائم است. برای ملاحظهٔ جزئیات محاسبهٔ عددی مرجع [۹] را ببینید.

چون $\sigma(۴) \approx (۰, ۴{,}۷)$، سن دوقلوی مانده در منزل حدود ۴٫۷ سال و سن دوقلوی مسافر ۴ سال است. چون $\kappa$ در این مثال ثابت تکه‌ای است،



شکل۲

جهان‌ـخط متشکل است از قوسهای هذلولوی و محاسبهٔ دقیق امکان‌پذیر است. (بخش ۲.۳ را ببینید).

### ۲.۳ تکانه، انرژی و نیرو

**قانون پایستگی.** بردار مماس یکه، در محاسبات روزمرهٔ فیزیک، مانند تحلیل اثر برهمکنشهای ذره در شتاب‌دهنده، نقش مهمی بازی می‌کند. این امر بدین دلیل است که به‌جای قوانین پایستگی کلاسیک تکانه، انرژی و جرم، یک قانون واحد برحسب بردارهای مماس بر جهان‌ـخطها به‌کار می‌رود.

به هر «ذره» عددی به‌نام جرم سکون وابسته است. در این مقاله ذراتی را که جرم سکون مثبت دارند بررسی می‌کنیم. برای تعمیم این بحث به ذرات نورگونه، مرجعهای [۱۱]، [۱۶] و [۱۷] را ببینید. اگر ذره‌ای با جرم سکون $m_۰$ دارای جهان‌ـخط $\sigma$، بردار مماس یکهٔ $\mathrm{T}$، و بردار قائم یکهٔ $\mathrm{N}$ باشد، بردار تکانهٔ نسبیتی یا چهار‌ـتکانهٔ ذره را به‌صورت بردار $P = m_۰\mathrm{T}$ تعریف می‌کنیم. اکنون فرض کنید در لحظهٔ $t = t_۰$ در چارچوب آزمایشگاه، ذراتی با بردارهای تکانهٔ نسبیتی $P_۱$، $P_۲$، ... $P_n$ برهمکنش می‌کنند. بعد از برهمکنش، ذرات جدیدی با بردارهای تکانهٔ نسبیتی $Q_۱$، $Q_۲$، ... ، $Q_m$ به‌جای می‌ماند. قانون پایستگی عبارت است از

$$P_۱ + P_۲ + \cdots + P_n = Q_۱ + Q_۲ + \cdots + Q_m$$

فرض کنید که، در چارچوب آزمایشگاه، جهان‌ـخطِ ذره‌ای با جرم سکون $m_۰$ عبارت است از $\sigma(t) = (x(t), t)$ که در آن $t$ از روی ساعت آزمایشگاه اندازه‌گیری می‌شود. بردار تکانهٔ نسبیتی برابر است با

$$m_۰(d\sigma/d\tau) = m_۰(dt/d\tau)(d\sigma/dt) = \left(\frac{m_۰\nu}{\sqrt{۱-\nu^۲}}, \frac{m_۰}{\sqrt{۱-\nu^۲}}\right)$$

مؤلفهٔ اول تکانهٔ آزمایشگاه یا سه‌ـتکانه و مؤلفهٔ دوم انرژی ذره براساس اندازه‌گیری در آزمایشگاه است. جرم سکون ذره، کمینهٔ انرژی آن است و هنگامی که ذره در چارچوب مرجع آزمایشگاه ساکن است انرژی آن کمینه است. در دستگاه یکاهایی که در آنها سرعت نور واحد نیست این برابری جرم و انرژی با رابطهٔ $E = m_۰c^۲$ داده می‌شود که در آن ضریب تبدیل $c$ سرعت نور است.

برای توضیح این قانون پایستگی فرض کنید که در چارچوب مرجع آزمایشگاه دو تکه بتونهٔ مساوی به جرم $m_۰$ در حال سکون وجود دارد. این دو تکه را با شتاب به طرف همدیگر می‌فرستیم و این دو در زمان $t_۰$ با سرعتهای مساوی و مخالف $\pm\nu_۰$ با هم برخورد می‌کنند. فرض کنید که آنها

بر اثر برخورد به هم می‌چسبند و یک تکهٔ جدید تشکیل می‌دهند. همچنین فرض کنید که $P_1$ و $P_2$ بردارهای تکانهٔ تکه‌های اصلی در زمان $t_0$، و $Q$ بردار تکانهٔ تکهٔ جدید باشد. در این صورت

$$Q = P_1 + P_2$$
$$= \left(\frac{m_0\nu_0}{\sqrt{1-\nu_0^2}}, \frac{m_0}{\sqrt{1-\nu_0^2}}\right) + \left(\frac{-m_0\nu_0}{\sqrt{1-\nu_0^2}}, \frac{m_0}{\sqrt{1-\nu_0^2}}\right)$$
$$= \left(0, \frac{2m_0}{\sqrt{1-\nu_0^2}}\right)$$

از این‌رو تکهٔ جدید بتونه در چارچوب مرجع آزمایشگاه ساکن و جرم سکون آن بزرگتر از مجموع جرم سکون‌های تکه‌های اولیه است. جرم سکون تکهٔ جدید برابر است با

$$\frac{2m}{\sqrt{1-\nu_0^2}} = 2m_0 + 2\left(\frac{1}{2}m_0\nu_0^2\right) + \frac{3}{4}m_0\nu_0^4 + \cdots.$$

بنابراین، با چشم‌پوشی از جمله‌های با توان بیشتر از دو، افزایش جرم سکون برابر است با انرژی جنبشی کلاسیک سیستم قبل از برخورد. انرژی به جرم تبدیل شده است.

براساس این محاسبات، می‌توان یک متمم کاربردی برای درس جبر خطی مقدماتی عرضه کرد که در آن از جبر برداری چهاربُعدی استفاده شود.

**نیرو.** سه‌ـنیرو مشتق سه‌ـتکانه نسبت به زمان آزمایشگاه است. با فرض ثابت بودن جرم سکون، نوشتن تکانهٔ نسبیتی به‌صورت $P = (p, E)$، و به‌کار بردن فرمولهای فرنه داریم

$$\left(\frac{dp}{dt}, \frac{dE}{dt}\right) = \frac{dP}{dt} = \frac{d\tau}{dt}\frac{dP}{d\tau} = m_0\kappa J\frac{d\mathrm{T}}{dt} = \left(m_0\kappa, m_0\kappa\frac{dx}{dt}\right)$$

از این‌رو، با فرض ثابت بودن جرم سکون و حرکت راستخط، سه‌ـنیروی وارد بر یک ذره برابر است با

$$p' = m_0\kappa = m_0\frac{x''}{(1-(x')^2)^{3/2}}$$

این رابطه جانشینِ نسبیتی قانون $F = ma$ نیوتن است. براساس این معادله می‌توان یک مجموعه مسألهٔ کاربردی برای یک درس معادلات دیفرانسیل مقدماتی مطرح کرد. با فرض ثابت بودن نیروی $F$ داریم

$$\frac{x''}{(1-(x')^2)^{3/2}} = \frac{F}{m_0} = \alpha$$

با قرار دادن $\nu = x'$ یک معادلهٔ تفکیک‌پذیر با جواب زیر به‌دست می‌آید

$$\nu(t) = \frac{\alpha t + a}{\sqrt{(\alpha t + a)^2 + 1}} \qquad (4)$$

که در آن

$$a = \frac{\nu_0}{\sqrt{1-\nu_0^2}}, \qquad \nu_0 = \nu(0)$$

از این رو با بزرگ شدن $t$ سرعت به سرعت نور نزدیک می‌شود و می‌توان به‌ازای یک $\alpha$ی مشخص زمان آزمایشگاه را برای شتاب دادن به ذره، مثلاً، تا ۹۹٪ سرعت نور به‌دست آورد. زمان ثبت‌شده توسط ساعت همراه ذره را می‌توان با استفاده از انتگرال زیر محاسبه کرد

$$\int \sqrt{1-\nu^2}\,dt = \frac{1}{\alpha}\sinh^{-1}(\alpha t + a) + C$$

با انتگرال گرفتن از $\nu(t)$ به‌دست می‌آید

$$x(t) = \alpha^{-1}\sqrt{(\alpha t + a)^2 + 1} + b \qquad (5)$$

که در آن

$$b = x_0 - \frac{1}{\alpha\sqrt{1-\nu_0^2}}, \qquad x_0 = x(0)$$

این یک هذلولی است و حرکت با شتاب ثابت را غالباً حرکت هذلولوی می‌گویند. با فرض $\nu_0 = x_0 = 0$ برای سهولت، حرکت عبارت است از

$$x(t) = \frac{1}{\alpha}\left(\sqrt{\alpha^2 t^2 + 1} - 1\right) = \frac{\alpha}{2}t^2 - \frac{\alpha^3}{8}t^4 + \cdots$$

که نتیجهٔ نیوتنی $x(t) = \alpha t^2/2$ به‌اضافهٔ جمله‌های مرتبهٔ بالاتر است.

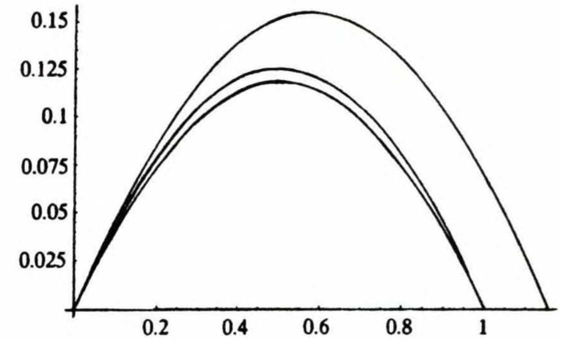
**آلبرت و ایزاک توپ‌بازی می‌کنند.** آلبرت اینشتین و ایزاک نیوتن، شانه به شانهٔ همدیگر ایستاده‌اند و توپهای یکسانی را با سرعت و ارتفاع اولیهٔ یکسان مستقیماً به طرف بالا پرتاب می‌کنند. توپ کدام‌یک بیشتر بالا می‌رود؟ توپ کدام‌یک مدت طولانیتری در هوا می‌ماند؟

با فرض اینکه $x(t)$ ارتفاع باشد و $x(0) = 0$، فرمولهای (۴) و (۵) نشان می‌دهند که توپ آلبرت در مدت زمان زیر به زمین برمی‌گردد ($\alpha$ در اینجا منفی است)

$$t = \frac{-2\nu_0}{\alpha\sqrt{1-\nu_0^2}}$$

که در آن $\nu_0$ سرعت اولیه است. چون توپ ایزاک در زمان $t = -2\nu_0/\alpha$ به زمین برمی‌گردد، توپ آلبرت مدت زمان بیشتری در هوا باقی می‌ماند و مسافت بیشتری هم بالا می‌رود. دو خم بالایی در شکل ۳ نمودارهای ارتفاع دو توپ برحسب زمان هستند.

اما برای اینکه توپ آلبرت با همان سرعت اولیهٔ توپ نیوتن بالا برود باید انرژی بیشتری به آن بدهد. فرض کنید آلبرت و ایزاک توپهای خود را با انرژی یکسانی پرتاب می‌کنند. یعنی، آنها توپهای خود را با نیروی ثابتِ مساوی و




**شکل۳** جهان‌ـخط توپهای آلبرت و ایزاک

برای زمانهای مساوی شتاب می‌دهند. با فرض $\nu_0 = 0$، سرعت توپهای آلبرت و ایزاک، به ترتیب، برابر است با

$$\nu_A(t) = \frac{\alpha t}{\sqrt{\alpha^2 t^2 + 1}}, \qquad \nu_1(t) = \alpha t$$

از این رو، به‌ازای نیروی یکسان، اگر توپ ایزاک با سرعت اولیهٔ $\nu_1$ رها شود آنگاه توپ آلبرت با سرعت اولیهٔ زیر رها می‌شود

$$\nu_A = \frac{\nu_1}{\sqrt{1+\nu_1^2}}$$

و در مدت زمان

$$t = \frac{-2\nu_A}{\alpha\sqrt{1-\nu_A^2}} = -2\nu_1/\alpha$$

به زمین برمی‌گردد. دو توپ مدت زمان یکسانی در هوا باقی می‌مانند. توپ ایزاک تا ارتفاع بیشتری بالا می‌رود.

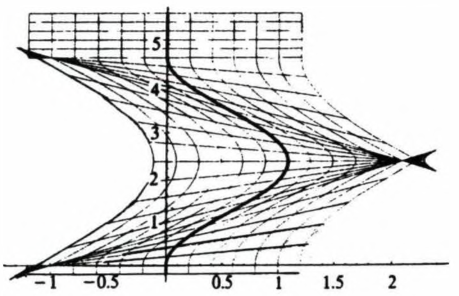
مسیرهای فضا‌ـزمان دو توپ، با فرض جرم واحد در شکل ۳ رسم شده است. نیرو، در میدان گرانش زمین، ۱ سال نوری بر سال بر سال تقریب زده شده است. سرعت اولیه، به خاطر رسم بهتر نمودار، نسبتاً بزرگ و برابر $\nu_1 = 0.5$ فرض شده است. نمودار وسط مربوط به توپ ایزاک، نمودار بالا مربوط به توپ آلبرت با سرعت اولیهٔ $\nu_A = 0.5$ و نمودار پایین مربوط به توپ آلبرت با سرعت اولیهٔ $\nu_A = \nu_1/\sqrt{1+\nu_1^2} = 1/\sqrt{5}$ است.

مسیرهای فضا‌ـزمان توپهای آلبرت و ایزاک، به ترتیب، خمهایی با انحنای ثابت در $\mathbb{P}_{-1}$ و $\mathbb{P}_0$ هستند. در مرجعهای [۲۱] و [۲۲] خمهای با انحنای ثابت در $\mathbb{P}_\epsilon$ را چرخه می‌گویند. به‌ازای $\epsilon = 1, -1$، چرخه‌ها (شاخه‌های) دایره‌ها و خطهای راست و به‌ازای $\epsilon = 0$، چرخه‌ها سهمیها و خطهای مستقیم (ناقائم) هستند.

## ۴. قطعه‌های شتابدار

انقباض طول مخصوصِ حرکت جسم گسترده است نه نقطه‌ای. در این موقعیت، برای تعریف دستگاه مختصات یک ناظر متحرک با مسیر فضا‌ـزمان $\sigma$، دستگاه کامل فرنه را به‌کار می‌بریم.

فرض کنید $\sigma(\tau)$ یک خم زمان‌گونه در $\mathbb{P}_{-1}$ باشد که نسبت به ویژه‌زمان پارامتری شده است و دارای چارچوب مرجع فرنه $\{\mathrm{T}(\tau), \mathrm{N}(\tau)\}$ است.




**شکل۴** شبکهٔ مختصات دوقلوهای مسافر

بنا به تعریف، مختصات $(\sigma,\tau)$ در چارچوب مرجع فرنه برای $\sigma$، به صورت زیر با نقاط در چارچوب $\mathbb{P}_{-1}$ مربوط می‌شوند

$$(x,t) = \sum(\sigma,\tau) = \sigma(\tau) + \sigma\mathrm{N}(\tau) \qquad (6)$$

مفهوم فیزیکی چارچوب مرجع فرنه عبارت است از: مجموعهٔ رویدادهای همزمان با ویژه‌زمان $\tau$ برای ناظر شتابدار با جهان‌ـخط $\sigma$ متشکل است از خط عمود بر بردار مماسِ یکهٔ $\mathrm{T}(\tau)$، یعنی در راستای $\mathrm{N}(\tau)$ قرار دارد، مقدار $\sigma$ فاصله‌ای است که ناظر شتاب‌دار اندازه می‌گیرد.

اگر سرعت، $\nu$، در چارچوب مرجع آزمایشگاه ثابت باشد، آنگاه با شرایط اولیهٔ مناسب، $\sigma(t) = (\nu t, t)$، $d\tau/dt = \sqrt{1-\nu^2}$ مقدار ثابت است و می‌توان نوشت $\tau = t\sqrt{1-\nu^2}$. در این صورت فرمول (۶) به صورت زیر در می‌آید،

$$x = (\nu\tau + \delta)/\sqrt{1-\nu^2}, \qquad t = (\nu\delta + \tau)/\sqrt{1-\nu^2}$$

که همان تبدیل لورنتس است.

شکل ۴، شبکهٔ مختصات $\sigma\tau$ را برای دوقلوی شتابدار بخش ۱.۳ نشان می‌دهد. جهان‌ـخط $\sigma$ پررنگ است.

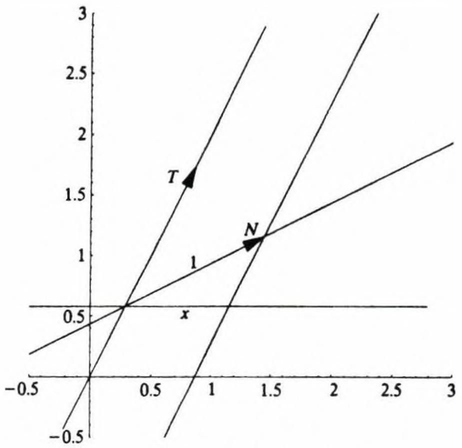
توجه کنید که این شبکه تکینگی دارد و آن در نقاطی است که همزمان با دو یا چند مقدار متمایز از $\tau$ هستند. هرگاه بردار $\mathrm{N}(\tau)$ ثابت نباشد وجود این تکینگیها اجتناب‌ناپذیر است؛ فرمولهای فرنه با $\epsilon = 0$ نشان می‌دهند که $\mathrm{N}$ در حالت نیوتنی ثابت است. برای اجتناب از این نقاط تکینه، توجه خود را به ناحیه‌هایی محدود می‌کنیم که در آنها خمهای موازی $\sum(\delta,\tau)$، با $\delta$ی ثابت، منظم‌اند. با استفاده از فرمولهای فرنه با $\epsilon = -1$ داریم

$$\frac{d}{d\tau}(\sigma(\tau)+\delta\mathrm{N}(\tau)) = \mathrm{T}(\tau)+\delta\kappa(\tau)\mathrm{T}(\tau) = (1+\delta\kappa(\tau))\mathrm{T}(\tau)$$

دستگاه مختصات یک ناظر شتابدار در ناحیه‌ای تعریف می‌شود که در آن وقتی $\kappa(\tau) > 0$، داریم $\delta > -1/\kappa(\tau)$ و وقتی $\kappa(\tau) < 0$ داریم $\delta \ll -1/\kappa(\tau)$.

ماتریس صورت دوخطی مینکوفسکی برحسب مختصات $\delta\tau$ عبارت است از

$$\begin{pmatrix} \langle\partial_\delta \sum, \partial_\sigma \sum\rangle & \langle\partial_\delta \sum, \partial_\tau \sum\rangle \\ \langle\partial_\tau \sum, \partial_\delta \sum\rangle & \langle\partial_\tau \sum, \partial_\tau \sum\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -(1+\delta\kappa(\tau))^2 \end{pmatrix}$$




**شکل۵** انقباض طول برای ناظرهای لخت

برای حالت اقلیدسی، مرجع [۸، صفحهٔ ۱۱۵] در مورد خمهای موازی را ببینید.

شبکهٔ مختصات ناظر متحرک نمونهٔ جالبی از دستگاه مختصات خمیده خطی است که می‌تواند در درس حسابان چندمتغیره به دانشجویان آموخته شود. یک تابع متمتیکا برای رسم شبکهٔ مختصات ناظران شتابدار را می‌توان در مرجع [۹] یافت.

## ۱.۴ انقباض طول

جهان‌ـخط یک ناظر لخت ($\kappa(\tau) \equiv 0$) که با سرعت $\nu$ در $\mathbb{P}_{-1}$ حرکت می‌کند عبارت است از $\sigma(\tau) = a + \tau\mathrm{T}$ که در آن $\mathrm{T} = (\sinh(\phi), \cosh(\phi))$ بردار مماس یکه است و $\tanh(\phi) = \nu$. بردار قائم یکه عبارت است از $\mathrm{N} = (\cosh(\phi), \sinh(\phi))$. طول «میله»ای که این ناظر لخت با خود حمل می‌کند و از $\sigma(\tau)$ تا $\sigma(\tau) + l\mathrm{N}$ گسترده است در چارچوب مرجع ناظر برابر $l$ است. طول آن در چارچوب مرجع آزمایشگاه چقدر است؟

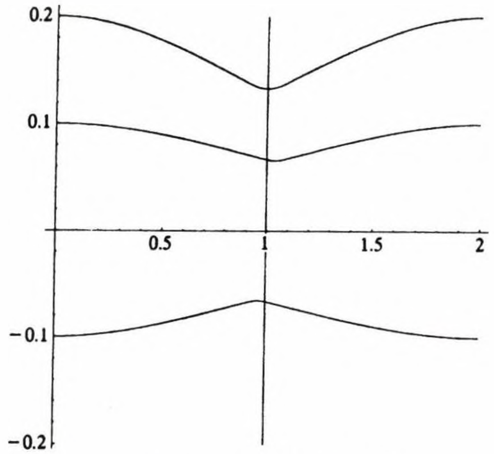
در شکل ۵ طول میله، $x$، در زمان $t$ در چارچوب مرجع آزمایشگاه برابر طول خط افقی‌ای است که از $t$، واقع بر محور دوم، می‌گذرد یعنی توسط جهان‌ـخطهای دو انتهای میله قطع می‌شود. چون $\mathrm{T}$ بر $\mathrm{N}$ عمود است، $l$ برابر است با طول تصویر عمودی بردار $(x, 0)$ بر امتداد $\mathrm{N}$. با استفاده از فرمول تصویر عمودی داریم

$$\begin{aligned} l = \frac{\langle (x,0), \mathrm{N}\rangle}{\langle \mathrm{N}, \mathrm{N}\rangle} &= \|(x,0)\|\,\|\mathrm{N}\| \cosh(\phi) \\ &= x\cosh(\tanh^{-1}(\nu)) = \frac{x}{\sqrt{1-\nu^2}} \end{aligned}$$

از این رو، $x = l(\sqrt{1-\nu^2})$ کوتاهتر از $l$ است.

اگر $\kappa(\tau)$ ناصفر باشد، جهان‌ـخطهای موازی خمیده هستند و هیچ فرمول ساده‌ای برای آنها وجود ندارد. با وجود این، این مسأله با روشهای عددی حل می‌شود. برای ملاحظهٔ جزئیات محاسبه، مرجع [۹] را ببینید.

فرض کنید دوقلوی مسافر پارادُکس ساعت (بخش ۱.۳) میله‌ای را با خود حمل می‌کند که طول آن به نظر مسافر ثابت است. طول این میله به نظر دوقلوی مانده در منزل چگونه تغییر می‌کند؟




**شکل۶** میله‌ای که دوقلوی مسافر با خود می‌برد.

در شکل ۶ نمودارهایی برای طول میله، آن‌طورکه توسط دوقلوی «ساکن» در چارچوب مرجع آزمایشگاه اندازه‌گیری می‌شود، رسم شده که در دستگاه مختصات دوقلوی مسافر از $\delta = -0.1$ تا $\delta = 0.1$ گسترده است. این نمودارها مربوط به قسمت رفتِ مسافرت‌اند؛ نمودارهای قسمت برگشت با این نمودارها متقارن‌اند.

نمودار بالا طول میله است. نمودار پایین، فاصلهٔ (علامت‌دار) از دوقلوی مسافر تا انتهای پسین میله و نمودار وسط، فاصله از دوقلو تا انتهای پیشین میله است. توجه کنید که وقتی سرعت دوقلوی مسافر بیشینه است ($\tau = 1$) کل طول میله کمینه است، اما طول نیمهٔ پسین اندکی قبل و طول نیمهٔ پیشین اندکی بعد کمینه می‌شود. یعنی، از دید دوقلوی ساکن، طول نیمه‌های پیشین و پسین میله به‌طور متفاوت منقبض می‌شود.

جزئیات انقباض نه تنها به سرعت لحظه‌ای چارچوب شتاب‌دار بلکه به مکان میله در آن چارچوب نیز وابسته است.

آیا در این مثال چیز خاصی در بارهٔ «مرکز» میله وجود دارد؟ مرکز میله جایی است که ما چارچوب مرجع مسافر را در آن قرار داده‌ایم. در مورد چارچوبهای وابسته به نقاط مختلف میله چه می‌توان گفت؟ این سؤال را در بخش بعد پاسخ می‌دهیم.

## ۲.۴ دو ناظر ساکن در یک چارچوب مرجع شتاب‌دار

فرض کنید $\sigma(\tau)$ جهان‌ـخط یک ناظر با چارچوب مرجع فرنه $\{\mathrm{T}(\tau), \mathrm{N}(\tau)\}$ است که در آن $\tau$ ویژه‌زمان، و ناظر دوم نسبت به ناظر اول ساکن است و در فاصلهٔ (علامت‌دار) $\delta$ از او قرار دارد. جهان‌ـخط ناظر دوم را می‌توان به صورت زیر پارامتری کرد

$$\rho(\tau) = \sigma(\tau) + \delta\mathrm{N}(\tau)$$

این نمایش پارامتری ویژه‌زمان نیست. اگر فرض کنیم $s$ پارامتر ویژه‌زمان برای $\rho$ باشد آنگاه

$$\frac{ds}{d\tau} = \left\|\frac{d\rho}{d\tau}\right\| = \|\mathrm{T}(\tau) + \delta\kappa(\tau)\mathrm{T}(\tau)\| = 1 + \delta\kappa(\tau)$$

وقتی شتاب وجود دارد، ساعتهای «سربالا» تندتر کار می‌کنند.

توجه کنید که وقتی $\delta$ به تکینگی در $-1/\kappa(\tau)$ نزدیک می‌شود ساعت ناظر دوم، به نظر ناظر اول، کُند و سرانجام متوقف می‌شود.

بنا به اصل نظریهٔ نسبیت عام اینشتین، ناظر در یک سفینهٔ بدون پنجره در فضا نمی‌تواند میان شتاب ناشی از موتورهای سفینهٔ خود و شتاب حاصل از میدان گرانشی اجرام نزدیک تمایز قائل شود. از این اصل نتیجه می‌شود که ساعتهای «سربالا» در میدان گرانشی تندتر کار می‌کنند. درستی این گفته به تجربه ثابت شده است.

توجه کنید که اگرچه $\tau$ ویژه‌زمان $\rho$ نیست، چارچوب مرجع فرنه در $\rho(\tau)$ عبارت است از $\{\mathrm{T}(\tau), \mathrm{N}(\tau)\}$، زیرا اینها بردارهای یکه در جهت درست هستند. در نتیجه خمهای مختصات فضاگونه برای هر دو ناظر یکسان هستند و خمهای مختصات زمان‌گونهٔ دو ناظر فقط در پارامتری‌سازی با هم متفاوت‌اند. به نظر ناظر لخت، اثرهای انقباض میله، اگر با $\sigma$ یا با $\rho$ حمل شوند، یکسان‌اند. اما، تابعهای انحنا متفاوت‌اند. درواقع داریم

$$\frac{d\mathrm{T}(\tau(s))}{ds} = \frac{d\mathrm{T}}{d\tau}\frac{d\tau}{ds} = \frac{\kappa(\tau)}{1+\delta\kappa(t)}\mathrm{N}(\tau)$$

به‌خصوص، ویژه‌شتاب انتهای پیشین میله که توسط یک ناظر شتاب‌دار حمل می‌شود، از ویژه‌شتاب انتهای پسین کمتر است. درواقع، ویژه‌شتاب در $\rho(\tau)$، همچنان که $\delta$ در $-1/\kappa(\tau)$ به تکینگی نزدیک می‌شود، به بینهایت گرایش پیدا می‌کند. این بدان معنی است که لبهٔ پیشین میله که تحت ویژه‌شتابی به بزرگی $g$ قرار دارد نمی‌تواند ویژه‌طولی بزرگتر از $g^{-1}$ داشته باشد، زیرا در این صورت می‌شکند. برای توضیح بیشتری در این مورد مرجع [۱۹] و مرجع [۲۰، مسألهٔ L.۱۴] را ببینید.

### ۳.۴ سرعت نور در یک چارچوب مرجع شتاب‌دار

یک اصل اساسی نظریهٔ نسبیت خاص این است که سرعت نور در همهٔ چارچوبهای لخت برابر است. این اصل در مورد ناظرهای شتاب‌دار معتبر نیست. چارچوب مرجع فرنه تصویر روشنی از این پدیده به‌دست می‌دهد.

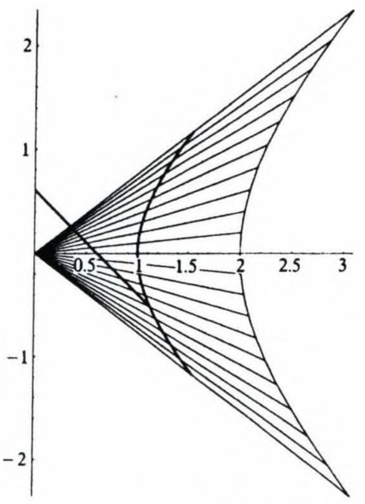
فرض کنید $\sigma(\tau)$ یک ناظر شتاب‌دار با چارچوب فرنه $\{\mathrm{T}(\tau), \mathrm{N}(\tau)\}$ باشد که در آن $\tau$ ویژه‌زمان است. مسیر مختصات ناظر شتاب‌دار با رابطهٔ $\sum(\delta,\tau) = \sigma(\tau)+\delta\mathrm{N}(\tau)$ داده می‌شود. یک ذرهٔ نورگونه در فضای $\delta\tau$ چه مسیری را دنبال می‌کند؟ فرض کنید رابطهٔ $\rho(u) = (\delta(u)), \tau(u))$ نمایش پارامتری جهان‌ـخطِ یک ذرهٔ نورگونه است. در این صورت، با استفاده از فرمول صورت دوخطی در مختصات $\delta\tau$ که در بخش ۴ داده شد داریم

$$\left(\frac{d\delta}{du}\right)^2 - (1+\delta\kappa(\tau))^2\left(\frac{d\tau}{du}\right)^2 = 0$$

سرعت ذره برابر است با

$$\frac{d\delta}{d\tau} = \pm(1+\delta\kappa(\tau))$$

که در آن علامت به جهت بستگی دارد. اگر $\kappa(\tau) \neq 0$، آنگاه سرعت، اگر و فقط اگر $\delta = 0$، برابر ۱ است. با گرایش $\delta$ به تکینگی در $-1/\kappa(\tau)$، سرعت نور به صفر نزدیک می‌شود.



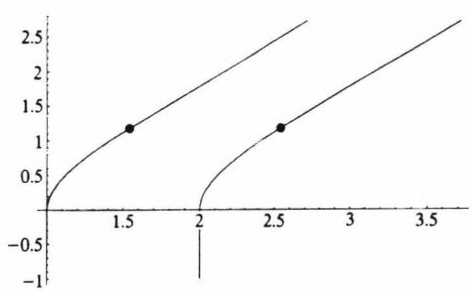

**شکل ۷** سرعت نور برای یک ناظر شتاب‌دار

چون سرعت نور وقتی $\delta = 0$ برابر ۱ است، ممکن است گفته شود که سرعت نور در «نزدیک» ناظر شتاب‌دار ثابت است. شکل ۷ نشان می‌دهد که برای ناظر شتاب‌داری که در چارچوب مرجع آزمایشگاه با سرعت نور در حرکت است چگونه فوتونها می‌توانند کند و متوقف شوند. خم پررنگ، جهان‌ـخطِ ناظری است که از $\tau = -1$ تا $\tau = 1$ شتابش ثابت و برابر ۱ است (حرکت هذلولوی). این جهان‌ـخط قطعه‌ای از دایرهٔ یکه در $\mathbb{P}_{-1}$ است. تکینگی این ناظر در مبدأ چارچوب مرجع آزمایشگاه واقع است. دستگاه مختصات ناظر شتاب‌دار تمامی $\mathbb{P}_{-1}$ را نمی‌پوشاند، بلکه فقط ربع خارج از مخروط نوری به طرف سمت راست مبدأ را می‌پوشاند. پرتوهای آمده از مبدأ، محورهای فضایی لحظه‌ای ناظر شتاب‌دار هستند. این پرتوها توسط بازه‌های مساوی ویژه‌زمان ناظر شتاب‌دار از هم جدا شده‌اند.

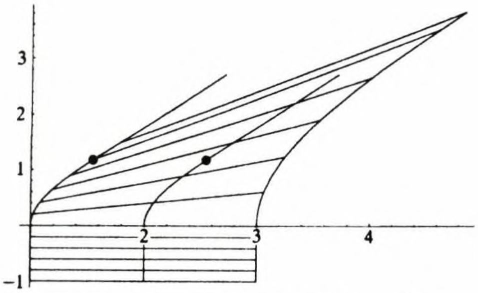
خط مستقیم پررنگ، جهان‌ـخطِ فوتون گسیل شده در $\sigma(-0{,}5)$ است. این فوتون در چارچوب مرجع آزمایشگاه در مدت زمان متناهی به $x = 0$ نزدیک می‌شود. برای ناظر شتاب‌دار، جهان‌ـخط فوتون باید بینهایت از خطهایی را که فاصلهٔ آنها به طور موقت برابر است قطع کند و بنابراین هرگز به تکینگی نمی‌رسد. چون فاصلهٔ تکینگی از ناظر شتاب‌دار متناهی است، سرعت باید کند شود.

### ۴.۴ موشکهای به‌هم بسته شده و قطار روی خط آهن

قطارِ روی یک خط دایره‌ای را در نظر بگیرید. ما این وضعیت را یک چندضلعی در نظر می‌گیریم که رأسهای آن روی یک دایرهٔ ثابت قرار دارند. ابتدا، با اِعمال شتاب یکسان به هر رأس، چندضلعی را از حالت سکون به دوران وامی‌داریم. براساس استدلال نوروود، اضلاع چندضلعی که (تقریباً) در جهت حرکت هستند، باید در چارچوب مرجع آزمایشگاه منقبض شوند. شعاع چندضلعی، که بر راستای حرکت عمود است، منقبض نمی‌شود. از این رو، نسبت محیط به قطر (در بحث نوروود، $\pi$) تغییر می‌کند. توجه کنید که ما به همهٔ رأسها شتاب یکسان می‌دهیم و بنابراین، در چارچوب مرجع آزمایشگاه، فاصلهٔ میان آنها نباید کاهش یابد که متناقض با اصل انقباض طول است. نوع راستخط این معما را پارادُکس موشکهای به هم بسته می‌گویند.




**شکل۸** جهان‌ـخطهای موشکهای به‌هم بسته شده




**شکل۹** دستگاه مختصات دوقلوی شتاب‌دار سمت چپ

دو سفینهٔ یکسان در $x = ۱$ و $x = ۲$ ساکن هستند. سپس موتور موشکها برای مدت زمان مشخصی روشن می‌شود و بعد از خاموش شدن موتورها، موشکها شتاب و سرعت نهایی یکسان می‌گیرند. به‌عبارت دقیقتر، فرض کنید دو سفینه، که در این مثال آنها را بدون بعد فرض می‌کنیم، تابع شتاب مشترکی دارند که در $۰ \leq \tau \leq ۱$ همواره برابر ۱ و در غیر این صورت برابر صفر است. جهان‌ـخطهای حاصل در شکل ۸ رسم شده است. در نقطه‌های روی نمودار، موتورها خاموش‌اند و سفینه‌ها به حالت خلاص پیش می‌روند.

چون توابع انحنا یکسان‌اند و شرایط اولیهٔ $\sigma_۲$ همان شرایط اولیهٔ $\sigma_۱$ هستند که یک واحد به سمت راست جابه‌جا شده است، خم دوم نیز همان خم اول است که یک واحد به راست منتقل شده است (قضیهٔ بنیادی خمهای مسطح). به نظر ناظر ساکن در چارچوب مرجع آزمایشگاه، فاصلهٔ میان دو سفینه هرگز تغییر نمی‌کند. از این‌رو، هیچ‌گونه کششی به طناب کشیدهٔ میان دو سفینه وارد نمی‌شود و طناب دست‌نخورده باقی می‌ماند.

اما بعد از خاموش شدن موتورها دو سفینه در یك چارچوب لخت متحرک ساکن هستند و بنابراین، بر اثر انقباض طول، طول طناب در این چارچوب مرجع بلندتر از طول آن قبل از روشن شدن موتورهاست. از این‌رو طناب کشیده می‌شود و اگر میزان کشیدگی بیشتر از حد کشسانی باشد پاره می‌شود.

حق با کیست؟ ناظر آزمایشگاه که می‌گوید هیچ فشاری به طناب وارد نمی‌شود یا ناظرهای متحرک که می‌گویند طناب پاره می‌شود؟ بحثهای مفصل مربوط به این مسأله را می‌توان در مرجعهای [۵]، [۱]، [۳]، [۴]، [۶]، [۱۲]، و [۱۸] یافت.

در بیان مسألهٔ پارادُکس معلوم نیست که طناب دقیقاً چگونه شتاب می‌گیرد. اما، بدون توجه به نحوهٔ شتاب طناب، آنچه مسلم است این است که لبهٔ پسین توسط $\sigma_۱$ و لبهٔ پیشین توسط $\sigma_۲$ حمل می‌شود. برای سادگی، به جای طناب از دو میله استفاده می‌کنیم: میلهٔ اول از $x = ۰$ تا $x = ۰٫۵$ در $\tau = ۰$ گسترده است و توسط $\sigma_۱$ حمل می‌شود و دومی از $x = ۰٫۵$ تا $x = ۱$ در $\tau = ۰$ گسترده است و توسط $\sigma_۲$ حمل می‌شود. همچنان که سفینه‌ها شتاب می‌گیرند، ناظر واقع در آزمایشگاه مشاهده می‌کند که انتهای پیشین میلهٔ اول به عقب به طرف $\sigma_۱$ و انتهای پسین میلهٔ دوم به جلو به طرف $\sigma_۲$ منقبض می‌شود؛ بخش ۱.۴ را ببینید. ناظر آزمایشگاه می‌بیند میله‌ها، که ابتدا با هم در تماس‌اند، از هم جدا می‌شوند. نتیجه می‌گیریم که اگر سرعت نهایی به اندازهٔ کافی زیاد باشد، هر طنابی پاره خواهد شد.

به عبارت دیگر، از بخش ۲.۴ می‌دانیم که شتاب لبهٔ پسین میله که توسط $\sigma_۲$ حمل می‌شود بیشتر از $\sigma_۲$ و شتاب لبهٔ پیشین که توسط $\sigma_۱$ حمل می‌شود کمتر از $\sigma_۱$ است. دو موشک شتاب یکسانی دارند. در این صورت، میله‌ها از هم جدا خواهند شد.

اکنون چندضلعی چرخان را در نظر بگیرید. اگر به همهٔ رأسها شتاب یکسان بدهیم، آنگاه چنانچه اضلاع چندضلعی رابط فیزیکی میان رأسها باشند می‌شکنند مگر اینکه شعاع را کم کنیم. اما، در این صورت، قطار از خط خارج می‌شود. شق دیگر این است که فرض کنیم طول اضلاع به صفر می‌رسد و به همهٔ نقاط حلقهٔ حاصل شتاب یکسان بدهیم. در این حالت محیط منقبض نمی‌شود و قطار روی خط باقی می‌ماند.

### ۵.۴ دوقلوهای با شتاب یکسان

پارادُکس دیگری در مورد مطالب بخش ۴.۴ وجود دارد. این پارادُکس در مرجع [۲۰، مسألهٔ L.۱۳] تحت عنوان پارادُکس دوقلوهای با شتاب یکسان بررسی شده است.

فرض کنید خلبانهای دو سفینه دوقلوها هستند. وقتی سفینه‌ها در $\tau = ۰$ شتاب می‌گیرند دوقلوها دقیقاً همسن هستند. از دید ناظر آزمایشگاه، شتاب دوقلوها یکسان است و سن و سال آنها همواره یکسان باقی می‌ماند. با وجود این، بعد از اینکه سفینه‌ها از نقطهٔ قطع در $\tau = ۱$ می‌گذرند هر دو در یک چارچوب مرجع لخت ساکن هستند و می‌پذیرند که دوقلوی سمت راست از دوقلوی سمت چپ پیرتر است.

برای اینکه بفهمیم چرا این‌طور است، شکل ۸ را بر شبکهٔ مختصات دوقلوی دست چپ می‌گذاریم (شکل ۹).

وقتی رویدادهای روی دو جهان‌ـخط در دستگاه مختصات خلبان سمت چپ همزمان‌اند، در چارچوب مرجع آزمایشگاه رویداد سمت راست دیرتر از رویداد سمت چپ به وقوع می‌پیوندد. چون دو خلبان در چارچوب مرجع آزمایشگاه همسن و سال هستند، خلبان سمت چپ باید به این نتیجه برسد که خلبان سمت راست از او پیرتر است زیرا، در سیستم شتاب‌دار، خطوطِ همزمانی به سمت راست اوج می‌گیرند. به طریق مشابه، خلبان سمت راست نتیجه می‌گیرد که خلبان سمت چپ جوانتر است زیرا خطوط همزمانی خلبان سمت راست (که در شکل نشان داده نشده است) نیز به طرف راست صعود می‌کند.

این تصویر همچنین ایراد بحث موشکهای به‌هم‌بسته را که در [۵] توصیف شده است مشخص می‌کند. بعد از این نتیجه‌گیری که طناب پاره می‌شود، مؤلفان این سؤال را مطرح می‌کنند که وضعیت از دید خلبانِ سمت چپ چگونه به نظر می‌رسد. با ادامهٔ بحث، در ضمن زمانهای کوتاهتر و کوتاهتر

شتاب، آنها نتیجه می‌گیرند که شتاب را می‌توان یک «پرش لحظه‌ای» در نظر گرفت. در نتیجه، خلبان سمت چپ «می‌بیند» که ابتدا خلبان سمت راست حرکت می‌کند ــ چیزی که به صورت آشکار در شکل ۹ غلط است. مسأله این است که «پرش لحظه‌ای»، تکینگی در دستگاه مختصات ناظر شتاب‌دار را درست تا مبدأ بالا می‌آورد. یک چنین ناظری دیگر دستگاه مختصات ندارد.

## مراجع


1. Graciela S. Birman, and Katsumi Nomizu, Trigonometry in Lorentzian Germetry, *Amer. Math. Monthly* 91 (1984) 543-549.
2. Dario, Costellanos, Making $\pi$ Equal 3, *Amer. Math. Monthly* **100** (1993) 69.
3. Edmond M. Dewan, Stress Effects due to Lorentz Contraction, *Amer. J. Physics* **31** (1963) 383-386.
4. E. Dewan, and M. Beran, Note on Stress Effects due to Relativistic Contraction, *Amer. J. Physics* **27** (1959) 517-518.
5. George F. R. Ellis, and Ruth M. Williams, *Flat and Curved Space-Times*, Oxford University Press, New York, 1988.
6. Arthur A. Evett, Note on the Separation of Relativistic Moving Rockets, *Amer. J. Physics* **28** (1960) 566.
7. Arthur A. Evett, A Relativistic Rocket Discussion Problem, *Amer. J. Physics* **40** (1972) 1170-1171.
8. Alfred Gray, *Modern Differential Geometry of Curves and Surfaces*, Second Edition, CRC Press, Boca Raton, 1997.
9. Garry Helzer, Relativity and Plane Curves, *Mathematica* notebook, http//www.math.umd.edu/~gah/Preprints.html.
10. Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler, *Gravitation*, W. H. Freeman and Co., San Francisco, 1973.
11. Gregory L. Naber, *The Geometry of Minkowski Spacetime*, Applied Mathematical Sciences no. 32, Srpringer-Verlag, New York, 1992.
12. Paul J. Nawrocki, Stress Effects due to Relativistic Contraction, *Amer. J. Physics* **30** (1962) 771-772.
13. Rick Norwood, How to Make Pi Equal to Three, *Amer. Math. Monthly* **99** (1992) 111.
14. Rick Norwood, More on pi, *Amer. Math. Monthly* **100** (1993) 577.
15. Barret, O'Neill, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, New York, 1983.
16. Wolfgang Rindler, *Introduction to Special Relativity*, second edition, Oxford University Press, Oxford, New York, 1991.
17. Wolfgang Rindler, *Essential Relativity*, revised second edition, Springer-Verlag, New York, 1977.
18. J. E. Romain, A Geometrical Approach to Relativistic Paradoxes, *Amer. J. Physics* **31** (1963) 576-585.
19. Edwin R. Taylor, and A. P. French, Limitation on Proper Length in Special Relativity, *Amer. J. Physics* **51** (1983) 881-893.
20. Edwin Taylor, and John Archibald Wheeler, *Spacetime Physics*, Second Edition, W. H. Freeman and Co., New York, 1992.
21. I. M. Yaglom, *A Simple Non-Euclidean Geometry and its Physical Basis*, Springer-Verlag, New York, Berlin, Heidelberg, 1979.
22. I. M. Yaglom, B. A. Rozenfel'd, and E. U. Yasinskaya, Projective Metrics, *Russian Math. Surveys* **19** (1964) 49-107.


******


• Garry Helzer, "Special relativity with acceleration", *Amer. Math. Monthly*, (3) **107** (2000) 219-237.



* گری هلزر، دانشگاه مریلند، آمریکا

gah@math.umd.edu


کواین فیلسوف معروف کتابهایش را با یک ماشین تحریر مدل ۱۹۲۷ تایپ می‌کرد که رسالهٔ دکتریش را با آن تایپ کرده بود. او در این ماشین تحریر بعضی نشانه‌ها (ازجمله «؟» و «!» و «1») را با نمادهای ریاضی جایگزین کرده بود. یک بار از او پرسیدند که بدون علامت سؤال چه می‌کند؟ گفت: خُب، می‌دانید، من با قطعیات سروکار دارم.

******

منبع:

http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Quine.html

نشر ریاضی، سال ۱۲، شمارهٔ ۱ و ۲، صص ۶۱-۶۴

# آموزش و مسأله

## اتحادهایی دترمینانی و کاربردی از آنها

رحیم زارع‌نهندی*

### ۱. مقدمه

دترمینان در قسمتهای مختلف ریاضیات از مقدماتی تا پیشرفته، کاربردهای اساسی دارد. با این‌حال، بعضی از ویژگیهای اولیهٔ آن هنوز نیاز به بررسی دارد. برای مثال، اگر درایه‌های یک ستون دترمینان را جابه‌جا کنیم مقدار دترمینان چگونه تغییر می‌کند؟ روشن است که نباید انتظار جوابی ساده داشت. مثال زیر را ببینید:

$$D = \begin{vmatrix} a_{۱} & b_{۱} & c_{۱} \\ a_{۲} & b_{۲} & c_{۲} \\ a_{۳} & b_{۳} & c_{۳} \end{vmatrix}, \quad \begin{vmatrix} a_{۲} & b_{۱} & c_{۱} \\ a_{۳} & b_{۲} & c_{۲} \\ a_{۱} & b_{۳} & c_{۳} \end{vmatrix} + \begin{vmatrix} a_{۱} & b_{۲} & c_{۱} \\ a_{۲} & b_{۳} & c_{۲} \\ a_{۳} & b_{۱} & c_{۳} \end{vmatrix} + \begin{vmatrix} a_{۱} & b_{۱} & c_{۲} \\ a_{۲} & b_{۲} & c_{۳} \\ a_{۳} & b_{۳} & c_{۱} \end{vmatrix} = ۰$$

که اتحادی است در مورد دترمینانهای حاصل از اثر جایگشت دوری (۳ ۲ ۱) به‌ترتیب روی اندیسهای ستون اول، دوم و سوم دترمینان $D$.

در این مقاله به تعمیم اتحاد بالا و بررسی اتحادهای دیگر در همین زمینه خواهیم پرداخت (قضیه‌های ۱، ۲ و ۳ را ببینید). این اتحادها که حاصل جنبی مطالعات مؤلف در زمینهٔ نظریهٔ تکینه‌ها در هندسهٔ جبری بوده است (رجوع شود به [SZ])، ماهیت نسبتاً ساده‌ای دارند و می‌توان آنها و کاربردهایشان را به‌طور مستقل بررسی کرد. در این مقاله به کاربردی از این اتحادها در کاهش تعداد مولدهای ایدآل کهادهای ماتریسهای گردان و چندگردان می‌پردازیم (قضیه‌های ۴ و ۵ را ببینید).

درواقع، هر یک از خاصیتهای مقدماتی دترمینان، مانند خطی بودن مقدار دترمینان نسبت به هر ستون، یک اتحاد دترمینانی است. قضیهٔ کیلی-همیلتن که هر ماتریس ریشهٔ چندجمله‌ای مشخصهٔ آن ماتریس است، یک اتحاد ماتریسی است. برهان یکی از قضیه‌های اصلی این مقاله (قضیهٔ ۲) نتیجهٔ یکی از ویژگیهای چندجمله‌ای مشخصهٔ یک ماتریس خاص است.

تحقیقات جالبی نیز با هدفهای گوناگون در مورد اتحادهای دترمینانی انجام شده که در این زمینه می‌توان به [AB] و [K] مراجعه کرد.

بخشی از مقالهٔ حاضر، با دیدگاهی متفاوت و به صورت فشرده‌تر، در مقاله دیگری از مؤلف آمده است [Z].

### ۲. اتحادهایی دترمینانی

در این مقاله $R$ یک حلقهٔ جابجایی یکدار و $M = [C_{۱}C_{۲}\cdots C_n]$ یک ماتریس $n \times n$ با درایه‌های متعلق به $R$ و با ستونهای $C_i$ فرض می‌شود. نگاشت دلخواه $\theta = \{۱, ۲, \cdots, n\} \to \{۱, ۲, \cdots, n\}$ را در نظر می‌گیریم. برای یک ستون $C = \begin{bmatrix} x_{۱} \\ x_{۲} \\ \vdots \\ x_n \end{bmatrix}$ از ستونهای ماتریس $M$ قرار می‌دهیم:

$$\theta(C) = \begin{bmatrix} x_{\theta(۱)} \\ x_{\theta(۲)} \\ \vdots \\ x_{\theta(n)} \end{bmatrix}$$

و آن را اثر $\theta$ روی ستون $C$ گوییم. به‌طور مشابه، ماتریس $\theta(M)$ ماتریسی است که از اثر $\theta$ روی هر یک از ستونهای $M$ به‌دست می‌آید. روشن است

که $\theta(M) = \theta(I)\cdot M$ که $I$ ماتریس همانی مرتبهٔ $n$ است. اگر $\theta$ یک جایگشت باشد، $\theta(I)$ را یک «ماتریس جایگشت» گویند.

**قضیهٔ ۱.** فرض می‌کنیم $\theta = \{1,2,\cdots,n\} \to \{1,2,\cdots,n\}$ نگاشتی با $r$ نقطهٔ ثابت است، اگر $M_j$ ماتریس حاصل از $M$ تحت اثر $\theta$ روی ستون $j$ام آن باشد، آنگاه

$$\sum_{j=1}^{n} \det(M_j) = r\cdot\det(M)$$

برهان، هر چند برهان این قضیه از روش اثبات قضیهٔ ۲ نیز نتیجه می‌شود، برهانی مقدماتی هم وجود دارد که در اینجا می‌آوریم.

فرض می‌کنیم $M=(m_{ij})_{1\le i,j\le n}$ و $\overline{M}_{ij}$ ماتریس به‌دست آمده از $M$ پس از حذف سطر $i$ام و ستون $j$ام آن باشد. مقدار دترمینان $M_j$ پس از بسط آن نسبت به ستون $j$ام چنین است:

$$\det(M_j)=\sum_{i=1}^{n}(-1)^{i+j}m_{\theta(i)j}\overline{M}_{ij}$$

با جمع طرفین این برابریها به‌ازای $j=1,2,\cdots,n$ و تعویض ترتیب جمعبندی در عبارت حاصل داریم

$$\begin{aligned}\sum_{j=1}^{n}\det(M_j) &= \sum_{j=1}^{n}\sum_{i=1}^{n}(-1)^{i+j}m_{\theta(i)j}\overline{M}_{ij}\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{i+j}m_{\theta(i)j}\overline{M}_{ij}\end{aligned}$$

ولی $\sum_{j=1}^{n}(-1)^{i+j}m_{\theta(i)j}\overline{M}_{ij}$ دترمینان ماتریسی است که از جایگزینی سطر $i$ام ماتریس $M$ با سطر $\theta(i)$ام آن حاصل می‌شود. این دترمینان برابر دترمینان $M$ است اگر $i$ یک ثابت $\theta$ باشد و در غیر این‌صورت برابر صفر است. با این ترتیب قضیه ثابت می‌شود. □

**قضیهٔ ۲.** برای هر نگاشت دلخواه

$$\theta=\{1,2,\cdots,n\}\to\{1,2,\cdots,n\}$$

و هر عدد $p$، $0<p<n$، و اعداد $i_1$، $i_2$، ...، $i_p$ که $1\le i_1<i_2<\cdots<i_p\le n$، فرض می‌کنیم $M_{i_1,i_2,\dots,i_p}$ ماتریس به‌دست آمده از $M$ با جایگزینی $C_{i_j}$، ستون $i_j$ام $M$، توسط $\theta(C_{i_j})$ باشد که $j=1,\dots,p$. اگر $\theta$ جایگشتی دوری به طول $n$ باشد آنگاه

$$\sum_{i\le i<i_2<\ldots<i_p\le n}\det(M_{i_1,i_2,\dots,i_p})=0$$

برهان، ابتدا $\theta$ را یک نگاشت دلخواه روی $\{1,2,\dots,n\}$ می‌گیریم. اگر $\lambda$ یک متغیر باشد، آنگاه برای $M=[C_1C_2\cdots C_n]$ داریم

$$\begin{aligned}&|M+\lambda\theta(M)|\\ &=|C_1+\lambda\theta(C_1)\ \ C_2+\lambda\theta(C_2)\cdots C_n+\lambda\theta(C_n)|\end{aligned}$$

دترمینان طرف دوم برابری بالا را «برحسب توانهای مختلف $\lambda$» بسط می‌دهیم

$$\begin{aligned}|M+\lambda\theta(M)| &= |M|+\lambda\sum_{1\le j\le n}|M_j|+\cdots+\\ &\lambda^p\sum_{1\le i_1<\ldots<i_p\le n}\left|M_{i_1,\dots,i_p}\right|+\cdots+\lambda^n\left|M_{1,2,\dots n}\right| \quad (*)\end{aligned}$$

از طرف دیگر اگر $J=\theta(I)$ ماتریس متناظر با $\theta$ باشد، داریم

$$|M+\lambda\theta(M)|=|M+\lambda JM|=|I+\lambda J|\cdot|M| \qquad (**)$$

توجه داریم که

$$|I+\lambda J|=\lambda^n\left|\frac{1}{\lambda}I+J\right|=\lambda^nQ(\frac{1}{\lambda})$$

که در آن $Q(\lambda)$ چندجمله‌ای مشخصهٔ ماتریس $-J$ است. بنا به ویژگیهای چندجمله‌ای مشخصهٔ یک ماتریس، ضریب $\lambda^{n-1}$ در $Q(\lambda)$ برابر اثر ماتریس $J$ است که همان تعداد نقاط ثابت نگاشت $\theta$ است و لذا با مساوی قراردادن ضریب $\lambda$ در $(*)$ و $(**)$، برهان دیگری برای قضیهٔ ۱ به‌دست می‌آید.

برای اثبات قضیهٔ ۲، کافی است توجه شود که به‌ازای هر جایگشت دوری به طول $n$ داریم

$$|I+\lambda J|=1+(-1)^{n-1}\lambda^n$$

بنابراین نتیجه می‌شود که ضریب $\lambda^p$ در $(*)$ صفر است، که حکم قضیه است. □

**قضیهٔ ۳.** با نمادگذاری قضیهٔ ۲، به‌ازای هر عدد صحیح $q$، $0\le q\le n$ و هر جایگشت دوری $\theta$ به طول $n$ داریم

$$\begin{aligned}&|C_1\cdots C_{q-1}\ \theta^p(C_q)\ C_{q+1}\cdots C_n|\\ &=(-1)^p\sum_{\substack{1\le i_1<\cdots<i_p\le n\\ \forall_j,\, i_j\ne q}}\det(M_{i_1},\cdots,i_p)\end{aligned}$$

برهان، بدون از دست‌دادن کلیت، می‌توان فرض کرد $q=1$. قضیه را با استقرا روی $p$ ثابت می‌کنیم. حالت $p=1$ حکم قضیه ۱ است زیرا به‌ازای هر جایگشت دوری $\theta$ به طول $n$ داریم $r=0$. با فرض درستی حکم برای $p$، آن‌را در مورد ماتریس

$$[\theta(C_1)C_2\cdots C_n]$$

به‌کار می‌بریم. در طرف راست اتحاد حاصل، ستون اول همهٔ دترمینانها، $\theta(C_1)$ است. حال با به‌کارگیری قضیهٔ ۲ در مورد ماتریس $M=[C_1C_2\cdots C_n]$ برای $p+1$ و جایگذاری، طرف راست اتحاد حاصل فوق‌الذکر به‌صورت مجموع دترمینانهایی که ستون اول آنها $C_1$ و $p+1$ ستون از ستونهای دیگرشان اثر $\theta$ روی ستونهای متناظر از $M$ است، در می‌آید که حکم موردنظر برای $p+1$ است. □

**توضیح.** اثر یک نگاشت $\theta=\{1,2,\cdots,n\}\to\{1,2,\cdots,n\}$ روی یک سطر از ماتریسی $n\times n$، مشابه اثر آن روی یک ستون تعریف می‌شود. روشن است که اگر در عبارتهای قضیه‌های ۱، ۲ و ۳ به‌جای ستونها، سطرها را در نظر بگیریم، این قضیه‌ها با اِعمال تغییرات بدیهی معتبرند و برهان آنها نیز با اعمال تغییرات متناظر پابرجاست.

## ۳. ماتریسهای گردان و چندگردان و ایدآلهای کهادهای آنها

ماتریس گردان[1] [یا دوری] ماتریسی است مربع که ستون اول آن عناصری در حلقهٔ $R$اند و ستون $i$ام آن با انتقال $i-1$ گامیِ ستون اول به‌صورت دوری به‌دست می‌آید، یعنی

$$M=\begin{bmatrix} a_n & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_n & \cdots & a_{n-3} & a_{n-2} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_2 & a_3 & \cdots & a_n & a_1 \\ a_1 & a_2 & \cdots & a_{n-1} & a_n \end{bmatrix}, a_i\in R$$

ماتریس چندگردان[2] با $b$ بلوک، ماتریسی است مانند $P$ که از کنار هم گذاشتن $b$ ماتریس گردان هم‌مرتبه به‌صورت افقی ، حاصل می‌شود . یعنی

$$P=[M_1\, M_2 \cdots M_b]$$

که $M_i$ یک ماتریس گردان است. در این مقاله $b=2$ فرض خواهد شد.

ماتریسهای گردان کاربردهای زیادی در محاسبات ماتریسی، نظریهٔ عملگرها، آنالیز عددی و جبر ترکیبیاتی دارند (رجوع شود به [D] و مراجع ذکرشده در آن). ماتریسهای چندگردان به‌طور طبیعی در مطالعهٔ ایدآلهای تکنیه‌های عام[3] در هندسهٔ جبری ظاهر می‌شوند [SZ].

حال کاربردهایی از نتایج بخش ۲ را در مورد ماتریسهای گردان و چندگردان ارائه می‌کنیم. یادآوری می‌کنیم که دترمینان یک زیرماتریس مربع $r\times r$ از یک ماتریس را یک کهاد $r\times r$ آن ماتریس گوییم. به‌ازای $r$ ماکسیمم، کهاد را کهاد ماکسیمال ماتریس می‌نامیم.

**تعریف ۱.** فرض می‌کنیم $P=[M_1M_2]$ یک ماتریس چندگردان با دو بلوک $n\times n$ است. برای هر زیرماتریس مربع ماکسیمال $A$ در $P$، اگر $t$ تعداد ستونهای $A$ از ماتریس $M_1$ باشد، $t$ را نوع[4] $A$ گوییم. $t$ را نوع کهاد $|A|$ نیز می‌گوییم.

**تعریف ۲.** فرض می‌کنیم $A$ یک زیرماتریس مربع ماکسیمال $P=[M_1M_2]$ از نوع $t$ است. $A$ را می‌توان به‌صورت $[D_1D_2]$ نوشت که $D_1$ یک زیرماتریس $t$ ستونی $M_1$ و $D_2$ یک زیرماتریس $n-t$ ستونی $M_2$ است. $|A|$ را یک کهاد ماکسیمال اصلی $P$ گوییم هرگاه ستونهای $D_1$ دقیقاً $t$ ستون اول $M_1$ باشند.

**قضیهٔ ۴.** هر کهاد ماکسیمال $P=[M_1M_2]$ از نوع $t$ ترکیبی $\mathbb{Z}$ـخطی از کهادهای ماکسیمال اصلی از همان نوع است.

برهان. $|A|$ را یک کهاد ماکسیمال $P$ از نوع $t$ می‌گیریم. داریم $A=[D_1D_2]$ که $D_1$ یک زیرماتریس $M_1$ و $D_2$ یک ماتریس $M_2$ است. به‌ازای $p\geq 1$ فرض کنید $p-1$ ستون اول $D_1$ با $p-1$ ستون اول $M_1$ یکسان باشد ولی ستون $p$ام $D_1$، ستون $k$ام $M_1$ باشد که $k>p$. ستون $p$ام $D_1$ را $C$ می‌گیریم. فرض کنید $E$ ماتریس حاصل از $D$ با جایگزینی $C$ توسط ستون $p$ام $M_1$ باشد. قضیهٔ ۳ را در مورد $E$ برای $\theta^{k-p}$ با فرض $\theta=(1,2,\ldots,n)$ به‌کار می‌بریم. با توجه به اینکه $M_1$ یک ماتریس‌گردان است، طرف چپ اتحاد قضیهٔ ۳ پس از کاربرد آن روی $E$، برابر $|A|$ است. در طرف راست اتحاد قضیهٔ ۳، وقتی $\theta$ روی $p-1$ ستون اول $E$ اثر می‌کند، حداقل دو ستون ماتریس حاصل مساوی می‌شوند و در نتیجه دترمینان مربوطه صفر خواهد بود. در بقیهٔ دترمینانهای طرف راست اتحاد قضیهٔ ۳، $p$ ستون اول دترمینانها، با $p$ ستون اول ماتریس $M_1$ یکسان خواهد بود. در نتیجه $|A|$ به‌صورت یک ترکیب $\mathbb{Z}$ـخطی از کهادهای ماکسیمال $P$ از نوع $t$ که در آنها $p$ ستون اول با $p$ ستون اول $M_1$ یکسان است، نوشته می‌شود. با اِعمال این روش برای هر یک از جمعوندها در ترکیب $\mathbb{Z}$ـخطی حاصله برای $|A|$، که روی ستون $p+1$ آنها انجام می‌شود، و تکرار این روش برای ستونهای بعدی جمعوندها در هر مرحله، در نهایت $|A|$ به‌صورت یک ترکیب $\mathbb{Z}$ـخطی از کهادهایی که $t$ ستون اول آنها $t$ ستون اول $M_1$اند، نوشته می‌شود، یعنی به‌صورت ترکیبی $\mathbb{Z}$ـخطی از کهادهای اصلی از نوع $t$. □

از قضیهٔ ۴ نتیجه می‌شود که مجموعهٔ کهادهای ماکسیمال اصلی، تشکیل یک مولد برای ایدآل پدید آمده به‌وسیلهٔ کهادهای ماکسیمال $P$ می‌دهد. تعداد عناصر این مجموعهٔ مولد، با شمارش تعداد کهادهای ماکسیمال اصلی از نوع $t$، $t=0,1,2,\cdots,n$، برابر است با $1+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n}=2^n$، در حالی‌که تعداد کل کهادهای ماکسیمال $P$ برابر $\binom{2n}{n}$ است.

در خاتمه به بررسی کهادهای $r\times r$ یک ماتریس گردان می‌پردازیم.

**قضیهٔ ۵.** فرض کنید $M$ یک ماتریس گردان $n\times n$ است. به‌ازای هر $r$، $1\leq r\leq n$، هر کهاد $r\times r$ از $M$ یک ترکیب $\mathbb{Z}$ـخطی از کهادهای $r\times r$ ماتریس $N$ است که ماتریس $N$ از $r$ ستون اول $M$ تشکیل یافته است.

برهان. کافی است $1\leq r<n$ فرض شود. اگر $I$ ماتریس همانی $n\times n$ باشد، ماتریس $P=[M\ I]$ یک ماتریس چندگردان با دو بلوک است. یک نگاشت دوسویی بین مجموعهٔ زیرماتریسهای $r\times r$ از $M$ و مجموعهٔ زیرماتریسهای مربع ماکسیمال $P$ از نوع $r$ تعریف می‌کنیم: فرض می‌کنیم $D$ یک زیرماتریس $r\times r$ در $M$ باشد که با سطرهای $i_1$، $i_2$، $\cdots$، $i_r$ از $M$ حاصل شده است. اگر $J$ زیرماتریس $n\times(n-r)$ای از $I$ باشد که با حذف ستونهای $i_1$، $i_2\cdots$، $i_r$ در $I$ به‌دست می‌آید و اگر $E$ زیرماتریس $n\times r$ای از $M$ باشد که شامل ستونهای $D$ است، آنگاه تناظر $[E\ J]\leftrightarrow D$، تناظری یک‌به‌یک بین زیرماتریسهای $r\times r$ از $M$ و زیرماتریسهای مربع ماکسیمال $P$ از نوع $r$ است. حال با توجه به اینکه $|E\ J|=|D|$، و اینکه $|E\ J|$ کهاد ماکسیمال اصلی $P$ است اگر و تنها اگر $E=N$، حکم قضیه، از قضیه ۴ نتیجه می‌شود چون $D$های متناظر $[N J]$ زیرماتریسهایی $r\times r$ از $N$ خواهند بود. □

از قضیهٔ ۵ نتیجه می‌شود که مجموعهٔ کهادهای $r\times r$ از $N$ یک مولد ایدآل پدید آمده توسط کهادهای $r\times r$ در $M$ است. این نتیجه قبلاً توسط واتانابه در ([W]، گزارهٔ ۷) با روشی متفاوت ثابت شده است.

1. circulant matrix
2. pluri-circulant matrix
3. generic singularities
4. type

برای تعمیم قضیه‌های ۴ و ۵ به ماتریسهای چندگردان با تعداد بلوکهای دلخواه، خوانندهٔ علاقه‌مند می‌تواند به [Z] مراجعه نماید.

## سپاسگزاری

نویسندهٔ مقاله لازم می‌داند از آقای عبدالحسین هورفر دانشجوی دکتری ریاضی دانشگاه تهران به جهت تبادل نظری که منجر به اثبات قضیه ۲ به روش مندرج در این مقاله شد، تشکر نماید. همچنین مؤلف از حمایت مالی دانشگاه تهران سپاسگزاری می‌کند.

## مراجع

[AB] G. E. Andrews and W. H. Burge, "Determinant identities", *Pacific J. Math.* **158** (1993) 1-14.

[D] P.J. Davis, *Circulant Matrices*, John Wiley & Sons, New York (1979).

[K] C. Krattenthaler, Determinant identities and a generalization of the number of totally symmetric self-complementary plane partitions", *arXiv: math. CO/9712202 26 Nov* (1997).

[SZ] P. Salmon, R. Zaare-Nahandi, "On the local equations at ordinary unibranched singularities", preprint.

[W] J. Watanabe, "Hankel matrices and Hankel ideals", *Proc. Schl. Sci., Tokai Univ.* **32** (1997) 11-21; *Queen's Papers in Pure and Appl. Math.* **102** (1996) 351-363.

[Z] R. Zaare-Nahandi, "On the ideals of minors of pluri-circulant matrices", preprint.

* * * * * *

* رحیم زارع‌نهندی، دانشگاه تهران و مرکز فیزیک نظری و ریاضیات سازمان انرژی اتمی ایران

rahimzn@khayam.ut.ac.ir

## نخستین کاربرد نماد انتگرال

لایب‌نیتس پیش از معرفی نماد انتگرال، علامت اختصاری. onm (مخفف omina به‌معنی مجموع یا کل) را در جلو عبارتی که باید از آن انتگرال گرفته شود می‌نوشت.

نماد انتگرال را نخستین بار گوتفرید ویلهلم لایب‌نیتس (۱۶۴۶ـ۱۷۱۶) در ۲۹ اکتبر سال ۱۶۷۵، در دستنوشته‌ای منتشر نشده به‌کار برد. چند هفته بعد، در ۲۱ نوامبر، برای اولین بار $dx$ را بعد از نماد انتگرال قرار داد. او بعداً در همان سال ۱۶۷۵ استفاده از این نماد را در نامه‌ای به هنری اولدنبرگ، دبیر انجمن سلطنتی [آکادمی علوم] پیشنهاد کرد: «مناسب است به جای .omn بنویسیم $\int$، بنابراین $\mathrm{omn}.l = \int l$، و یا $\int l$ برابر است با مجموع همهٔ $l$ها.» نخستین کاربرد نماد انتگرال در یک متن چاپی، در مقاله‌ای به قلم لایب‌نیتس در «آکتا اِرودیتوروم» بود. نماد انتگرال درواقع صورت کشیدهٔ حرف $S$ است که حرف اول summa [مجموع] است.

نیوتن در سال ۱۷۰۴ در رسالهٔ تعیین مساحت زیرخم، برای نشان دادن انتگرال $x$، خط عمودی کوچکی در بالای $x$ گذاشته بود، و برای انتگرال $x$ای که یک خط در بالای آن بود، دو خط کنار هم به‌کار برده بود. نماد دیگری که نیوتن برای نمایش انتگرال به‌کار می‌برد، محصورکردن عبارت در یک مستطیل بود. کایوری [Cajori] می‌نویسد که نمادهای نیوتن برای انتگرالگیری، اشکال داشت زیرا $x$ با خطی در بالایش، با $x$ پریم اشتباه می‌شد و قرار دادن عبارت در مستطیل برای چاپچی مشکل بود و به این دلیل نمادهای او، حتی در انگلستان، مقبولیت نیافت.

**نمادهای حدود انتگرالگیری.** حدود انتگرالگیری در آغاز فقط با کلمات نشان داده می‌شد. اویلر برای نخستین بار در رسالهٔ تأسیس حساب «انتگرال» از نماد استفاده کرد. او حدود را در کروشه می‌گذاشت و کلمات لاتینی $ab$ و $ad$ را به‌کار می‌برد.

مبتکر نماد جدید انتگرال معین، ژان باتیست فوریه (۱۷۶۸ـ۱۸۳۰) است که در سال ۱۸۲۲ در کتاب معروفش، نظریهٔ تحلیلی گرما، آن را به‌کار برد. البته این نماد را قبلاً در «گزارشهای آکادمی فرانسه (۱۸۱۹ـ۱۸۲۰)» در مقاله‌ای که قسمت اول کتاب نامبرده تجدید چاپ آن است، به‌کار برده بود.

* * * * * *

منبع:
http://members.aol.com/jeff570/calculus.html

نشر ریاضی، سال ۱۲، شمارهٔ ۱ و ۲، صص ۶۵-۶۷

# مسأله یا نظریه

مریم میرزاخانی*

در جای‌جای ریاضیات به شگردهایی برای اثبات قضایا برمی‌خوریم که اگر به‌طور منفرد در نظر گرفته شوند می‌توانند منشأ مسائل گوناگون و جالب‌توجهی باشند. اگر این شگردها، و قضایای خاصی که به وسیله آنها ثابت می‌شوند خارج از متن نظریهٔ مربوط مطرح شوند واقعیتهای حیرت‌آوری به نظر می‌رسند. این نوشته گردایه‌ای از این نوع مسائل از دیدگاه نویسنده است. در بعضی از این مسائل فرض بر این است که خواننده معلوماتی کمی فراتر از ریاضیات سالهای اول دانشگاه دارد.

**مسألهٔ ۱.** فرض کنید $f:\mathbb{C}\to\mathbb{C}$ تابعی تحلیلی باشد. ثابت کنید که اگر $M(r)=\sup\limits_{|z|=r}|f(z)|$، و $r$ای بزرگتر از ۱ وجود داشته باشد به‌طوری که $M(r)\cdot M(r^3)=M(r^2)^2$، آنگاه به‌ازای عدد مختلط مناسب $\alpha$ و عدد طبیعی مناسب $n$ داریم $f(z)=\alpha z^n$.

حل. ایدهٔ حل مسأله، استفاده از اصل ماکسیمم برای توابع تحلیلی است. اگر

$$\alpha=\frac{\log M(r)-\log M(r^3)}{2\log r}$$

و تابع $h(z)=z^\alpha f(z)$ را در نظر بگیریم، این «تابع» درواقع یک‌مقداری نیست ولی تابعی تحلیلی روی رویهٔ ریمانی مناسب است، و مهمتر اینکه $|h(z)|$ تابعی یک‌مقداری است. پس بنابر اصل ماکسیمم، این تابع ماکسیمم قدر مطلق خود در $A=\{z\,\big|\,r<|z|<r^3\}$ را روی مرز می‌گیرد. ولی به سادگی می‌توان دید که درواقع $\alpha$ چنان انتخاب شده که این ماکسیمم روی دو دایرهٔ $|z|=r$ و $|z|=r^3$ مساوی است، و درواقع

$$r^{3\alpha}M(r^3)=r^{2\alpha}M(r^2)=r^\alpha M(r)=\operatorname*{Max}_{r\le|z|\le r^3}|z^\alpha f(z)|$$

بنابراین از آنجاکه $r<r^2<r^3$، طبق اصل ماکسیمم داریم $f(z)\cdot z^\alpha=c$؛ در نتیجه $f(z)=c\cdot z^{-\alpha}$ و از آنجا که $f$ تابعی تحلیلی است، $-\alpha\in N\cup\{0\}$.

درواقع با همین ایده می‌توان ثابت کرد که برای هر تابع تحلیلی $f$ تابع $\log M(r)$ تابعی محدب از $\log r$ است.

موضوع مسألهٔ فوق برگرفته از مطالب آنالیز مختلط حول‌وحوش «قضیهٔ سه دایرهٔ آدامار» [۴] است.

حتماً شما هم تجربهٔ برخورد با مسائلی را داشته‌اید که علی‌رغم ظاهر ساده‌شان، حل آنها بسیار دشوار است. گاهی برای حل این مسائل به استفاده از ساختارهای پیشرفته‌تر ریاضی نیاز داریم و در غیر این‌صورت حل آنها به میزان قابل ملاحظه‌ای پیچیده‌تر می‌شود. توپولوژی جبری منبعی غنی از این‌گونه مسائل است. در اینجا معمولاً مسائل توپولوژی با بهره‌گیری از روشهای توپولوژی جبری به مسائل جبری مبدل می‌شوند. مرجع متعارفی برای مطالب توپولوژی جبری، مرجع [۳] است.

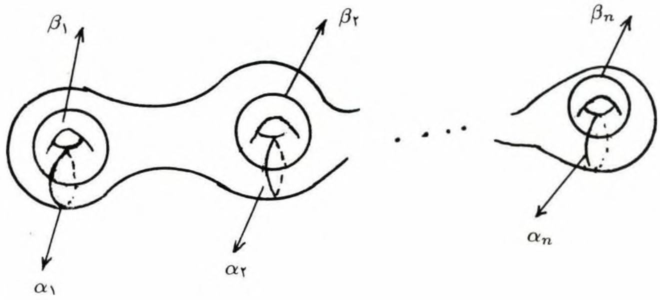
**مسألهٔ ۲.** اگر $M_n$ فضایی همسانریخت با چنبرهٔ $n$ سوراخه باشد و

$$f:M_g\to M_h\qquad g<h$$

تابعی پیوسته باشد، آنگاه $\deg f=0$.

حل. تابع $f$، تابع $f^*:H^*(M_h,\mathbb{Z})\to H^*(M_g,\mathbb{Z})$ را القا می‌کند که درواقع یک همریختی حلقه‌ای از $H^*(M_h,\mathbb{Z})$ است. در اینجا $H^*(M_g,\mathbb{Z})$ حلقهٔ کوهومولوژی مجهز به ضرب ناوی[1] است. اما ساختار حلقه‌ای $H^*(M_n,\mathbb{Z})$ به‌سادگی قابل شناسایی است. درواقع اگر $\widehat{\alpha}_i$ و $\widehat{\beta}_i$

1. cup product




عناصر $H^1(M_n,\mathbb{Z})$ متناظر با خمهای سادهٔ بستهٔ شکل باشند، ساختار ضربی $H^*(M_n,\mathbb{Z})$ با روابط

$$\widehat{\alpha_i}\smile\widehat{\alpha_j}=0$$
$$\widehat{\beta_i}\smile\widehat{\beta_j}=0$$
$$\widehat{\alpha_i}\smile\widehat{\beta_j}=\widehat{\gamma}$$

که $\gamma$ مولد $H^2(M_n,\mathbb{Z})$ است مشخص می‌شود. با توجه به این ساختار به‌سادگی می‌توان دید که به‌ازای هر $0\neq\alpha\in H^1(M_n,\mathbb{Z})$، $\beta\in H^1(M_n,\mathbb{Z})$ وجود دارد که $\gamma=\alpha\smile\beta$ یک مولد $H^2(M,\mathbb{Z})$ است.

حال به مسأله برمی‌گردیم. $f_1^*: H^1(M_h,\mathbb{Z})\to H^1(M_g,\mathbb{Z})$ یک همریختی گروهی است اما $H'(M_k,\mathbb{Z})=\underbrace{\mathbb{Z}\oplus\cdots\oplus\mathbb{Z}}_{k}$. پس $f_1^*$ درواقع یک همریختی گروهی ازگروه آزاد آبلی با رتبهٔ $h$ به گروه آبلی آزاد با رتبهٔ $g$ است که $h>g$؛ بنابراین $\ker(f_1^*)\neq 0$ و $\alpha\neq 0$ در $H^1(M_k,\mathbb{Z})$ وجود دارد که $f_1^*\alpha=0$. حال اگر $\beta$ و $\gamma$ را طبق (۲) بگیریم، $\gamma$ مولد $H^2(M,\mathbb{Z})$ است، $f^*(\gamma)=f^*(\alpha)\smile f^*(\beta)=0\smile f^*(\beta)=0$. بنابراین $\deg f=0$.

توجه: برخلاف حالت $h=0$، در حالت کلی لزومی ندارد که چنین تابعی با تابع ثابت هوموتوپ باشد.

مسألهٔ زیر مثال دیگری است از استفادهٔ مشابه از ساختار حلقه‌ای کوهومولوژی فضاهای توپولوژیک در مسأله‌ای که برای بیان آن فقط از مفاهیم ابتداییتر استفاده می‌شود.

**مسألهٔ ۳.** فرض کنید $f: S^{2n}\times S^{2n}\to S^{2n}$ تابعی پیوسته باشد و $(x_0,y_0)\in S^{2n}\times S^{2n}$. ثابت کنید که اگر $\alpha=\deg(f|_{S^{2n}\times\{y_0\}})$ و $\beta=\deg(f|_{\{x_0\}\times S^{2n}})$، آنگاه $\alpha\beta=0$.

حل. فرض کنید $s\in H^{2n}(S^{2n},\mathbb{Z})$ یک مولد باشد. بنابر قضیهٔ کونت[1] $H^{2n}(S^{2n}\times S^{2n},\mathbb{Z})$ توسط دو عضو $s\times 1$ و $1\times s$ تولید می‌شود و

$$f^*(s)=\alpha(s\times 1)+B(1\times s)$$

1. künneth

اما از آنجا که $S^{2n}$ کره‌ای زوج بعدی است،

$$(1\times s)\smile(s\times 1)=(s\times 1)\smile(1\times s)=s\times s$$

بنابراین

$$0=f^*(s\smile s)=f^*(s)^2=[\alpha(s\times 1)+\beta(1\times s)]^2$$
$$=(s^2\times 1)+\beta^2(1\times s^2)+2\alpha\beta(s\times s)=2\alpha\beta(s\times s)$$

پس $\alpha\beta=0$!

شایان ذکر است که در حالت کرهٔ فرد بعدی مسأله غلط است. به عنوان مثال، برای تابع $f:S^1\times S^1\to S^1$ داریم

$$f(e^{i\theta_1},e^{i\theta_2})=e^{i(\theta_1+\theta_2)}$$

مسألهٔ بعدی ما مثالی از نظریهٔ اندازه است که قبل از بیان آن بهتر است کمی راجع به قضیهٔ مفید و زیبای «چگالی لبگ» توضیح دهیم. اگر بخواهیم بعضی از خواص روی بازه‌های $\mathbb{R}$ را به مجموعه‌ای اندازه‌پذیر مثل $E$ (تقریباً همه جا) گسترش دهیم، احتیاج به پیداکردن نقاطی در $E$ داریم که در نزدیکی آنها نقاط بسیار دیگری از $E$ وجود دارد. قضیهٔ چگالی لبگ وجود چنین نقاطی را تضمین می‌کند.

به بیان دقیقتر اگر $E\subset\mathbb{R}$ مجموعه‌ای اندازه‌پذیر باشد، گوییم نقطهٔ $x\in E$ دارای چگالی $d$ است اگر $\lim\limits_{h\to 0}\frac{(E\cap[x-h,x+h]}{2h}$ وجود داشته و برابر $d$ باشد.

اگر مجموعهٔ نقاطی از $\mathbb{R}$ را که دارای چگالی ۱ هستند $\phi(E)$ بنامیم، محتوای قضیهٔ لبگ این است که $\phi(E)$ اندازه‌پذیر و شامل تقریباً تمام نقاط $E$ می‌شود و تقریباً زیرمجموعهٔ $E$ است! به‌عنوان مثال، امکان ندارد که $E$ و مکملش هرکدام تقریباً نصف هر بازهٔ $\mathbb{R}$ را اشغال کنند.

**قضیهٔ چگالی لبگ.** برای هر مجموعهٔ اندازه‌پذیر $E\subset\mathbb{R}$، $\mu(E\Delta\phi(E))=0$.

(برای اثبات و توضیح بیشتر، مراجعه به [۲] توصیه می‌شود.)

مسألهٔ زیر مثالی زیبا و نابدیهی از استفاده از قضیهٔ چگالی لبگ در حل مسائل نظریهٔ اندازه است.

**مسألهٔ ۴.** فرض کنید $f$ یک همسانریختی $f: D_1\to D_2$ است که $D_1$ و $D_2$ در $\mathbb{C}$ باز هستند به‌طوری که مشتقات جزئی $f_x$ و $f_y$ تقریباً همه جا وجود دارند. ثابت کنید که در این صورت $f$ تقریباً همه جا مشتق‌پذیر است این مطلب در نظریهٔ نگاشتهای شبه‌همدیس[1] کاربردهای مهمی دارد [۱].

حل. بی‌کاستن ازکلیت مسأله می‌توان فرض کرد مساحت $D_1$ متناهی است. اگر $\epsilon>0$ به‌ازای هر $n$،

$$g_n(z)=\sup_{0<|h|<\frac{1}{n}}\left|\frac{f(z+h)-f(z)}{h}-f_x(z)\right|$$

1. quasi-conformal mappings

بنابراین (تقریبا همه جا) روی $D_1$، $\lim g_n = 0$. پس با توجه به قضیهٔ اگوروف[1]، $E \subset D$ وجود دارد که مساحت $(D-E)$ کوچکتر از $\epsilon$ باشد (حد فوق به سمت ۰ یکنواخت-همگراست). به‌عبارت دیگر $E$ با این خاصیت وجود دارد که به‌ازای هر $z \in E$، $\frac{f(z+ih)-f(z)}{h}$ به‌طور یکنواخت به $f_y(z)$ و $\frac{f(z+h)-f(z)}{h}$ به‌طور یکنواخت به $f_x(z)$ میل کند.

توجه کنید که همگرایی یکنواختْ پیوستگی $f_x$ و $f_y$ را روی $E$ ایجاب می‌کند ($f$ پیوسته است). حال برای اثبات حکم مسأله کافی است ثابت کنیم که $f$ تقریباً همه جا روی $E$ مشتق‌پذیر است. بنابر قضیهٔ فوبینی و «چگالی لبگ» روی تقریباً تمام نقاط $z_0 = x_0 + iy_0 \in E$، $x_0$ و $y_0$ نقاط با چگالی ۱ برای $E_{x_0} = \{y \in \mathbb{R} | x_0 + iy \in E\}$ و $E_{y_0} = \{x \in \mathbb{R} | x + iy_0 \in E\}$ هستند.

حال ثابت می‌کنیم $f$ در چنین نقطه‌ای مشتق‌پذیر است. برای راحتی نوشتن نمادها فرض کنید $z_0 = 0$ و $0 < \eta < 1$ ثابت باشد. داریم

$$I(z) = |f(z) - f(0) - xf_x(0) - yf_y(0)|$$

$$I(z) \le |f(x+iy) - f(x) - yf_y(x)| + |f(x) - f(0) - xf_x(0)| + |y(f_y(x) - f_y(0))|$$

پس $\delta$ی مثبتی وجود دارد که

$I(z) \le \eta|z|$ به‌ازای هر $|z| < \delta$ که $x \in E$

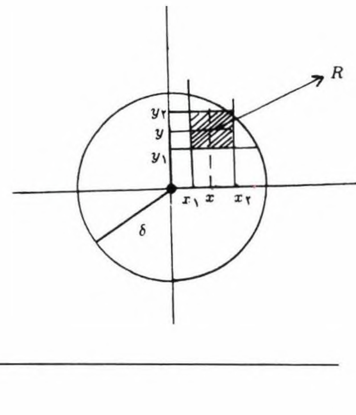
$I(z) \le \eta|z|$ به‌ازای هر $|z| < \delta$ که $iy \in E$

اما از آنجا که $x_0 = 0$ و $y_0 = 0$ نقاط با چگالی ۱ در $E_{y_0}$ و $E_{x_0}$ هستند، می‌توان $\delta$ی کوچکتری پیدا کرد که به‌ازای هر $x, y > 0$ که $z = x + iy$ و $|z| < \delta/2$، نقاط $x_1$، $x_2$ و $iy_1$، $iy_2$ از $E$ وجود داشته باشند که

$$(1-\eta)x < x_1 < x < x_2 < (1+\eta)x$$

$$(1-\eta)y < y_1 < y < y_2 < (1+\eta)y$$

در این صورت بنابر (۱)، به‌ازای هر $z^*$ در مرز $R$، $\partial R$، که $R = [x_1, x_2] \times [y_1, y_2]$، داریم $I(z^*) \le \eta(z^*)$. از طرف دیگر، بنابر




همسانریخت‌بودن $f$، $f$ یک تابع باز است و در اصل ماکسیمم صدق می‌کند. بنابراین

$$I(z) \le |f(z^*) - f(0) - xf_x(0) - yf_y(0)|$$
$$\le I(z^*) + |z - z^*|(|f_x(0)| + |f_y(0)|)$$

اما $|z - z^*| < \eta|z|$؛ در نتیجه به‌ازای هر $z$ با $x, y > 0$ که $|z| < \delta/2$، $I(z) \le 2\eta|z|(1 + |f_x(0)| + |f_y(0)|)$. با استدلالی مشابه، این نابرابری به‌ازای هر $|z| < \delta/2$ برقرار است و این، مشتق‌پذیری $f$ را در صفر نتیجه می‌دهد!

در پایان چند مسأله برای خوانندگان علاقه‌مند به حل مسائل چالش‌برانگیز می‌آوریم.

۱. فرض کنید $G$ گروهی توپولوژیک و ساده-همبند است و $f : U \to H$ یک همریختی از مجموعهٔ باز و همبند $U$ حول $e$ به گروه $H$ است. ثابت کنید تابع $f$ به‌طور یکتا قابل گسترش به $G$ است. به‌عبارت دیگر، همریختی یکتای $\psi$ از $G$ به $H$ وجود دارد که $f|_U = \psi|_U$.

۲. فرض کنید $L_1, L_2, L_3, L_4 \subseteq \mathbb{P}^3$، ۴ خط دوبه‌دو متنافر باشند. ثابت کنید اگر بیش از ۲ خط وجود داشته باشند که هر چهار خط را قطع کنند، آنگاه بینهایت خط با این خاصیت وجود دارند.

۳. ثابت کنید که اگر $f : \mathbb{C} \to \mathbb{C}$ تابعی تحلیلی باشد که همه جا ناصفر است، و عدد طبیعی $d$ وجود داشته باشد که به‌ازای هر $z \in \mathbb{C}$، $|f(z)| \le e^{|z|^d}$، آنگاه $f(z) = e^{Q(z)}$ که $Q$ یک چندجمله‌ای با ضریب در $\mathbb{C}$ است.

۴. فرض کنید $f : \mathbb{R}^2 \to \mathbb{R}$ تابعی باشد به قسمی که تحدید $f$ به هر خط $x = c$، $y = d$ پیوسته باشد. ثابت کنید نقاط پیوستگی $f$ در $\mathbb{R}^2$ چگال هستند.

۵. آیا می‌توان برای چندجمله‌ای‌های همگن درجهٔ $d$ی ۳ متغیری $(d \ge 2)$ ریشهٔ گویا برحسب ضرایب پیدا کرد (به‌ازای $d = 1$ این کار به‌سادگی امکان‌پذیر است زیرا $(1, 1, \frac{-a-b}{2})$ ریشهٔ $ax + by + cz = 0$ است و تابعی گویا برحسب ضرایب این چندجمله‌ای است.)

## مراجع


1. Imayoshi, Y. and M. Taniguchi, *An Introduction to Teichmüller Spaces*, Springer (1992).
2. Oxtoby, J. C. *Measure and Category*, Springer (1971).
3. Spanier, E. *Algebraic Topology*, Mc Graw-Hill (1966).
4. Titchmarsh, *The Theory of Functions* (2nd. ed.), Oxford University Press, reprint 1976.


* * * * * *


* مریم میرزاخانی، بخش ریاضی دانشگاه هاروارد، آمریکا

mirzak@math.harvard.edu


1. Egorov

# کتاب

## ریاضیات چیست؟

سیاوش شهشهانی *

**ریاضیات چیست؟ ریچارد کورانت، هربرت رابینز، ویراست دوم (به اهتمام یان استیوارت). ترجمهٔ سیامک کاظمی. نشر نی. تهران، ۱۳۷۹، بیست و نُه + ۵۹۹ صفحه.**

نام «کورانت» برای بسیاری از دست‌اندرکاران ریاضی، «مؤسسهٔ علوم ریاضی کورانت» را تداعی می‌کند که یکی از معتبرترین مراکز ریاضی جهان به‌خصوص در زمینه‌های آنالیز، معادلات دیفرانسیل، محاسبات عددی و کاربردهای ریاضیات است. مؤسسهٔ کورانت درواقع بخش تحصیلات تکمیلی ریاضیات در دانشگاه نیویورک آمریکاست. ریچارد کورانت، که این مؤسسه به یاد او نامگذاری شده است در سال ۱۹۳۴ جزو اولین موج دانشمندان یهودی تباری بود که از آلمان به امریکا مهاجرت کردند. وی تقریباً تمامی سی‌وهشت سال باقیماندهٔ عمرش را در خدمت این مؤسسه گذراند. او شاگرد سابق هیلبرت و وارث سنت ریاضی پربار گوتینگن، قطب ریاضی جهان در پیش از جنگ جهانی دوم، بود. کوششهای بی‌وقفهٔ کورانت، بخش ریاضی دانشگاه نیویورک را از یک بخش ریاضی متوسط به یک مرکز جهانی معتبر ارتقا داد. کورانت از اولین سالهای اشتغالش در دانشگاه نیویورک، سخنرانیها و درسهایی عمومی برای غیرمتخصصان ارائه می‌کرد که هدف آن بالابردن فرهنگ عمومی ریاضیات بود. وی تصمیم گرفت بر مبنای یادداشتهای دانشجویان، کتابی از این درسها تهیه کند که نهایتاً چاپ اول آن در سال ۱۹۴۱ با همکاری هربرت رابینز با عنوان «ریاضیات چیست؟» انتشار یافت. همکاری رابینز در چاپ این کتاب داستان پرماجرایی است که بعداً به آن اشاره‌ای مختصر خواهیم کرد.

در پیشگفتار چاپ اول، ابراز تأسف ریچارد کورانت را می‌خوانیم که می‌گوید جایگاه دو هزار سالهٔ ریاضیات در فرهنگ بشری به دلایلی که یکی از آنها بی‌اعتنایی دست‌اندرکاران علم ریاضی است به‌خطر افتاده است و یک هدف او از تهیهٔ این کتاب، کوشش برای قابل فهم ساختن اندیشه‌ها و روشهای ریاضیات برای عامهٔ فرهیخته است. او معتقد است که هر چند کتب مهیج تاریخی و زندگینامه‌ها تا حدی جای خالی ریاضیات را در فرهنگ عمومی پر کرده‌اند، لیکن هدف او فهماندن واقعیت ریاضیات است به‌نحوی که محتوا و مفاهیم فدا نشوند و در عین حال جزئیات و ریزه‌کاری‌های


فنی موجب ملال خاطر خواننده و دافع ذوق و شوق او نگردند. مطالبی که کورانت به‌این منظور برای درج در کتاب انتخاب کرده است از میان مباحثی برگزیده شده‌اند که نه فقط قابل دسترسی و جذاب‌اند، بلکه پُلی میان ریاضیات دبیرستانی و بدنهٔ اصلی ریاضیات فعال روز را، به‌زعم کورانت، تشکیل می‌دهند. بدین‌ترتیب هشت فصل کتاب را موضوعهای متعارفی تشکیل می‌دهند که عبارت‌اند از (۱) عددهای طبیعی، (۲) دستگاه اعداد در ریاضیات، (۳) ترسیمهای هندسی. جبر هیأتهای اعداد، (۴) هندسهٔ تصویری. رهیافت اصل موضوعی. هندسه‌های نااقلیدسی، (۵) توپولوژی، (۶) تابع و حد، (۷) ماکسیمم و مینیمم، و (۸) حساب دیفرانسیل و انتگرال. تلاش برای عامه فهم کردن ریاضیات و علوم در میان اسلاف کورانت در گوتینگن از زمانی‌که فلیکس کلاین عملاً رهبری جامعهٔ علمی آلمان را به عهده گرفت رواج یافت. کلاین که در اثر فعالیتهای اجتماعی‌اش از نیاز علوم محض به حمایت عمومی آگاهی داشت و سخنران و نویسندهٔ چیره‌دستی بود، سخنرانیها و نوشتگان متعددی با این هدف عرضه کرد که بعضی از این نوشته‌ها با گذشت بیش از یک قرن هنوز تجدید چاپ می‌شوند. هیلبرت، جانشین بزرگ کلاین در مؤسسهٔ ریاضیات گوتینگن، شاید بدیعترین و ماندنی‌ترین نمونهٔ این نوع آثار را به‌نام *هندسهٔ قابل تجسم* [۲] در سال ۱۹۳۲ به چاپ رساند، کتابی که به چند ده زبان ترجمه شده و هنوز منشأ الهام و مملو از نکات بدیع حتی برای متخصصان است. اثر جذاب و معروف دیگری، این نیز مربوط به دههٔ ۱۹۳۰ در گوتینگن، مجموعه‌ای است که هانس رادِماخر

و اتو توپلیتس فراهم آوردند و اکنون با پاره‌ای تغییرات زیرعنوان انگلیسی «*The Enjoyment of Mathematics*» در دسترس است [۳]. روند هیلبرت در کتاب خود تأکید بر محتوای شهودی و بصری هندسه به مفهوم عام آن است. مجموعهٔ رادِماخر و توپلیتس درواقع جُنگی از مطالب معروف، جذاب و شوق‌آفرین است که با معلومات اندک قابل دسترسی است. در مقابل، کورانت قصد داشت که در کتاب خود، علاوه بر پیگیری هدف اشاعهٔ ریاضیات، نوعی جهان‌بینی نسبت به ریاضیات و جایگاه آن درعلم و فرهنگ را، که از درون و بیرون در مخاطره می‌دید، به خواننده القا کند. اینکه تا چه حد او در این راه موفق بوده است قابل بحث است. گفته می‌شود خود او این اثر را ناقص تلقی می‌کرده هر چند که این کتاب با فروش چند صد هزار نسخه به زبانهای مختلف، یکی از موفقترین آثار چاپ‌شدهٔ تاریخ ریاضیات به‌شمار می‌آید. در هر حال می‌توان شمه‌ای از دیدگاه کورانت نسبت به ریاضیات را در چند صفحه پیش‌درآمد کتاب با عنوان «ریاضیات چیست؟» (صفحات بیست‌وپنج تا بیست‌ونه ترجمهٔ حاضر) مشاهده کرد. در این چند صفحه دودلی و شاید سردرگمی فلسفی حاکم بر گوتینگنِ آن زمان مشهود است. از یک طرف مکتب ریاضی گوتینگن مکتب هیلبرتِ است و هیلبرت بنیانگذار مکتب فلسفی صورتگرایی. از طرفی دیگر، عملکرد عملی هیلبرت و گوتینگن کاملاً به دور از افراطهای جزمی صورتگرایی و وفادار به ریاضیات سنتی و کاربردهای آن است. کورانت، مانند هیلبرت، تناقضی میان این «شیٔ‌زدایی» از ریاضیات از یک سو و اتکا به بدنهٔ تاریخی ریاضیات از سوی دیگر قائل نیست، و در عین حال از پیگیری بحث فلسفی پیرامون ماهیت ریاضیات سرباز می‌زند. پیش‌درآمد با این عبارات به‌پایان می‌رسد که « ... نه فلسفه، بلکه تجربهٔ فعال در خود ریاضیات است که می‌تواند به این پرسش پاسخ دهد: ریاضیات چیست؟». در این راستا کتاب حاضر بی‌شک اثر بسیار موفقی بوده است.

اما در مورد عنوان «ریاضیات چیست؟» نقل‌قولی از زندگینامهٔ کورانت [۴] به قلم کنستانس رید خالی از لطف نیست. در یک میهمانی در منزل هرمان وایل که نویسندهٔ معروف آلمانی توماس مان[1] هم در آن حضور داشته، کورانت از مان در مورد عنوان کتاب نظرخواهی می‌کند به این صورت که عنوان «ریاضیات چیست؟» برای کتاب «کمی غیرصادقانه» است و عنوان دقیقتری برای آن شاید «بحثهای ریاضی پیرامون مسائل ابتدایی برای عامه» باشد. توماس مان جواب می‌دهد که در این موردنظری ندارد ولی یک تجربهٔ شخصی خودش را برای کورانت بازگو می‌کند که وقتی ترجمهٔ انگلیسی رمان او با عنوان لوته در وایمار[2] برای چاپ آماده می‌شده ناشر پیشنهاد می‌کند آن‌را بازگشت معشوقه[3] بنامند. مان می‌خواهد به این تغییر نام رضایت ندهد ولی ناشر می‌گوید «اگر لوته در وایمار باشد ده هزار نسخه، شاید هم بیست هزار، به فروش می‌رسد، ولی از بازگشت معشوقه صد هزار نسخه خواهیم فروخت. البته حق‌التألیف متناسب با فروش است». مان به بازگشت معشوقه رضایت می‌دهد. می‌گویند کورانت بلافاصله به چاپخانه تلفن می‌زند و نام «ریاضیات چیست؟» ابقا می‌شود.

از زمان ظهور اولین چاپ این کتاب نقش هربرت رابینز در تهیهٔ آن موضوع شایعات و نقل‌قولهای گوناگونی بوده است. مجدداً زندگینامهٔ کورانت [۴] را به مدد می‌گیریم. کنستانس رید ظاهراً موفق شد پس از مصاحبه‌هایی با کورانت، رابینز و دیگران تصویر دقیقی از این ماجرا ترسیم کند. در سال ۱۹۳۹ کورانت توپولوژیدان جوانی به‌نام هربرت رابینز را که چندی پیش از هاروارد دکترای ریاضی خود را گرفته بود، با این انگیزه در دانشگاه نیویورک استخدام می‌کند که علاوه بر انجام وظایف عادی تدریس، او را در تکمیل کتاب «ریاضیات چیست؟» یاری دهد، رابینز که نمی‌خواست وقت محدودی را که برای تحقیق داشت به‌ازای دستمزد اندکی برای همکاری در نوشتن یک کتاب توصیفی صرف کند دچار دودلی می‌شود ولی بالاخره در مقابل این قول ضمنی که نام او همراه با نام کورانت به‌عنوان مؤلف کتاب ذکر خواهد شد به این کار تن درمی‌دهد. پس از دو سال کوشش و همکاری، رابینز کتاب را بازنویسی می‌کند و بخشهایی مانند فصل توپولوژی را به‌آن می‌افزاید ولی در نمونه‌خوانی درمی‌یابد که قرار است فقط نام کورانت به‌عنوان مؤلف ذکر شود. رابینز پس از مشورت با استاد ذی‌نفوذ رسالهٔ خود ویتنی[1] که قول می‌دهد در صورت محقق شدن این امر، موضوع اخراج کورانت از انجمن ریاضی امریکا را مطرح خواهد ساخت، نامه‌ای مؤدبانه به کورانت می‌نویسد که در آن ضمن اذعان به اینکه کتاب زادهٔ فکر کورانت است خاطرنشان می‌سازد که زحمات دوساله و درگیری فکری و روانی که در تهیهٔ کتاب داشته است او را مستحق سهمی از تألیف کتاب می‌سازد. ظاهراً شایعهٔ تهدید ویتنی و احتمال پیگیری قضایی نیز از محافل دیگر به گوش کورانت می‌رسد و او بالاخره رضایت می‌دهد که نام هر دو به‌عنوان نویسنده ذکر گردد.

پس از استقبال کم‌سابقه‌ای که از چاپ اول این کتاب در سال ۱۹۴۱ صورت گرفت و تجدید چاپهای مکرر با تغییرات جزئی در طول حیات کورانت، ویراست دوم کتاب به اهتمام ریاضیدان انگلیسی یان استوارت در سال ۱۹۹۶ توسط انتشارات دانشگاه آکسفرد به چاپ رسید. در این ویراست، تغییر اصلی افزودن یک فصل جدید (فصل ۹) به‌عنوان «پیشرفتهای جدید» است که در آن یان استوارت عمدتاً به توصیف مختصر تعدادی از مباحث که در چاپهای پیشین به‌عنوان مسائل حل‌نشده مطرح شده بودند ولی در سالهای اخیر به جواب رسیده‌اند یا گشایش عمده‌ای در بررسی آنها حاصل شده است پرداخته است. روح اصلی کتاب همچنان دست‌نخورده، و همچنان زنده و شاداب باقی مانده است.

ویراست دوم موقعیت مغتنمی بود برای یک ترجمهٔ جدید کتاب به فارسی که خوشبختانه مترجم حاضر با چیره‌دستی از عهدهٔ آن برآمده است. صورت فارسی پیشین این کتاب [۱] که در سال ۱۳۴۹ منتشر شده بود قطعاً نیاز به بازنگری اساسی داشت. در این سی سال فرهنگ ریاضیات جدید درکشور پا گرفته و از جمله، واژه‌گزینی ریاضیات در حدی پیشرفت کرده است که مترجم ویراست جدید هیچ‌گونه مشکلی برای یافتن مترادفهای فارسی نداشته است، و خوشبختانه با احتراز از افاضات و اضافات بیمورد، موفق شده است متنی خوشخوان، روان، دقیق که مهمتر از همه حافظ روح اصلی کتاب است به زبان فارسی ارائه کند.

این نگارنده حدود ۳۰ صفحهٔ متن اصلی کتاب را به‌طور پراکنده، و پیشگفتارها را به‌طور کامل، دقیقاً مورد بررسی قرار داده و مشکل عمده‌ای در ترجمهٔ آن نیافته است. فقط ذکر چند نکته را که بعضی صرفاً سلیقه‌ای هستند در اینجا لازم می‌داند.

1. Thomas Mann 2. Lotte in Weimar

3. The Beloved Returns

1. Hassler Whitney

۱. در «پیشگفتار ویراست اول»، صفحهٔ دوازده، سطرهای ۵ و ۹، کلمهٔ «مؤلفان» به‌کار رفته که با توجه به اینکه فقط نام یک مؤلف در پایان پیشگفتار ظاهر می‌شود عجیب می‌نماید. مراجعه به متن انگلیسی نشان می‌دهد که در اصل، در هر دو مورد فاعل مجهول به‌کار رفته است. این انتخاب توسط کورانت کاملاً عمدی بوده است زیرا همچنان که در بالا توضیح داده شد از یک سو او خود را تنها مؤلف واقعی کتاب قلمداد می‌کرد و از سوی دیگر به میل یا اجبار نمی‌خواست موجب کدورت خاطر رابینز شود. به این ترتیب، بهتر است در ترجمه نیز فاعل مجهول حفظ شود.

۲. در پیش‌درآمد «ریاضیات چیست؟»، صفحهٔ بیست و نه، سطر چهارم اصطلاح «واقعیت‌زدایی» در مقابل dissubstantiation آمده است که به نظر نگارنده گمراه‌کننده است. «واقعیت‌زدایی» احتمالاً به گوش اکثر خوانندگان زدودن واقعیت را تداعی می‌کند درحالی‌که در اینجا مراد نوعی زدودن شیئیت فیزیکی از مفاهیم ریاضی است. اصطلاح «شیء‌زدایی» نزدیکتر به معنای مورد نظر به نظر می‌رسد.

۳. در مورد واژگان ریاضی، مترجم از مجموعهٔ تدوین‌شده توسط مرکز نشر دانشگاهی و انجمن ریاضی ایران استفاده کرده است که در اکثریت نزدیک به اتفاق موارد مقبول جامعهٔ ریاضی کشور است. در یکی دو مورد اصطلاحات دیگری امروزه تفوق پیدا کرده‌اند. مثلاً امروز معمولاً کلمهٔ «میدان» به جای «هیأت» به کار گرفته می‌شود که اولی ترجمه از انگلیسی و دومی ترجمه از فرانسه و آلمانی است. لغت «تلاشی» برای decay این اشکال را دارد که با افزودن «ی» اضافه اشکال تلفظی ایجاد می‌کند (مثلاً «تلاشیِ نمایی») که در کتاب ظاهر می‌شود). در این مورد لغت «زوال» قابل بررسی است.

۴. مترجم سعی کرده است برای اجتناب از آوردن مخفف اسامی کوچک، که در فارسی گاهی مشکل‌دار است، اسامی را به‌طور کامل بیاورد. در حداقل یک مورد یعنی G. D. Birkhoff، این کوشش منجر به اشتباه شده است. مترجم این نام را «گِرت بیرکاف» نقل کرده است (ص ۲۷۲) درحالی‌که مقصود G. D. Birkhoff یعنی «جرج دیوید برکُف» است. ریاضیدان مورد نظر در متن اصلی جرج دیوید (پدر) است نه گرت (فرزند). ضمناً «برکُف» نزدیکتر از «بیرکاف» به تلفظ اصلی است.

۵. در منابع آخر کتاب بهتر بود مواردی که کتابْ ترجمهٔ فارسی دارد موضوع برای خوانندهٔ فارسی زبان تصریح می‌شد. نگارنده به وجود حداقل یک مورد (کتاب مفاهیم نوین ریاضیات اثر یان استوارت) وقوف دارد.

این نکات جزئی، قابل اصلاح، و بعضاً قابل مناقشه، چیزی از ارزش ترجمهٔ ارزشمند این کتاب کلاسیک نمی‌کاهد. امیدواریم به زودی شاهد ترجمه‌های فارسی همترازی از آثار کلاسیک ذکرشدهٔ [۲] و [۳] نیز باشیم.

## مراجع

۱. ریاضیات چیست؟ نوشتهٔ ریچارد کورانت و هربرت رابینز، ترجمهٔ حسن صفاری، انتشارات خوارزمی، ۱۳۴۹.

1. 2. D. Hilbert & S. Cohn-Vossen, *Geometry and Imagination*, Chelsea, New York (1952).

[این کتابْ ترجمه از متن آلمانی با مشخصات

*Anschauliche Geometrie*, Julius Springer, Berlin (1932)

است.]

2. 3. H. Rademacher & O. Toeplitz, *The Enjoyment of Mathematics*, Dover, New York (1990).

3. 4. C. Reid, *Courant in Göttingen and New York*, Springer-Verlag, New York (1976).

* * * * * *

* سیاوش شهشهانی، دانشگاه صنعتی شریف و پژوهشگاه دانشهای بنیادی

من گمان می‌کنم حقیقت تقریباً از این قرار است ــ حقیقت آنقدر پیچیده است که جز به تقریب نمی‌شود از آن سخن گفت ــ که ایده‌های ریاضی از تجربه منشأ می‌گیرند، هر چند شجره‌نامهٔ آنها گاهی دور و دراز و تیره و تار است. ولی وقتی ایده‌ها از این طریق پدید آمدند، موضوعِ متشکل از آنها زندگی خاصی در پیش می‌گیرد و بیشتر شبیه یک هنر خلاقه می‌شود، که تقریباً به‌طورکامل تحت تأثیر انگیزه‌های زیبایی‌شناختی است، تا شبیه هر چیز دیگر، به‌خصوص یک علم تجربی. ولی نکتهٔ دیگری هست که به نظر من باید به آن توجه داشت. وقتی یک مبحث ریاضی از سرچشمهٔ تجربیش دور شده یا نسل دوم و سومی از مبحث اولیه باشد که فقط به‌طور غیرمستقیم از ایده‌های برخاسته از «واقعیت» الهام گرفته است، در معرض خطرات مهیبی قرار دارد. چنین مبحثی هر چه بیشتر تابع زیبایی‌شناسیِ صرف و «هنر برای هنر» می‌شود که الزاماً بد نیست به شرط اینکه موضوع موردنظر در احاطهٔ موضوعات وابسته‌ای باشد که هنوز ارتباطات نزدیکی با تجربه داشته باشند یا تحت سیطرهٔ اشخاصی بسیار باذوق باشد. اما خطر در این است که موضوع در راستای کمترین مقاومت رشد کند، جریان که از سرچشمه‌اش بسیار دور شده به چندین شاخهٔ بی‌اهمیت تجزیه شود، و به تودهٔ نابسامانی از جزئیات و مطالب پیچیده بدل گردد. به عبارت دیگر، مبحث ریاضی در فاصلهٔ زیاد از خاستگاه تجربیش یا پس از مقدار زیادی تنفس «انتزاعی»، در معرض تباهیدگی قرار می‌گیرد. . . .

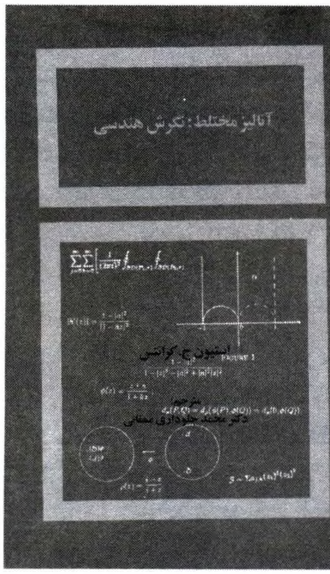
به هر صورت، هرگاه این مرحله فرا رسید، تنها راه چاره‌ای که به نظر می‌رسد، «جوان‌سازی» موضوع از طریق بازگشت به منشأ است یعنی تزریق ایده‌هایی که کمابیش به‌طور مستقیم از تجربه گرفته شده باشند. من متقاعد شده‌ام که این کار شرط لازم برای حفظ تر و تازگی و سرزندگی موضوعات ریاضی بوده است و در آینده نیز خواهد بود.

* * * * * *

مطلب بالا بخشی از مقالهٔ «ریاضیدان» اثر فون نویمان است.

# آنالیز مختلط: نگرش هندسی

سعید ذاکری*




**آنالیز مختلط: نگرش هندسی. استیون ج. کرانتس. ترجمهٔ محمد جلوداری ممقانی. انتشارات علمی و فرهنگی. تهران، ۱۳۷۹. ۲۱۸ صفحه.**

سخن طعن‌آمیزی را از سالهای اول دوران دانشجویی به خاطر می‌آورم که می‌گفتند دوران آنالیز مختلط دیگر به سر آمده است. در این ارزیابی شتابزده اندک مایه‌ای از حقیقت نیز وجود دارد؛ درواقع آنچه که اغلب غیرمتخصصان تحت عنوان کلی «آنالیز مختلط» می‌شناسند، میراث مکتبی است که سالها پیش از جنگ دوم جهانی به اعلا درجهٔ شکوفایی خود رسیده بود. این همان مکتب عمدتاً آلمانی آنالیز مختلط کلاسیک است که بیش از یک قرن سایهٔ سنگین خود را بر پیکرهٔ ریاضیات انداخته بود و از نگاشتهای همدیس تا نظریهٔ تحلیلی اعداد، از توابع بیضوی تا خمهای جبری، از نظریهٔ پتانسیل تا معادلات پاره‌ای و توابع خاص فیزیک ریاضی، همگی مظاهر سیطرهٔ آن به‌شمار می‌رفتند. نتایج منتشرشده در این دوران چنان گسترده، عمیق و کامل‌اند که نوآوری در این زمینه را، اگر نه ناممکن، بسیار دشوار می‌کنند. نگاهی گذرا به هر یک از کتابهای آنالیز مختلط یا توابع خاص پیش از جنگ کافی است تا خواننده را به شگفتی وا دارد. بسیاری از مطالب این کتابها امروزه جایی در فرهنگ عمومی ریاضی ندارند. نسل ریاضیدانان زندهٔ مسلط به برخی از این فنون نیز در خطر انقراض است. به این تعبیر، شاید دوران آنالیز مختلط واقعاً به سرآمده باشد.

اما از دهه‌های پایانی قرن نوزدهم، آنالیز مختلط کلاسیک جریان ماهیتاً متفاوتی را نیز در بطن خود پرورد که سردمداران آن وایرشتراس، هرویتس[1]، یا میتاگ‌لفلر[2] نبودند. خصیصهٔ این جریان، نگرش هندسی به نظریهٔ توابع تحلیلی بود که نمونه‌های آغازین آن را می‌توان در آثار ریمان سراغ گرفت، اما متولیان اصلیش فوکس[3]، شوارتس، کلاین، پوانکاره، پیکار[4]، و بعدها در اوایل قرن بیستم کاراتئودوری، مونتل، و کوبه[5] بودند. این جریان نه تنها هنوز به موزهٔ تاریخ ریاضیات نپیوسته، بلکه از راههای گوناگون و گاه شگفت‌آوری

1. Hurwitz 2. Mittag-Leffler 3. Fuchs 4. Picard 5. Koebe

به ریاضیات امروز سرایت کرده است. یافتن مثالهایی در تأیید این ادعا چندان دشوار نیست: تحقیقات هندسی فوکس، کلاین و پوانکاره دربارهٔ زیرگروههای گسستهٔ گروه خودریختیهای تحلیلی کرهٔ ریمان، مآلاً به نظریه‌ای مبدل شد که امروزه گروههای کلاینی خوانده می‌شود و در بررسی توپولوژی و هندسهٔ خمینه‌های هذلولوی، به‌ویژه برنامهٔ ترستن[1] برای رده‌بندی خمینه‌های سه‌بعدی، نقشی مرکزی دارد. صورتهای تعمیم‌یافته و کمّی قضایای پیکار که نوانلینا[2] به اثبات رساند، با رهیافت هندسی آلفرس به پیدایش نظریهٔ نوین توزیع مقادیر منجر شد و بعدها حتی همتایی شگفت‌انگیز در نظریهٔ اعداد (مشخصاً هندسهٔ دیوفانتی) پیدا کرد. محک مونتل برای نرمال بودن را فاتو و ژولیا در تکرار توابع گویا به‌نحو مؤثری به‌کار بستند، که بعدها نظریهٔ دینامیک مختلط را آفرید. نگرش هندسی کاراتئودوری و کوبه به توابع تک‌ارز[3]، گروچ[4] و تایشمولر[5] را به مطالعهٔ نگاشتهای شبه‌همدیس هدایت کرد که علاوه بر اهمیت ذاتیشان در هندسه و آنالیز، ابزاری اساسی در دینامیک همدیس به مفهوم عام به‌شمار می‌آیند.

سهم شوارتس در این میان تنها به «اصل انعکاس شوارتس»، «تبدیل شوارتس-کریستفل»، یا «مشتق شوارتسی»، که همه در دروس مقدماتی آنالیز مختلط می‌آموزند، خلاصه نمی‌شود. او که بینشی عمیقاً هندسی داشت، و آنالیز سخت را نیز از استادش وایرشتراس آموخته بود، پس از به‌دست آوردن

1. Thurston 2. Nevanlinna 3. univalent 4. Grötzsch
5. Teichmüller

نتایج مهمی در بارهٔ رویه‌های مینیمال به آنالیز مختلط روی آورد و در این قلمرو بود که آثار درخشانی از خود به‌جا گذاشت. مثلاً روش او بود که اولین اثبات درست قضیهٔ نگاشت ریمان را، در پی انتقاد ویرانگر وایرشتراس از اثبات اولیهٔ ریمان، به‌دست داد. همچنین، شوارتس نخستین کسی بود که رفتار مرزی نگاشتهای همدیس را، که سالها بعد به تحقیقات اساسی کاراتئودوری انجامید، بررسی کرد. روش هندسی و سادهٔ او برای ساختن توابع خودریخت بعدها به دست کلاین و پوانکاره توسعهٔ بسیار یافت. به‌خصوص اثبات او برای وجود تابع پیمانه‌ای بیضوی، که پیش از آن تنها با استفاده از نظریهٔ پیچیدهٔ انتگرالهای بیضوی میسر بود، رهیافت به مراتب مقدماتیتری به قضایای پیکار و مونتل را ممکن ساخت.

اما مشاهدهٔ تقریباً بدیهی شوارتس در ۱۸۶۹، که این روزها در کتابهای درسی به «لم شوارتس» موسوم است، احتمالاً بیش از هر نتیجهٔ دیگر نام او را در تاریخ زنده نگه داشته است. (این حقیقت وقتی طنزآمیزتر می‌شود که بدانیم صورت امروزی این لم درواقع از آنِ کاراتئودوری است!) لم شوارتس، که نتیجهٔ سادهٔ اصل ماکسیمم است، ادعا می‌کند که اگر $f$ تابعی تحلیلی از قرص واحد به خودش باشد که مبدأ را ثابت نگه می‌دارد، آنگاه $|f'(0)| \leq 1$، و تساوی تنها وقتی برقرار است که $f$ یک دوران باشد. شوارتس این حکم را گام کوچکی در اثبات قضیهٔ نگاشت ریمان می‌دید، و احتمالاً از عمق واقعی آن خبر نداشت. در ۱۹۱۶ پیک[1] دریافت که لم شوارتس تعبیر جالبی برحسب متریک هذلولوی روی قرص واحد $\mathbb{D}$ دارد. نقطهٔ $z$ در $\mathbb{D}$ و بردار مماس $v$ در $z$ را در نظر بگیرید. $v$ را با مشتق خودریختی تحلیلی $\varphi : \mathbb{D} \to \mathbb{D}$ که $z$ را به مبدأ می‌نگارد به بردار مماس $w$ در مبدأ منتقل کنید. در این صورت طول هذلولوی $v$ برابر با طول اقلیدسی $w$ تعریف می‌شود. این به خودریختی خاص $\varphi$ بستگی ندارد، چرا که بنا به لم شوارتس هر خودریختی دیگر با این خاصیت، ترکیب $\varphi$ با یک دوران حول مبدأ است. محاسبهٔ ساده‌ای بر مبنای این تعریف نشان می‌دهد که طول هذلولوی $v$ درواقع برابر است با $|v|/(1-|z|^2)$، و بنابراین متریک هذلولوی $\mathbb{D}$ را می‌توان به‌صورت

$$\rho_{\mathbb{D}}(z) = \frac{|dz|}{1-|z|^2}, \qquad |z| < 1$$

نمایش داد. از تعریف بالا بلافاصله نتیجه می‌شود که $\rho_{\mathbb{D}}$ تحت هر خودریختی تحلیلی $\mathbb{D}$ ناورداست. با استفاده از این مطلب، و نیز اینکه گروه خودریختیهای تحلیلی $\mathbb{D}$ ترایا عمل می‌کند، می‌توان به روایت پیک از لم شوارتس رسید که ماهیت هندسی آن را فاش می‌کند: هر نگاشت تحلیلی $f : \mathbb{D} \to \mathbb{D}$ فاصلهٔ هذلولوی را اکیداً کاهش می‌دهد مگر آنکه یک خودریختی باشد، به بیان معادل، برگردان[2] $\rho_{\mathbb{D}}$ تحت $f$ همه جا در نامساوی

$$f^*\rho_{\mathbb{D}} \leq \rho_{\mathbb{D}}$$

صدق می‌کند، و تساوی در نقطه‌ای برقرار است اگر و تنها اگر $f$ یک خودریختی باشد. این واقعیت بلافاصله به چارچوب رویه‌های ریمانی گسترش می‌یابد. درواقع از قضیهٔ یکنواخت‌سازی نتیجه می‌شود که هر رویهٔ ریمانی هذلولوی $X$، یعنی رویه‌ای که پوشش اکملش قرص واحد است، مجهز به یک متریک هذلولوی $\rho_X$ است که خمیدگی گاوسی آن در هر نقطه $-4$ است. در این صورت لم شوارتس-پیک می‌گوید که برای هر نگاشت تحلیلی $f : X \to Y$ بین رویه‌های ریمانی هذلولوی، نامساوی $f^*\rho_Y \leq \rho_X$ همه جا برقرار است، و تساوی در نقطه‌ای رخ می‌دهد اگر و تنها اگر $f$ یک ایزومتری موضعی (معادلاً یک نگاشت پوششی) باشد. این حقیقت ابتدایی، سنگ بنایی در هندسهٔ رویه‌های ریمانی است.

در ۱۹۳۸ آلفرس دریافت که آنچه لم شوارتس-پیک را برقرار می‌کند نه شکل خاص متریک هذلولوی، که منفی‌بودن خمیدگی آن است. او در مقالهٔ کوتاه [1] نشان داد که اگر $\sigma$ متریک همدیس دلخواهی بر رویهٔ ریمانی $X$ باشد که خمیدگی آن همه جا از ثابت منفی $-K$ کمتر است، آنگاه به‌ازای هر نگاشت تحلیلی $f : \mathbb{D} \to X$، نامساوی

$$f^*\sigma \leq \frac{2}{\sqrt{K}}\ \rho_{\mathbb{D}}$$

همه جا برقرار است. از این قضیه به‌سادگی نتیجه می‌شود که اگر $\sigma$ متریک همدیسی بر رویهٔ ریمانی $X$ باشد که خمیدگی آن کران بالای منفی دارد، آنگاه هر نگاشت تحلیلی از صفحه به $X$ لزوماً ثابت است. این تعمیمی از قضیهٔ لیوویل است و در عین حال نشان می‌دهد که وجود چنین متریکی شرط لازم و کافی برای هذلولوی بودن $X$ است.

بیش از ۴۰ سال بعد، آلفرس در یادداشتهایی بر مجموعهٔ آثارش [2] اعتراف می‌کند که در ۱۹۳۸ به عمق قضیه‌اش واقف نبوده و اگر کاربرد این نامساوی در یافتن تخمین بهتری برای ثابت بلاخ[1] نبود، حاضر نمی‌شده نتیجه‌ای چنین «کم‌مایه» را منتشر کند! پیشرفتهای سالهای بعد باید توهم او را از میان برده باشد، چرا که نتیجهٔ «کم‌مایهٔ» او به تنهایی آغازگر جریانی شد که ثمرهٔ آن تبیین ارتباطات عمیق میان آنالیز مختلط از یک سو و قضایای مقایسه در هندسهٔ ریمانی از سوی دیگر بوده است. برای مرور خوبی بر تاریخچهٔ این جریان و بعضی پیشرفتهای اخیر، [5] را ببینید.

ایدهٔ اساسی دیگر در آنالیز مختلط هندسی از اثبات کاراتئودوری-کوبه برای قضیهٔ نگاشت ریمان سرچشمه می‌گیرد (شرح زیبایی از تاریخچهٔ تلاشهای گوناگون برای اثبات این قضیه در [6] آمده است). نکته آن است که بر هر خمینهٔ مختلط می‌توان شبه‌متریکهایی تعریف کرد که تحت یکریختیهای تحلیلی ناوردا باشند. مثلاً نقطهٔ $p$ در خمینهٔ مختلط $X$ و بردار مماس $v$ در $p$ را در نظر بگیرید، و در خانوادهٔ همهٔ نگاشتهای تحلیلی $f : X \to \mathbb{D}$ که $p$ را به مبدأ می‌نگارند، سوپریمم $|df(p)v|$ را محاسبه کنید و آن را با $C_X(p,v)$ نشان دهید. این کمیت، که به مفهومی طول بردار $v$ را اندازه می‌گیرد، شبه‌متریک کاراتئودوری $X$ نامیده می‌شود، زیرا که کاراتئودوری آن را نخستین بار در ۱۹۲۸ برای تعمیم لم شوارتس-پیک به دامنه‌های کراندار در $\mathbb{C}^2$ معرفی کرد. مشابهاً، در خانوادهٔ همهٔ نگاشتهای تحلیلی $f : \mathbb{D} \to X$ که مبدأ را به $p$ می‌نگارند و مشتقشان بردار مماسی در مبدأ چون $u$ را به $v$ می‌فرستد، اینفیمم $|u|$ را محاسبه کنید و آن را با $K_X(p,v)$ نشان دهید. این کمیت شبه‌متریک کوبایاشی $X$ خوانده می‌شود. خاصیت اساسی این شبه‌متریکها آن است که هر دو تحت نگاشتهای تحلیلی کاهش می‌یابند. به‌عبارت دیگر، اگر $f : X \to Y$ نگاشتی تحلیلی باشد، آنگاه

$$C_Y(f(p), Df(p)v) \leq C_X(p,v)$$

1. Pick 2. pull-back

1. Bloch

و

$$K_Y(f(p), Df(p)v) \leq K_X(p, v).$$

به‌ویژه، $C_X$ و $K_X$ هر دو تحت یکریختیهای تحلیلی ناوردا هستند. در ساده‌ترین حالت $X = \mathbb{D}$، اینها همان متریک هذلولوی را به‌دست می‌دهند، اما در حالت کلی این شبه‌متریکها کاملاً متفاوت‌اند. مثلاً اگر $X$ صفحهٔ مختلطی باشد که از آن تعدادی متناهی نقطه حذف‌شده، آنگاه $C_X$ متحداً صفر است. از سوی دیگر می‌توان نشان داد که رویهٔ ریمانی $X$ هذلولوی است اگر و تنها اگر $K_X$ ناتبهگون باشد. از این رو در حالت رویه‌های ریمانی، ناتبهگون بودن $K_X$ را می‌توان به عنوان تعریف هذلولوی بودن گرفت. این دیدگاه، کوبایاشی را در ۱۹۶۷ به مفهوم هذلولوی بودن خمینه‌های با بعد بالاتر هدایت کرد [۴].

بررسی رفتار این شبه‌متریکها در سالهای اخیر به نتایج مهمی در مورد نگاشتهای تحلیلی بین خمینه‌های مختلط و نیز هندسهٔ ذاتی این خمینه‌ها انجامیده، و به یکی از مباحث روز آنالیز هندسی و چندمتغیرهٔ مختلط مبدل شده است. مقاله‌های [۸] و [۳] توصیف خوبی از تاریخچه و پیشرفتهای اخیر این شاخه به‌دست می‌دهند.

* * * * * *

کتاب کوچک استیون کرانتس شرحی است مقدماتی و بسیار خواندنی از این روشهای هندسی در آنالیز مختلط. این اثر، همچون کتابهای دیگر مجموعهٔ تکنگاشتهای کاروس[1]، برای استفادهٔ غیرمتخصصان نوشته شده است. با این حال، خوانندهٔ باتجربه‌تری هم که اغلب این مطالب را از کتابها، مقاله‌ها یا درسنامه‌های پراکنده فراگرفته است، بی‌شک از خواندن این کتاب لذت خواهد برد. اگرچه زبان مناسب برای بیشتر قضایای کتاب زبان رویه‌های ریمانی است، کرانتس با خردمندی بر این وسوسه چیره‌شده و برای حفظ ماهیت مقدماتی کتاب به حالتِ دامنه‌هایِ در صفحه بسنده کرده است. فصل صفر کتاب مروری است کوتاه بر چند قضیهٔ اساسی آنالیز مختلط یک‌متغیره، ازجمله صورت کلاسیک لم شوارتس، محک مونتل، قضیهٔ نگاشت ریمان، و قضایای کوچک و بزرگ پیکار. تأکید اصلی در اینجا بر شگردِ یافتنِ نگاشت ریمانِ یک دامنهٔ ساده‌ـهمبند به‌عنوان جواب یک مسألهٔ غایی است. اهمیت این شگرد در فصل سوم، هنگام تعریف شبه‌متریکهای کاراتئودوری و کوبایاشی، معلوم می‌شود. فصل یک کتاب مدخلی است بر مفاهیم مقدماتی هندسهٔ دیفرانسیل در صفحه. مفهوم (شبه)متریک همدیس روی یک دامنه، طول خم، برگردان یک متریک تحت یک نگاشت و نیز استفاده از نمادهای متغیر مختلط در هندسهٔ دیفرانسیل دوبعدی از آن جمله‌اند. به‌عنوان مثالی اساسی، متریک هذلولوی روی قرص مورد بررسی نسبتاً مفصلی قرارگرفته و در پی آن تعبیر پیک از لم شوارتس آمده است. فصل دوم کتاب با مفهوم خمیدگی گاوسی یک متریک آغاز می‌شود، و سپس تعمیم آلفرس از لم شوارتس‌ـپیک به اثبات می‌رسد. با استفاده از این نتیجه، اثبات جالبی از قضیهٔ کوچک پیکار به‌دست می‌آید. مابقی این فصل را نویسنده به محک مارتی[2] برای نرمال بودن، و کاربرد آن در اثبات قضیهٔ بزرگ پیکار اختصاص داده است.

فصل سوم، اساسیترین و در عین حال فنیترین مطالب کتاب را دربردارد. در آنجا ابتدا شبه‌متریکهای کاراتئودوری و کوبایاشی تعریف و خواص مقدماتی آنها بررسی می‌شود. قلب این فصل قضیه‌ای است که می‌گوید برای هر دامنهٔ متناهیاً همبند در صفحه که هیچ مؤلفهٔ همبندی مکملش تک‌نقطه‌ای نباشد، متریک کاراتئودوری، و در نتیجه کوبایاشی، کامل است. به‌عنوان یک کاربرد، نویسنده این نتیجهٔ کلاسیک را ثابت می‌کند که هر دامنهٔ کراندار با مرز هموار در صفحه که گروه خودریختیهای تحلیلیش فشرده نباشد، به‌طور همدیس هم‌ارز با قرص واحد است. این فصل با اثبات این مطلب اساسی به پایان می‌رسد که برای دامنه‌های واقع در صفحه، شرط ناتبهگون بودن شبه‌متریک کوبایاشی معادل آن است که دامنه متریکی بپذیرد که خمیدگیش کران بالای منفی داشته باشد. فصل چهارم و آخر کتاب، چشم‌انداز کوتاه اما مفیدی است از کاربرد این روشها در حالت دو متغیر مختلط. به زعم نویسنده، قضیهٔ پوانکاره در باب هم‌ارز تحلیلی نبودن گوی واحد و دوقرص[1] در $\mathbb{C}^2$، مثالی مناسب برای تشریح این روشهاست. بدین ترتیب، این قضیه نه به روش کلاسیک، که با استفاده از خواص متریک کوبایاشی ثابت شده است. کتاب با ضمیمهٔ کوتاهی در مورد تعریف خمیدگی گاوسی رویه‌ها در $\mathbb{R}^3$ و ارتباط آن با تعریف مجردتر خمیدگی برای متریکهای همدیس دامنه‌های در صفحه پایان می‌یابد.

گزینش مطالب مناسب در این موضوع نسبتاً تخصصی، و توصیف آنها در حد درک دانشجویان دورهٔ کارشناسی، کاری است بس دشوار، و نکتهٔ فوق‌العاده برجسته در مورد کتاب این است که کرانتس با چیره‌دستی از عهدهٔ آن برآمده است. بیان روان، پرهیز از زیاده‌گویی، و دوری از قالب خشک و انعطاف‌ناپذیر کتابهای درسی، از محاسن سبک توصیفی و شسته‌ورفتهٔ کتاب است که نظیرش را این روزها کمتر می‌توان یافت. من در سرتاسر کتاب به هیچ اشکال جدی ریاضی برنخوردم، اما به نظرم در موارد معدودی امکان اصلاحات جزئی هست. مثلاً در صورت قضیه‌ای که می‌گوید مجموعهٔ صفرهای یک تابع تحلیلی گسسته است (فصل ۰ بخش ۱ قضیهٔ ۴) باید فرض می‌شد که تابع متحداً صفر نباشد؛ یا در صورت قضیهٔ نگاشت ریمان (فصل ۰ بخش ۳ قضیهٔ ۶) به‌جای این فرض که دامنه با قرص واحد همسانریخت است، بهتر بود فرض می‌شد که ساده‌ـهمبند است. برگردان یک متریک $\rho$ تحت یک نگاشت هموار $f$ با صفرهای منزوی (فصل ۱ بخش ۳ تعریف ۱) چنین تعریف شده است:

$$f^*\rho(z) = \rho(f(z)) \cdot \left|\frac{\partial f}{\partial z}\right|$$

این تعریف تنها وقتی درست است که $f$ تحلیلی باشد (برگردان یک متریک همدیس تحت یک نگاشت هموار غیرتحلیلی دیگر همدیس نیست). از آنجا که کتاب صرفاً با متریکهای همدیس و نگاشتهای تحلیلی سروکار دارد، درست آن بود که نویسنده به جای اصرار بر توجیه تعریف بالا، آن را به نگاشتهای تحلیلی منحصر می‌کرد. اثبات این قضیه که هر دامنهٔ کراندار با مرز هموار و گروه خودریختیهای غیرفشرده، به‌طور همدیس هم‌ارز با قرص است (فصل ۳ بخش ۴ قضیهٔ ۳) می‌توانست به کمک قضیهٔ هرویتس (فصل ۰ بخش ۱

1. Carus 2. Marty

1. bi-disk

قضیهٔ ۹) کوتاهتر و سرراست‌تر شود. با این حال، هیچ‌یک از این موارد جزئی ذره‌ای از قدر کتاب نمی‌کاهد.

******

انتشار ترجمهٔ فارسی کتاب کرانتس را بی‌تردید باید به فال نیک گرفت. تلاش مترجم محترم و همت ناشر در عرضهٔ کتابی که به هر حال چشم‌انداز مالی چندانی ندارد، قابل تقدیر و ستودنی است. با این حال، هنگام مطالعهٔ نسخهٔ فارسی به کاستیهایی برخوردم که از نظر امانت ریاضی یادآوری آنها را بجا می‌دانم. برای سادگی، این نکات به ۴ بخش تقسیم شده و جداگانه مورد نقد قرار گرفته‌اند.

۱. ترجمهٔ نادرست یا مغایر با متن اصلی. اغلب این موارد از نظر ریاضی اشتباه برانگیزند؛ در زیر نمونه‌هایی می‌آورم:

• صفحهٔ ۷۷ سطر ۳: «نکتهٔ اصلی این مثال ساده آن است که برای این که تابع تامی ثابت باشد به غیر از کراندار بودن باید یک پاره خط را از مقادیرش حذف کند.» این به‌وضوح ترجمهٔ نادرستی است. دیدن متن اصلی نشان می‌دهد که مقصود نویسنده چه بوده است: «نکتهٔ این مثال ساده آن است که نبودن پاره‌خطی در مجموعهٔ مقادیر یک تابع تام به‌تنهایی کافی است تا ثابت بودنش را تضمین کند، چه رسد به شرط کرانداری.» در ادامهٔ ترجمهٔ این پاراگراف می‌خوانیم: «پیکار اندازهٔ این پاره‌خط را که یک تابع تام غیرثابت می‌تواند از مقادیرش حذف کند مورد بررسی قرار داده است.» این با ادعای جملهٔ پیش در تناقض آشکار است. متن اصلی درواقع می‌گوید: «پیکار این سؤال را مورد بررسی قرار داد که مکمل تصویر یک تابع تام غیرثابت چقدر بزرگ می‌تواند باشد.»

• صفحهٔ ۸۷ قضیهٔ ۵: «فرض کنید ... خانواده‌ای از توابع $\hat{\mathbb{C}}$ـمقدار روی $U$ (معادلاً، توابع هولومرفیک) باشد.» ترجمهٔ درست چنین است: «فرض کنید ... خانواده‌ای از توابع هولومرفیک $\hat{\mathbb{C}}$ـمقدار (معادلاً، توابع مرومرفیک) روی $U$ باشد.»

• صفحهٔ ۱۲۴ سطر ۵−: «اگر $\partial U$ دارای $k$ مؤلفه باشد آنگاه با به کار بردن $k$ بار متوالی قضیهٔ نگاشت ریمان، نگاشت همدیسی از $U$ به قرص واحد که از آن $(k-1)$ زیر قرص مجزا حذف شده است، به دست می‌آید.» چنین ادعایی در متن اصلی نشده، و درواقع از قضیهٔ نگاشت ریمان هم نتیجه نمی‌شود. آنچه متن اصلی ادعا می‌کند، به دست آمدن نگاشت همدیسی است از $U$ به قرص واحد که از آن $k-1$ دامنهٔ ساده‌ـهمبند با مرز $C^2$ (درواقع تحلیلی حقیقی) حذف شده‌اند. ضمناً برای اثبات این ادعا نویسنده به کتاب آلفرس ارجاع می‌دهد، که در ترجمه حذف شده است.

• صفحهٔ ۱۳۲ سطر ۴−: « ... اگر تعداد مؤلفه‌های همبند دامنهٔ مسطحی متناهی و بزرگتر از ۲ باشد آنگاه گروه خودریختیهای آن متناهی است.» این به‌وضوح اشتباه است. نویسنده در اینجا از عبارت connectivity استفاده می‌کند که نه تعداد مؤلفه‌های همبندی خود دامنه، بلکه تعداد مؤلفه‌های همبندی مکمل آن است، یعنی یکی بیشتر از عدد بتی دامنه.

• صفحهٔ ۱۵۸ پاراگراف آخر: « ... اگر یک تابع با حداقل سه متغیر مختلط همدیس باشد (به این معنا که در بینهایت کوچک حافظ زاویه‌ها و طولهاست) در این صورت باید خطی باشد (تمرین: این ادعا را ثابت کنید. رک: «DFN»).» این مطلب به کلی نادرست است زیرا از یک سو وابرریختی‌ای که در بینهایت کوچک حافظ طول باشد، درواقع یک ایزومتری است، و از سوی دیگر یک نگاشت همدیس لزوماً خطی نیست. دیدن متن اصلی نشان می‌دهد که در ترجمهٔ این عبارت تا چه اندازه بی‌دقتی شده است:

... if a smooth function of at least three real variables is conformal (in the usual sense of being infinitesimally angle preserving and, up to scaling, length preserving) then that function must be the composition of isometries, dilations, and inversions (exercise: prove this, or see the discussion of Liouville's theorem in [DFN]).

که ترجمه‌اش مثلاً چنین است: « ... اگر یک تابع هموار از دست‌کم سه متغیر حقیقی همدیس باشد (به این معنای متعارف که در بینهایت کوچک حافظ زاویه و، در حد تغییر مقیاس، حافظ طول است)، آنگاه این تابع لزوماً ترکیبی است از ایزومتریها، مضارب همانی، و انعکاسها (تمرین: این ادعا را ثابت کنید، یا بحث قضیهٔ لیوویل در [DFN] را ببینید).»

۲. انتخاب معادلهای نادرست یا ناهماهنگ. مثالهایی از این دست در زیر آمده است:

• «ساختار عالی» برای extra structure، «منتظم» برای taut، «فروریختن» برای shrinking [مثلاً در صفحهٔ ۱۸۴ سطر ۱۰ آمده است: «با استفاده از ... و در صورت لزوم فروریختن $U$ و $M$ میدانهای ... را تعریف می‌کنیم» که بی‌معنی است. منظور نویسنده صرفاً این است که در صورت لزوم دامنه‌های $U$ و $M$ را کوچکتر انتخاب می‌کنیم]، «ماتریس ارتفاع» برای attitude matrix که بی‌شک با altitude اشتباه گرفته شده است.

• صفحهٔ ۳۳ سطر ۵−: «در نتیجه می‌توان طرحی از مطلب و تا حدودی روش هندسهٔ دیفرانسیل را بدون درگیرشدن با نمادگذاری و اعمال ماشینی یاد گرفت.» خواننده ممکن است بپرسد که اعمال ماشینی چگونه اعمالی هستند. حقیقت با مراجعه به متن اصلی روشن می‌شود:

The result is that one can learn the flavor and some of the methodology of differential geometry without being encumbered by its notation and machinery.

که ترجمه‌اش مثلاً چنین است: «نتیجه آنکه می‌توان بی‌گرفتارشدن در بند نمادگذاری و ساز و برگ هندسهٔ دیفرانسیل، حال و هوای آن را دریافت و برخی از روشهایش را آموخت.»

• بعضی معادلهای دیگر با آنکه ذاتاً اشتباه نیستند، با سلیقه و دقت کافی انتخاب نشده‌اند. از آن جمله‌اند «تابع مدولار بیضوی» به‌جای «تابع پیمانه‌ای بیضوی» (elliptic modular function)، «همسایگی تیوبی» به‌جای «همسایگی لوله‌ای» (tubular neighborhood)، «مفهوم عاقلانه» به‌جای «مفهوم معقول» (reasonable notion)، «نظر برجسته» به‌جای «بینش

هوشمندانه» (brilliant insight)، «طول کوتاه‌کن» به‌جای «کاهندهٔ فاصله» یا حتی «فاصله کم‌کن» (distance decreasing).

• ضبط برخی اسامی خاص باید مورد بازنگری و دقت بیشتری قرار می‌گرفت. تنها به ذکر سه نمونه بسنده می‌کنم: «کسراتی» به‌جای «کازُراتی» (Casorati)، «پین‌لو» به‌جای «پَنلِوِه» (Painlevé)، «لیندولف» به‌جای «لیندلُف» (Lindelöf).

• مسألهٔ دیگر ناهماهنگی و یک دست نبودن معیار انتخاب معادلهاست، که این یکی قطعاً به گردن ویراستار است. کم نیستند واژه‌هایی حتی فنی که به دو نوع متفاوت ضبط شده باشند. از آن جمله‌اند «معادلات کوشی-ریمان» و «رابطه‌های کوشی-ریمان»، «تایلور» و «تیلور»، «دامنه» و «ناحیه»، «استانده» و «استاندارد»، «تکین» و «منفرد»، «متریک ریمان» و «کرهٔ ریمانی» به‌جای «متریک ریمانی» و «کرهٔ ریمان». همچنین است ناهماهنگی بین ضبط «شوارتس» و مثلاً «هورویتز» یا «استولتز».

۳. نثر فارسی ترجمه و ساختار انگلیسی آن. متأسفانه نثر فارسی کتاب روان و خوشخوان نیست. به‌علاوه در ترجمهٔ کتاب عبارات و تعبیراتی که ساختار کاملاً بیگانه دارند فراوان به چشم می‌خورند، که این برای خوانندهٔ فارسی‌زبان گاه آزاردهنده است. برای روشن‌شدن مطلب، مثالهایی می‌آورم:

• اگر به شما بگویند یک خم «مثبت-جهتدار» و «بدیهی-توپولوژیک» و «پاره-پیوسته-مشتقپذیر» در نظر بگیرید، واکنشتان چه خواهد بود؟ اختصار در ترجمه مهم است، اما زیبایی و کارایی از آن هم مهمتر است. حتی اگر نبودن معادلی بهتر مترجم یا ویراستار را وادار می‌کرد که از مثلاً «تکه‌ای $C^{1}$» استفاده کند، باز بهتر از «پاره-پیوسته-مشتقپذیر» می‌بود.

• در صفحهٔ ۴۷ سطر ۷ می‌خوانیم: «هر شاخهٔ ریاضی ریختارهای خاص خود را دارد: توابعی که ویژگیهای مورد مطالعه را حفظ می‌کنند.» به‌نظرم ساختار انگلیسی این عبارت، آن را از فصاحت و روشنی معنا بی‌بهره ساخته است.

• در صفحهٔ ۱۰۳ سطر ۵- آمده است: «نگاشت $h$ طول کوتاه‌کن مترهای کوبایاشی است.» به گمانم بهتر است که بگوییم «نگاشت $h$ نسبت به متریکهای کوبایاشی کاهندهٔ فاصله است.»

۴. نکات ظاهری‌تر. بجز کاستیهایی که بدان اشاره شد، وارسی مختصری نشان می‌دهد که در امور ظاهری کتاب نیز دقت لازم به کار نرفته است. متأسفانه تعداد اشکالات از این دست بیش از آن است که فهرست کردنشان در اینجا ممکن باشد. با این حال، به چند نمونهٔ مهمتر اشاره می‌کنم:

• غلطهای فراوانی در حروفچینی فرمولهای ریاضی با TeX-پارسی دیده می‌شود که گاه مانند فرمول انتهای صفحهٔ ۲۴ گمراه کننده‌اند، وگاه متن را به نمونه‌های غلطگیری نشدهٔ پیش از چاپ شبیه می‌کنند. مثلاً صفحات ۱۲۱ و ۱۲۲ کتاب را به دقت بخوانید و با متن اصلی مقایسه کنید. از این مسأله که در انتخاب قلم[1]های مناسب اشتباهات بی‌شمار شده، می‌گذرم.

• بی‌قاعدگی در نقطه‌گذاری متن ترجمه کاملاً مشهود است. یک مثال بارز، ظهور نابهنگام نقطه در میان جمله است.

• صورت انگلیسی برخی از اصطلاحات فنی غیر متداول و اسامی خاص در پانوشت یا واژه‌نامه نیامده است. مثلاً خواننده به اصطلاح «بافه» یا «فرین» (به ترتیب معادل sheaf و extreme) بر می‌خورد، بی‌آنکه در هیچ کجا اشاره‌ای به معنی یا معادل انگلیسیشان شده باشد.

1. font

• استفاده از TeX-پارسی ظرایفی دارد که در مرحلهٔ حروفچینی و نمونه‌خوانی باید آنها را مد نظر داشت. یک نمونهٔ سهل‌انگاری در این متن، استفادهٔ مکرر از حرف انگلیسی $o$ به جای علامت ترکیب توابع است که از چشم‌نوازی بسیاری از فرمولها کاسته است. از آوردن مثالهای بیشتر خودداری می‌کنم.

• کتاب فهرست راهنما ندارد، که نقص کوچکی نیست، اما در عوض دو واژه‌نامهٔ انگلیسی-فارسی و فارسی-انگلیسی دارد. من در ضرورت و فایدهٔ این واژه‌نامه‌ها، که سنتاً به انتهای کتابهای ریاضی فارسی اضافه می‌شوند، جداً شک دارم. اما بحث من در اینجا صرفاً کیفیت واژه‌نامه‌های این کتاب است که صرف‌نظر از غلطهای فراوان، به روشی بسیار ناشیانه تنظیم شده‌اند، به این ترتیب که فهرست راهنمای نسخهٔ انگلیسی کتاب را ترجمه و پس از حذف شمارهٔ صفحات (یعنی مهمترین اطلاعات نهفته در آن)، آن را به عنوان واژه‌نامهٔ انگلیسی-فارسی به کتاب افزوده‌اند. نتیجهٔ ناگزیر این روش غریب، حضور دهها واژه و عبارت بی‌ربط و تکراری است که هیچ جایی در یک واژه‌نامه ندارند. حتی خوانندگانی هم که همچون من از فن تهیهٔ کتاب سررشته‌ای ندارند، با نگاهی کوتاه به این واژه‌نامه‌ها، مسأله را درخواهند یافت.

• شکل صفحهٔ ۳۸ وارون است، و خلاصهٔ ۱۳ سطری پشت جلد کتاب، حتی اگر یک بار نمونه‌خوانی می‌شد، آن‌چنان مغلوط نمی‌بود. اشتباهاتی چنین پیش‌پا افتاده، آن هم از ناشری معتبر، جداً بعید است.

این خلاصهٔ نکاتی بود که در مطالعهٔ اجمالی ترجمه به نظرم رسید. البته انصاف نیست که مسؤولیت این کاستیها را تماماً بر عهدهٔ مترجم یا ویراستار بگذاریم، چه بخش بزرگی از آنها می‌توانست با نظارت بهتر ناشر به کتاب راه پیدا نکند. به هر حال باید امیدوار بود که ناشر محترم به ویرایش دوباره و دقیقی از این کتاب اقدام کند و با رفع کاستیهای موجود نسخهٔ پیراسته‌ای در اختیار علاقه‌مندان قرار دهد. بحث در این‌باره را با سخنی از سر دلسوزی کوتاه می‌کنم: انتشار کتابهای وزین و آراستهٔ ریاضی در سالهای اخیر، به‌ویژه توسط مرکز نشر دانشگاهی، فرهنگ نوشتاری ریاضیات در ایران را به مراتب ارتقا داده و علاقه‌مندان فارسی زبان را مشکل‌پسندتر کرده است. شایسته است قدر این تحول اساسی دانسته شود. از سوی دیگر، اکنون که امکان ترجمهٔ آثار خارجی بی‌پرداخت حق کپی‌رایت به راحتی فراهم است، دست‌کم می‌توان حقوق معنوی نویسنده را پاس داشت. شاید گام اصلی در این راه آن باشد که در ترجمه، ویرایش، و چاپ این آثار نهایت دقت و امانت مبذول شود.

******

تأکید نامعقولی که در درسهای مقدماتی آنالیز مختلط بر مطالب کهنه و کم‌اهمیت می‌شود، باعث شده که اغلب دانشجویان آن را به صورت فنون پراکنده‌ای در سریهای توانی، یا محاسبهٔ انتگرالهای پیچیده به کمک مانده‌ها، یا یافتن فرمولهایی صریح برای مزدوجهای همساز ببینند. حتی محتوای درس مقدماتی دورهٔ کارشناسی ارشد نیز از این بلیه مصون نمانده است. دیر زمانی است که تحولات مهم در این شاخه، بهانه‌ای برای بازنگری در مواد این درسها به دست داده است. رهیافت هندسی به لم شوارتس و کاربردهای آن، پیدایش روشهای گسسته در آنالیز مختلط (مثلاً اثبات ردین-سالیوان برای قضیهٔ نگاشت ریمان به روش دایره انبانی[1] [7] و پیامدهای زیبایش [9])،

1. circle packing

و امکانات جدید در بررسی هندسهٔ تبدیلات موبیوس و تکرار توابع گویا به کمک کامپیوتر، مثالهای خوبی از این تحولات مهم‌اند. کتاب کرانتس یکی از اولین تلاشهای موفق در معرفی پاره‌ای از این تحولات به عامهٔ ریاضی دوستان است. هر چند که این اثر به مفهوم رایج یک کتاب درسی نیست، تجربهٔ نگارندهٔ این سطور نشان داده است که می‌تواند به عنوان متمم یک درس آنالیز مختلط دورهٔ کارشناسی ارشد، یا در سمینارهای ریاضی دورهٔ کارشناسی، به‌خوبی مورد استفاده قرار بگیرد.

## مراجع

1. L. Ahlfors, "An extension of Schwarz's lemma", *Trans. Amer. Math. Soc.* **43** (1938), 359-364.

2. L. Ahlfors, *Collected Papers*, Birkhäuser, Basel (1982).

3. A. Isaev and S. Krantz, "Invariant distances and metrics in complex analysis", *Notices Amer. Math. Soc.* **47** (2000), 546-553.

4. S. Kobayashi, "Intrinsic distances, measures and geometric function theory", *Bull. Amer. Math. Soc.* **82** (1976), 357-416.

5. R. Osserman, "From Schwarz to Pick to Ahlfors and beyond", *Notices Amer. Math. Soc.* **46** (1999), 868-873.

6. R. Remmert, *Classical Topics in Complex Function Theory*, Springer-Verlag, New York (1998).

7. B. Rodin and D. Sullivan, " The convergence of circle packings to the Riemann mapping", *J. Differential Geom.* **26** (1987), 349-360.

8. H. L. Royden, "Hyperbolicity in complex analysis", *Ann. Acad. Sci. Fenn.* Ser. A I Math. **13** (1988), 387-400.

9. K. Stephenson, "Circle packing and discrete analytic function theory", preprint, http://www. math. utk. edu/~kens

* * * * * *

* سعید ذاکری، دانشگاه پنسیلوانیا در فیلادلفیا، امریکا

zakeri@math.upenn.edu

# نامهٔ مترجم در بارهٔ این نقد

همکاران گرامی

با سلام، گرچه دریافت متن نقد و بررسی کتاب آنالیز مختلط: نگرش هندسی به‌علت اشاره به کاستیهای آن تلخ و گزنده است ولی به دلایل زیر مایلم بی‌کم و کاست در نشر ریاضی چاپ و انتشار یابد.

۱. نقد کتاب علاوه بر این که موجب ارتقاء کیفیت انتشارات می‌شود یک عمل فرهنگی است، که در صورت تداوم می‌تواند علاوه بر تربیت متخصصان مورد نیاز خود به‌صورت گفتمانی اجتماعی ودمکراتیک نهادینه شود.

۲. وجود فرهیختگانی در زمینهٔ ادبیات ریاضی به زبان فارسی که مصلحت علمی را فدای منافع دوستانه و فردی نمی‌کنند.

۳. وجود متخصصی که کتاب را با دقت خوانده و به ندای اینجانب در مقدمهٔ مترجم پاسخی درخور داده است.

۴. پذیرش خطا و پاسخگو بودن به آن تأثیرهای شگرف اجتماعی و فرهنگی دارد که غفلت از آنها موجبات عقب‌افتادگی علمی جامعهٔ ایرانی را فراهم می‌آورد.

همکاران گرامی، اکنون که ترجمهٔ اینجانب راهگشای احیای امر خطیر نقد و بررسی کتابهای ریاضی به زبان فارسی شده است پیشنهاد می‌نمایم

الف) ترجمهٔ دیگری از اینجانب بنام تابع شوارتس و کاربردهای آن نیز مورد نقد و بررسی قرار گیرد.

ب) نقد و بررسی کتابهای ریاضی به زبان فارسی جایگاه ویژه‌ای در نشر ریاضی داشته باشد و استمرار یابد.

پ) همه‌ساله در جلسه‌ای با حضور مؤلفین، مترجمین و نقادان از نقادان کتاب تجلیل شایسته‌ای به‌عمل آید.

ت) با چاپ مقالاتی در نشر ریاضی روشهای نقد کتاب آموزش داده شود.

در خاتمه ضمن تشکر از آقای دکتر سعید ذاکری به خاطر ارائهٔ تاریخچه‌ای از سیر تحول آنالیز مختلط و بررسی جامعی که از ترجمهٔ اینجانب به عمل آورده‌اند، خاطرنشان می‌کنم که پس از سالها تحصیل و تدریس و سکونت در تهران هنوز با شعر «حیدربابایه سلام» شهریار مأنوس‌تر از هر شعر دیگری هستم.

با سپاس

**محمد جلوداری‌ممقانی**

# فراریاضیات منطق فازی•

فرانسیس جفری پله‌تیر*
ترجمهٔ سیامک کاظمی

**Metamathematics of Fuzzy Logic, Peter Hájek, Trends in logic, vol. 4, Kluwer Academic Publishers (1998) viii+297pp.**

«منطق فازی» از دید افراد مختلف معناهای متفاوتی دارد. از نظر بعضیها نوعی فلسفهٔ زندگی است، «راهی برای برون‌رفت از تنگنایی که تفکر سیاه و سفید رایج در فرهنگ غربی برای ما ایجاد کرده است». از نظر دیگران، راه دقیقتری است برای توصیف زبان (و تفکر) معمولی ما «که در آن فکر نمی‌کنیم هر چیزی یا درست است یا غلط، بلکه به سایه‌های خاکستری نیز که در اندیشه و تفهیم و تفهّم زبانی ما وجود دارد، توجه داریم». بعضیها منطق فازی را ابزاری هستی‌شناختی می‌دانند که «واقعیت را دقیقتر از این ادعا توصیف می‌کند که بگوییم به‌ازای هر فقرهٔ موجود $a$ و به‌ازای هر ویژگی $F$، یا $a$ نشان‌دهندهٔ $F$ است یا $a$ نشان‌دهندهٔ نقیض $F$». در نظر برخی از نظریه‌پردازان، منازعهٔ «فازی در برابر کلاسیک» نوعی منازعهٔ «آنالوگ در برابر دیجیتال» مدرن است؛ و در شاخه‌های متعددی از مهندسی، منطق فازی به‌عنوان چارچوبی نظری در نظر گرفته می‌شود که در آن، آوردن عقربهٔ شمارنده به درجهٔ $n$، در جایی از کارخانه (یا ماشین) تغییر بزرگتری ایجاد می‌کند تا آنکه عقربه کاملاً به درجهٔ $n$ آورده نشود.

شک نیست کارهای مهندسی مهمی که با استفاده از تکنیک اخیر صورت گرفته به پیشرفتهای چشمگیری در تکنولوژی دوربین و یخچال انجامیده است، ولی خیلیها می‌پرسند که آیا این واقعاً «منطق فازیِ کاربردی» است یا آیا اصلاً منطقی فراتر از منطق دو ارزشی کلاسیک در آن مفروض گرفته شده است؟ (آیا واقعاً برای مدلسازیِ شمارنده‌ای که مدت زمان پخت در اجاق برقی (برحسب دقیقه) را به درجهٔ پختگی می‌نگارد، به منطق فازی نیاز است؟) و نیازی به گفتن ندارد که ادعاهای نظریه‌پردازان فازی در زمینه‌هایی بجز زمینه‌های فوق‌الذکر از پذیرش عام برخوردار نشده است. در واقع، چندان مبالغه نیست اگر بگوییم که اصطلاح «منطق فازی» صاحبنظران را به دو گروه تقسیم می‌کند: آنها که با شور و اشتیاقی مبلغانه از آن هواداری می‌کنند و کسانی که آن را نوعی ابتذال منطقی می‌دانند (که در نخستین نگاه، جالب و مهیج است ولی اندکی بعد سطحی به نظر می‌رسد و توی ذوق می‌زند).

در این آشفته بازار دیدگاهها و چارچوبهای مختلف، کتاب هایک نمونهٔ خوشایندی عرضه می‌کند. نخستین تدبیر او تمایزگذاشتن بین «منطق فازی در معنای وسیع آن» و «منطق فازی در معنای محدود آن» است؛ و ادعا می‌کند که هدف کتاب حاضر، بررسی مفهوم محدود منطق فازی است که به نظر او همراه است با بررسی محاسبات صوری منطق چندارزشی. با این حال و علی‌رغم توجهی که در بیشتر کتاب به ویژگیهای صوری منطق فازی مبذول شده است، هایک صفحات آغازی را به بحث در بارهٔ مسائلی از این‌گونه اختصاص داده است که آیا منطق فازی عملاً کاربرد دارد یا آیا می‌توان آن را به عنوان مُدلی برای ابهام و عدم حتمیت به کاربرد یا نه. (هایک مصرّ است که این منطق، منعکس‌کنندهٔ نظریه احتمال نیست.) و بحث کوتاه اما جالبی هم در این باره آمده است که «ارزش صدق از کجا می‌آید» و «ارزشها را چگونه می‌توان تعبیر کرد».

هایک در مورد خوانندگان کتاب ادعا می‌کند: «فرض بر این است که می‌توانند استدلالهای ریاضی را تعقیب کنند و دست‌کم از معلومات ریاضی پایه برخوردارند». آنها همچنین به «دست‌کم تجربهٔ اندکی در مورد منطق گزاره‌ای (بولی) کلاسیک [و] قدری اطلاعات از حساب محمولات» نیاز دارند. او مدعی است که «دانش عمیقی از منطق ریاضی یا جبر را مفروض نگرفته است». تجربهٔ این ناقد نشان می‌دهد که هایک در مورد میزان مطالبی که خواننده‌ای با این زمینه می‌تواند به آسانی دریابد، مبالغه کرده است.

منطق فازی در «معنی محدود»ش، منطقی بینهایت ارزشی است. علی‌رغم اعتقاد بسیاری از نویسندگان متون عامه‌پسندِ منطق فازی به

اینکه منطق بینهایت‌ارزشی را لطفی زاده در ۱۹۶۵ ابداع کرد (رک. جورنال آو سیمبالیک لاجیک، جلد XXXVIII، ص ۶۵۶)، نخستین بار ووکاشه‌ویچ[1] و تارسکی بودند که در دهه‌های ۱۹۲۰ و ۱۹۳۰ به تحقیق در این مبحث پرداختند و در دههٔ ۱۹۵۰ و اوایل دههٔ ۱۹۶۰ دوباره علاقه زیادی به این منطق پیدا شد. هایک بعضی از این مطالب ذی‌ربط تاریخی را در فصل پایانی بسیار کوتاه کتاب (که در آن فقط دو صفحه به دورهٔ قبل از انتشار مقالهٔ زاده در ۱۹۶۵ اختصاص دارد) آورده است. این ناقد ترجیح می‌داد که مؤلفْ منطق فازی را به‌عنوان محصول تلاش تاریخی برای بررسی منطقهای بینهایت‌ارزشی تشریح و ارائه کند. هایِک فکر روشنی در بارهٔ نحوهٔ بررسی منطقهای بینهایت‌ارزشی دارد و این کار را از دیدگاه و طبق سلیقهٔ خودش به‌خوبی انجام می‌دهد، حتی بدون قیدوشرط می‌توان گفت به‌خوبی انجام می‌دهد.

تا آنجا که این ناقد دیده است، سایر متونی که به تشریح منطق فازی پرداخته‌اند با ادعایی از این‌گونه آغاز می‌شوند: «هر حرفِ گزاره‌ای، ارزشی در بازهٔ حقیقی [۱ , ۰] می‌گیرد. عطف [ترکیب عطفی] جمله‌ها حداقلِ ارزشِ مؤلفه‌های آن است، در حالی‌که فصل [ترکیب فصلی] جمله‌ها، حداکثرِ ارزشِ مؤلفه‌های آن است. نقیض به‌صورت ۱ منهای ارزش مؤلفهٔ نقض [نفی] نشده تعبیر می‌شود.» هایک، به جای این کار، بحث معنایی خود را با تشریح مفهومی بسیار کلی از تابع ارزش عطفی* شروع می‌کند، که باید با عطف کلاسیک روی ارزشهای ۰ و ۱ همخوان باشد؛ باید تعویض‌پذیر و شرکت‌پذیر، و نسبت به هر دو شناسهٔ غیرنزولی باشد. $x = (1 * x)$ و $0 = (0 * x)$؛ و باید تابعی پیوسته از $[0,1]^2$ به توی $[0,1]$ باشد. استراتژی هایِک، پروراندن یک نظریهٔ معنایی برای یک منطق بنیادی (BL) است که درست این نوع «ضعیف» از عطف را داشته باشد، و سپس افزودن محدودیتهای گوناگون به عملگر عاطف که معناهایی برای منطقهای فازی متفاوت و متعدد ایجاد می‌کنند. راههای گوناگون افزودن محدودیت به این عملگر، به اصطلاح هایِک، به «t_نرم»های متمایزی می‌انجامند. سه محدودیت متفاوت مورد نظرند:

| | |
|---|---|
| ووکاشه‌ویچ | $(x * y) = \max(0, x + y - 1)$ |
| گودل | $(x * y) = \min(x, y)$ |
| گوگوئن[2] | $(x * y) = xy$ (حاصلضرب عددهای حقیقی) |

(شایان ذکر است که هایک گاهی کلمهٔ «گوگوئن» را برای این آخری به‌کار می‌برد و گاهی «حاصلضرب» را.) با این تعریفهای عطف در منطق فازی، فقط منطق گودل است که با منطق فازی به مفهوم رایج شباهت پیدا می‌کند.

با مفروض بودن هر یک از این t_نرم‌ها می‌توان یک [ترکیب] شرطی (معنایی) $\Rightarrow$ را به‌عنوان متمم نسبی[3] اش به‌صورت زیر تعریف کرد

$$(x \Rightarrow y) = \text{ماکسیمالی که در } (x * z) \leq y \text{ صدق می‌کند } z$$

متناظر با این سه تعریف متفاوت *، سه [ترکیب] شرطی متفاوت داریم که همه مقدار یکسانی، یعنی ۱، را به‌دست می‌دهند به شرط آنکه مقدم دارای ارزش صدقی برابر یا کوچکتر از تالی باشد، ولی در حالاتی که مقدم «صادقتر» از تالی باشد، تفاوت دارند:

اگر $x \leq y$، آنگاه $1 = (x \Rightarrow y)$؛ در غیر این‌صورت

| | |
|---|---|
| ووکاشه‌ویچ | $(x \Rightarrow y) = 1 - x + y$ |
| گودل | $(x \Rightarrow y) = y$ |
| گوگوئن | $(x \Rightarrow y) = y/x$ (تقسیم اعداد حقیقی) |

از مدتها قبل در جامعهٔ منطقدانان فازی بر سرِ تعریف درست [ترکیب] شرطی اختلاف‌نظر وجود داشته است (و هنوز هم وجود دارد)، و در اینجا می‌بینیم که چگونه سه تا از این تعریفها پیامد انتخاب عطف‌اند. (در نوشتگان این مبحث، تعریفهای دیگری نیز برای استلزام رایج است که به این طریق به‌دست نمی‌آیند؛ بعضی از آنها حتی ۲ ارزشی‌اند.)

هایِک از $\Rightarrow$ برای تعریف $(-)$ که یک عملگر ناقض معناشناختی است، استفاده می‌کند: $(x \Rightarrow 0) =_{\text{تعریف}} (-)x$. به این ترتیب برای سه نوع $\Rightarrow$ متفاوت ما دو نقیض به‌دست می‌آید:

ووکاشه‌ویچ: $(-)x = 1 - x$

گودل: $1 = (-)x$ اگر $0 = x$، و $0 = (-)x$ در غیر این‌صورت

گوگوئن: $1 = (-)x$ اگر $0 = x$، و $0 = (-)x$ در غیر این‌صورت

خاطرنشان می‌شود که فقط نقیضِ نوع ووکاشه‌ویچ متناظر با تعریف نقیض در منطق فازی سنتی است، و نقیض دیگر اساساً یک عملگر دو ارزشی است که می‌گوید جملهٔ نقض‌نشده ارزش ۰ دارد یا نه.

معهذا، هر کدام از این t_نرم‌ها که داده شده باشد، با استفاده از & به‌عنوان اداتی در زبان که نشان‌دهندهٔ * است و استفاده از $\to$ به‌عنوان ادات نمایشگر $\Rightarrow$، می‌توانیم اداتهای فازی سنتی $\wedge$ و $\vee$ (عملگرهای مینیمم و ماکسیمم) را تعریف کنیم:

$$A \wedge B \text{ عبارت است از } A \mathbin{\&} (A \to B)$$

$$A \vee B \text{ عبارت است از } ((A \to B) \to B) \wedge ((B \to A) \to A)$$

(توجه کنید که $\wedge$، و نه &، عملگر عاطفی است که در تعریف $\vee$ به‌کار می‌رود.) با مفروض بودن ثابت ۰ در زبان، نقیض و هم‌ارزی تعریف می‌شوند:

$$\neg A \text{ عبارت است از } A \to 0$$

$$A \equiv B \text{ عبارت است از } (A \to B) \mathbin{\&} (B \to A)$$

(توجه کنید که &، و نه $\wedge$، عملگر عاطف به‌کار رفته در تعریف $\equiv$ است.)

خواننده ممکن است توجه کرده باشد که بسیاری از ویژگیهای متعارف اداتهای متفاوت در BL محفوظ می‌مانند، از قبیل قوانین معمولی توزیع‌پذیری، قوانین تعویض‌پذیری، قوانین دمورگن (برای $\wedge$، $\vee$، و $\neg$)، و نظایر اینها. توجه کنید که $(p \to \neg\neg p)$ یک قضیه است ولی $(\neg\neg p \to p)$ نیست (و در واقع با تعبیر گودلی_گوگوئنی صادق نیست). همچنین خاطرنشان می‌کنیم که قضیهٔ استنتاج با کلیت تمام در BL برقرار نیست. به جای آن داریم

$\Gamma \cup \{\varphi\} \vdash \psi$ اگر و تنها اگر $n$ای وجود داشته باشد که $\Gamma \vdash (\varphi^n \to \psi)$

1. Łukasiewicz 2. Goguen 3. residuum

که در آن $\varphi^n$ عبارت است از $(\varphi \& \varphi \& \cdots \& \varphi)$، $n$ بار. (قضیهٔ استنتاج نامقید فقط برای منطق گودل برقرار است.) همین‌طور ممکن است در بارهٔ فشردگی معناشناختی سؤال پیش آید: آیا درست است که «$\Gamma \models \varphi$ اگر و تنها اگر به‌ازای $\Delta$ای که زیرمجموعه‌ای متناهی از $\Gamma$ باشد، $\Delta \models \varphi$»؟ با در دست داشتن منابع منطق فازی، ممکن به نظر می‌رسد که مجموعهٔ متناقض نامتناهیی چون $\Gamma$ بسازیم که هیچ زیرمجموعهٔ متناهی متناقضی نداشته باشد. ولی در آن صورت به‌ازای هر $\Delta$ی متناهی، $\Gamma \models \varphi$ معتبر و $\Delta \models \varphi$ نامعتبر خواهد بود. این امر پیامدهایی برای تمامیت دارد زیرا در آن صورت، شناسهٔ معتبری چون $\Gamma \models \varphi$ وجود دارد که استنتاجی برای آن وجود ندارد (زیرا استنتاجها متناهی‌اند).

می‌توان دید که بسیاری از تابعهای ارزش با استفاده از این t_نرمها قابل تعریف‌اند. ولی استدلال کاردینالیِ ساده‌ای نشان می‌دهد که منطق فازی به‌وضوح ناتمام است زیرا تعداد ناشماراییِ تابع ارزش $n$تایی از $[0,1]^n$ به $[0,1]$ وجود دارد. هایک در قضیهٔ بسیار جالبی نشان می‌دهد که هر t_نرم پیوسته (با هر تعریف معقولی از عطف) ترکیبی از سه t_نرمی است که او انتخاب کرده است.

هایک دو فصل را به سه منطق فازی تعریف‌شده، اصل موضوعی‌سازی آنها، و نشان‌دادن حالاتی که آنها قابل تعبیر متقابل‌اند، اختصاص می‌دهد. در پایان فصل ۳ (در بارهٔ منطق فازی ووکاشه‌ویچ) شرح می‌دهد که این منطق چه ارتباطی با پارادُکس قیاس مسلسل[1] دارد. در انتهای این بحث، ادعا می‌کند که به این مثال در فصلهای بعدی که راجع به منطق گودل و گوگوئن هستند و نیز در منطق محمولات فازی بیشتر پرداخته خواهد شد. به نظر ناقد مایهٔ تأسف است که این بحث بسیار جالب ادامه نیافته است.

هایک در فصل ۵ تحلیل خود را به منطق محمولات گسترش می‌دهد. بار دیگر با BL شروع می‌کنیم و منطقی را که با افزودن محمولها، ثابتها، متغیرها، و سورهای $\forall$ و $\exists$ ایجاد می‌شود در نظر می‌گیریم. (این منطق را $BL\forall$ می‌نامند.) قضیه‌های مهم اینها هستند:

$$\exists x\varphi \to \neg\forall x\neg\varphi \qquad \text{و} \qquad \neg\exists x\varphi \equiv \forall x\neg\varphi$$

و این واقعیت که عکس قضیهٔ اولی، قضیه‌ای از $BL\forall$ نیست. از لحاظ معناشناختی ایده این است که ثابتها نشان‌دهندهٔ (دقیقاً!) عناصری از یک حوزه‌اند، در حالی که محمولهای $n$ موضعی نشان‌دهندهٔ یک «رابطهٔ $n$تایی فازی» هستند (که ارزش آن محمول را به‌ازای هر $n$تایی انتخاب‌شده از آن حوزه به‌دست می‌دهد). $\forall x\varphi$ دارای ارزشِ صدقِ «نمونه‌ای از $\varphi$ که کمترین صدق را دارد» می‌باشد در حالی که $\exists x\varphi$ ارزش صدقِ «صادق‌ترین نمونهٔ $\varphi$» را دارد. اگر چنین نمونه‌هایی وجود نداشته باشند، ارزش فرمول تعریف نمی‌شود. (این امر، انحرافی از منطق فازی سنتی به نظر می‌رسد که در آن به چنین فرمولهایی ارزش حدی $\varphi$های گوناگون نسبت داده می‌شود.) این محدودیت به این طریقِ اِعمال می‌شود که فقط مدلهای «سالم» در ملاحظات مربوط به تمامیتْ ذی‌ربط به حساب می‌آیند؛ تمامیت نسبت به ردهٔ $\mathcal{C}$ — جبرهای خطی — مرتبی که سالم‌اند، در نظر گرفته می‌شود. بنابراین هایک از این نتیجه‌گیری می‌پرهیزد که منطق فازی را از طریق گسترش ردهٔ جبرهای ارزش نمی‌توان به‌طور بازگشتی اصل موضوعی کرد. منطقدان فازی سنتی قاعدتاً می‌گوید که این تمامیت «واقعی» نیست بلکه فقط تمامیت نسبت به مجموعهٔ محدودی از مدلهاست. هم منطقدان فازی سنتی و هم فرد شکاک نسبت به منطق فازی خواهند گفت که اهمیت یا ذی‌ربط بودن این ردهٔ جدید از مدلها باید توجیه شود تا بتوانیم با احساس مسؤولیت بگوییم که منطق فازی صورت اصل موضوعی کامل و قابل‌قبولی یافته است.

1. sorites paradox

هایک پس از ارائهٔ $BL\forall$ به گسترش منطقهای ووکاشه‌ویچ، گودل، و گوگوئن با محمولها، نامها، و سورها می‌پردازد. وی نشان می‌دهد که همانگوهای منطق محمولات گودل مطابق‌اند با همانگوهای رویِ «$G$_جبر استاندارد $[0,1]_G$ از تابعهای ارزش [و] بنابراین $G\forall$ به‌طور بازگشتی اصل موضوعی می‌شود». خوانندهٔ علاقه‌مند ممکن است توجه کند که این کار نسبت به زیربازهٔ شمارای $[0,1]$ انجام می‌شود. هایک سپس نشان می‌دهد که هیچ نتیجه‌گیری مشابهی در مورد منطقهای ووکاشه‌ویچ یا گوگوئن میسر نیست: اصل موضوعی‌سازی بازگشتی برای هیچ‌یک از این منطقها ممکن نیست. (این امر از قضیه‌ای متعلق به اسکارپلینی[1] [رک. جورنال آو سیمبالیک لاجیک، جلد XXIX، ص ۱۴۵] نتیجه می‌شود). بسیاری از ناظران ممکن است فکر کنند که این موضوع ناقوس مرگ منطق فازی را به صدا درمی‌آورد. به نظر ناقد، هایک می‌توانست وقت بیشتری صرف این موضوع کند.

مباحث متعدد و بسیار جالب دیگری نیز در کتاب آمده است، از قبیل بررسی تساوی (تشابه؟) در منطق فازی، موضوعات پیچیدگی محاسبه و تصمیم‌ناپذیری، فصل کاملی در بارهٔ «استنباط تقریبی»، سورهای تعمیم‌یافته، مفاهیم وجهی، شرحی از منطقهای فازی دیگر، بحثی در بارهٔ پارادُکس دروغگو در منطق فازی («این جمله دست‌کم کمی غلط است»)، و بسیاری مباحث دیگر. گوهرهای فراوانی در این بحثها نهفته است. به نظر ناقد، به‌خصوص بحث مفصل هایک در فصل ۷ در بارهٔ مفهوم زاده از «قاعدهٔ ترکیبی استنباط» و امکان سایر «قواعد فازیِ استنباط» برای نظریه‌پردازانی که منطقدان نیستند ولی به توصیف استدلال یا توصیف کنترل‌کننده‌های فازی علاقه دارند بسیار جالب‌توجه و مهم است.

خطاهای چاپی (جزئی) که ناقد یافته است، آنقدر هست که حدس بزند تعداد این‌گونه خطاها درکتاب احتمالاً زیاد است. بعضی از اثباتها شبیه بازنویسیِ کلمه به‌کلمهٔ طرحهای اثباتی‌که هایک روی تخته‌سیاه نوشته باشد، به نظر می‌آیند. فهم اختصارات و قواعد غیررسمیِ به‌کار رفته در کتاب گاهی دشوار است.

هر کس توجهی به منطق فازی دارد — خواه از «معتقدان راستین» آن باشدو خواه ازکسانی که آن‌را نوعی ابتذال منطقی می‌دانند — باید این کتاب را داشته باشد و از آن چیز بیاموزد. اگر بگوییم این کتاب بهترین کتاب در بارهٔ ویژگیهای صوری منطق فازی است، تجلیل چندانی از آن نکرده‌ایم زیرا کتاب دیگری در این باره وجود ندارد. اما اگر کتابهای دیگری هم وجود می‌داشت باز هم می‌توانستیم «فازی‌وار» استنباط کنیم که این یکی از همه بهتر است.

******

• Francis Jeffry Pelletier, *The Bulletin of Symbolic Logic*, (3) 6 (2000) 342-346.

* فرانسیس جفری پله‌تیر، دانشگاه آلبرتا، کانادا

1. Scarpellini

# گزیدهٔ آثار الیس کولچین•

اندی مجید*

ترجمهٔ شهرام محسنی‌پور

***Selected Works of Ellis Kolchin*, H. Bass, A. Buium, and P. J. Cassidy (editors), American Mathematical Society (1999) xiii + 639 pp.**

«مجموعهٔ آثار» نوع متعارفی از کتب ریاضی در کنار سایر انواع آن از قبیل تکنگاشت، کتاب درسی، یادداشت درسی، و گزارش کنفرانس است. ولی برخلاف بقیهٔ انواع، چندان معلوم نیست که چقدر ارزش دارد که ریاضیدانی شخصاً مجموعهٔ آثاری در تملک داشته باشد. در اتاق من در بین صدها کتاب ریاضی موجود دقیقاً سه مجموعهٔ آثار وجود دارد که تنها یکی از آنها را زمانی طولانی است که دارم: مجموعهٔ آثار آبل، ویراستهٔ سیلو و لی به تاریخ چاپ ۱۸۸۱، از انتشارات انجمن ریاضی نروژ. من آن را برای یادگاری در سال ۱۹۷۰ در سفری به اسلو به مناسبت کنفرانس هندسهٔ جبری از ناشر خریدم. اینکه هنوز، نود سال بعد از چاپ اول آن، نسخه‌های ویراست اولیه با همان بهای اولیهٔ بسیارکم به فروش می‌رسید، نشانه‌ای است از وضع فروش مجموعه‌های آثار. مجموعهٔ آثار دیگر را در قسمت حراج یک کتابفروشی در یک مرکز خرید روستایی با بهایی کمتر از پنج درصد نرخ ناشر پیدا کردم. به‌علاوه چون مؤلف را می‌شناختم و برایش احترام قائل بودم به‌راحتی تصمیم به خرید گرفتم. و اما در بارهٔ سومی: می‌خواستم از سهمیهٔ خرید کتابم تماماً استفاده کنم و کتابی از سری مجموعه/گزیدهٔ آثار AMS [انجمن ریاضی آمریکا] کاملاً با آن جور درمی‌آمد. هر چند همهٔ این خریدها مربوط به ریاضیات و از جهات متعددی نیز مایهٔ رضایت خریدار بودند، نمی‌توانم بگویم که در هیچ مراجعهٔ حرفه‌ای و تخصصی آنها را خوانده‌ام.

در اینکه خیلی از مجموعه‌های آثار به این منظور خوانده شده باشند، شک دارم. ولی گزیدهٔ آثار الیس کولچین با شرح و حواشی می‌تواند استثنا باشد. اگر بخواهیم اولین جملهٔ ویراستاران را در مقدمه به عبارتی دیگر بیان کنیم، جبر دیفرانسیلی، به‌ویژه نظریهٔ گالوای دیفرانسیلی در قالب دقیق جبری و امروزیش عمدتاً آفریدهٔ الیس کولچین است. کولچین علاوه بر مقالات منتشرشده‌اش، نتایج خود را در شکل نهایی آن در دو تکنگاشت پژوهشی عمده، [۶] و [۷]، ارائه کرده است (که در این مجلد دوباره چاپ نشده‌اند). در این دو تکنگاشت درونمایه‌های اصلی یعنی نظریهٔ گالوای دیفرانسیلی و گروههای جبری دیفرانسیلی درکلیتی استادانه همراه با اثبات همهٔ نتایج مهم در وسیعترین چارچوب ممکن عرضه شده‌اند. ولی این دو تکنگاشت پژوهشی با وجود کامل و دقیق بودنشان نمی‌توانند انتخابی مناسب برای مطالعه در نزد خود به‌ویژه برای مبتدیان باشند.[۱]

با این‌حال، نیاز خواننده مبتدی با مراجعهٔ مستقیم به مقالات اولیهٔ کولچین که در اینجا دوباره چاپ شده‌اند، می‌تواند کاملاً تأمین شود: در [۸] نظریهٔ گالوای دیفرانسیلی پیکار-وسیو، یا نظریهٔ گالوای میدانهای دیفرانسیلی که در آن گروه گالوا گروهی جبری خطی است، تشریح شده است و در مرجع [۹] نظریهٔ گالوای دیفرانسیلیِ قویاً نرمال یا نظریهٔ گالوا برای میدانهای دیفرانسیلی که در آن گروه گالوای دیفرانسیلی یک گروه جبری خطی است، آمده است.

مثالی ساده از توسیع پیکار-وسیو به‌صورت زیر به‌دست می‌آید: با میدان توابع کسریِ یک متغیر مختلط، $F = C(t)$، به‌همراه مشتقگیری $D_F = d/dt$ شروع می‌کنیم و سپس میدان توسیعی دیفرانسیلی $E = F(y, w) \supset F$ را با افزودن یک پادمشتق $y = \log(t) = \int 1/t$ و/یا یک تابع نمایی $w = e^t$ تشکیل می‌دهیم (مشتقگیری $D_E$ به‌صورت $D_E(y) = 1/t$، $D_E(w) = w$ و برای $f \in F$ به‌صورت $D_E(f) = D_F(f)$ داده می‌شود). خودریختی دیفرانسیلی $F(w)$ روی

۱. در این کتابها از زبان و اصطلاحات مبانی ویل [۱۷] استفاده شده است که برای دانشجویانی که هندسهٔ جبری را براساس طرح (scheme)ها آموخته‌اند، کمتر آشناست.

$F$ باید $w$ را به مضرب اسکالر مختلط ناصفری از خودش یعنی $cw$ ببرد که $c \in \boldsymbol{C}$، و این امر خودریختی را مشخص می‌کند. به‌علاوه هر $c$ در اینجا ظاهر می‌شود. بنابراین مابین گروه $G(F(w))/F)$ از خودریختی‌های دیفرانسیلی $F(w)$ روی $F$ و گروه ضربی $\boldsymbol{G}_m(\boldsymbol{C})$ تناظری دوسویی وجود دارد. همین‌طور، چون خودریختی‌های دیفرانسیلی $F(y)$ روی $F$ باید $y$ را به $y+d$ ببرند که $d$ اسکالری مختلط است، معلوم می‌شود که تناظری دوسویی بین $G(F(y)/F)$ و گروه جمعی $\boldsymbol{G}_a(C)$ وجود دارد. آن چیزی که $F(w) \supset F$ (یا $F(y) \supset F$) را بدل به توسیع پیکارـوسیو می‌کند این است که میدان بالایی به‌عنوان یک میدان دیفرانسیلی روی میدان پایینی با مجموعهٔ پُری از جوابهای یک معادلهٔ دیفرانسیل خطی همگن روی میدان پایینی تولید می‌شود (به ترتیب $W' - W = 0$ و $Y'' + tY' = 0$). در اینجا «پُر» بدین معنی است که اندازهٔ بعد روی ثابتهای مجموعهٔ جوابها در توسیع، با مرتبهٔ معادله‌ها مساوی است. «ثابتها» عناصر با مشتق صفر هستند و همچنین لازم است که هر ثابتی در میدان توسیعی، از قبل در پایه بوده باشد. معمولاً فرض می‌شود که ثابتها زیرمیدانی تشکیل می‌دهند که بستهٔ جبری است و دارای مشخصهٔ صفر است.

برج توسیعهای پیکارـوسیو لزوماً پیکارـوسیو نیست (برای مثال $\boldsymbol{C}(t,\log(t),\log(\log(t))) \supset \boldsymbol{C}(t,\log(t)) \supset \boldsymbol{C}(t)$). برج توسیعهای پیکارـوسیو را توسیع لیوویل گوییم هرگاه گروه گالوای دیفرانسیلی هر طبقه از توسیعها $\boldsymbol{G}_m$ یا $\boldsymbol{G}_a$[1] باشد. یکی از نتایج کاربردی اصلی در [۸]، عبارت است از اینکه توسیع پیکارـوسیو در یک توسیع لیوویل می‌نشیند اگر و تنها اگر گروه گالوای دیفرانسیلی توسیع پیکارـوسیو حل‌پذیر باشد. چون گروه مفروض $\boldsymbol{G}_a$ با افزودن یک پادمشتق و نیز گروه مفروض $\boldsymbol{G}_m$ با افزودن یک نمایی به‌دست می‌آید، طریقهٔ دیگر بیان قضیهٔ بالا این است که بگوییم مجموعهٔ پُری از جوابهای یک معادلهٔ دیفرانسیل را می‌توان با افزودن مکرر نمایی‌ها و پادمشتق‌ها به‌دست آورد اگر و تنها اگر گروه گالوایِ دیفرانسیلی متناظر گروهی حل‌پذیر و همبند باشد.

مثالی از توسیع قویاً نرمالی که پیکارـوسیونیست، با شروع دوباره از $F = \boldsymbol{C}(t)$ و افزودن یک $p$ تابع وایرشتراس $y$ و مشتق آن ($E = F(y, y') \supset F$) به‌دست می‌آید. از اینجا معلوم می‌شود که گروه گالوای دیفرانسیلی $E$ روی $F$، $G(E/F)$، یک چندگونای آبلی یک‌بعدی (خم بیضوی) متناظر با $y$ است.

قبلاً گفتیم که نیاز خوانندهٔ مبتدی با مراجعهٔ مستقیم به [۸] و [۹] می‌تواند کاملاً تأمین شود. ولی این مجلد چیز بیشتری دربردارد: مقالهٔ بورل [۲] که یکی از چهار مقالهٔ عالی توصیفی و مروریِ جدید است که ویراستاران در بخش آخر این کتاب، بخش «شرح و حواشی»، آورده‌اند. بورل علاقهٔ کولچین را به نظریهٔ گروههای جبری روی میدانی دلخواه و کارهای او را در این زمینه تشریح می‌کند. او همچنین نظریهٔ گالوای کولچین را به طرز درخشانی خلاصه می‌کند: هم در حالت توسیعهای پیکارـوسیو و هم در حالتِ به‌مراتب پیچیده‌تر توسیعهای قویاً نرمالِ عام (پیچیدگی از آنجا ناشی می‌شود که لازم می‌آید یکریختیهای توسیع $E \supset F$ را به توی یک میدان دیفرانسیلی عمومی $U \supset F$ و شامل $E$ در نظر بگیریم). «عمومی» بدین معناست که همهٔ توسیعهای دیفرانسیلیِ متناهیْ‌مولدِ زیرمیدانهای دیفرانسیلیِ $U$ که روی $F$ متناهی‌مولدند، در $U$ نشانده شوند) خوانندهٔ مبتدی که از [۲] شروع می‌کند و سپس به سراغ [۸] و [۹] می‌رود، دیگر مبتدی نخواهد بود.

کولچین همچنین تا حدی به حالتهای خاص و کاربردهای نظریهٔ گالوای دیفرانسیلی علاقه‌مند بود و این مجلد شامل آثار وی در این زمینه نیز هست. معهذا خوانندهٔ علاقه‌مند به این مباحث چشم‌اندازی بسیار کاملتر و امروزیتر را در مقالهٔ سینگر [۱۵] در بخش «شرح و حواشی» خواهد یافت. سینگر مسأله‌های معکوس و مستقیم نظریهٔ گالوای دیفرانسیلی [پیکارـوسیو] را بررسی می‌کند: «مستقیم» بدین معناست که «با مفروض بودن یک معادلهٔ دیفرانسیل خطی همگن، گروه جبری دیفرانسیلی خطی متناظرش را معین کنید» و «معکوس» بدین معناست که «معادلهٔ دیفرانسیلی پیدا کنید که گروه گالوای دیفرانسیلی آن گروه جبری خطیِ داده شده باشد»[1]. (در هر دو حالت، میدان دیفرانسیلی پایه معین است) و برای هر دو مسأله روشهای ساختیْ مطلوب است.

برای مسألهٔ مستقیم، راه‌حل کاملی وجود ندارد. منظور از راه‌حل، الگوریتمی است برای محاسبهٔ گروه گالوای دیفرانسیلی $G(E/F)$ برای توسیع پیکارـوسیوی $E$ی معادلهٔ دیفرانسیل $L(Y) = 0$ روی میدان $F$، حتی در حالتی که $F$ میدانی ساختنی مانند $\bar{Q}(t)$ باشد. روش محاسبه را، حتی در صورتِ بدیهی بودن و یا متناهی بودن $G(E/F)$ نیز نمی‌دانیم. با وجود این در صورتی که شرطهایی اضافی روی $G$ یا روی $L$ (برای مثال، تقریباً‌حل‌پذیری برای $G$ و داشتن مرتبهٔ ۲ برای $L$) گذاشته شود، جوابهای متعددی برای مسألهٔ مستقیم وجود دارد و همچنین در بسیاری ازمواردی که بتوان شرطهایی روی کنش $G(E/F)$ بر مجموعهٔ جوابهای $V = \{y \in E | L(y) = 0\}$ از معادلهٔ $L$ در $E$ گذاشت، نیز جوابهای مختلف وجود دارد. (یادآوری می‌کنیم که یکی از شرطها برای توسیع پیکارـوسیو این بود که $V$ پُر باشد، یعنی روی میدان ثابتهای $C$ در $F$ یک فضای برداری باشد و همچنین بعد $E$ برابر با مرتبهٔ $L$ باشد). $G = G(E/F)$ گروهی جبری روی $C$ است و کنش آن روی $V$ جبری (و صادق[2]) است. در حالتی که کنش $G$ روی $V$ مثلاً تحویل‌پذیر باشد، به‌ویژه هنگامی‌که $V$ زیرمدول یک‌بعدی یا جمعوند زوج داشته باشد و بعد $G$ کوچک باشد، الگوریتمها شناخته شده‌اند. در حالت کلیتر از زوج $V$ و $G$ می‌توان زیررستهٔ آبلی تانسوردار از رستهٔ همهٔ $G$ مدولهای جبری را تولید کرد. بر طبق نظریهٔ رسته‌های تاناکایی، این رسته $G$ را تعیین می‌کند و در نتیجه شرطهای مربوط به کنش $G$ روی $V$ گاهی می‌تواند مشخص‌کنندهٔ $G$ باشد.

اولین بار در سال ۱۹۸۰، دُلینی [۴] رسته‌های تاناکایی را وارد نظریهٔ گالوای دیفرانسیلی کرد؛ ادامهٔ مطلب را ببینید.

بررسی مسألهٔ معکوس قبلاً روی میدان پایهٔ $F = \boldsymbol{C}(t)$ انجام می‌شد. همان‌طور که س. ترتکوف و م. ترتکوف نشان داده‌اند این واقعیت که هر گروه

---

۱. یا متناهی

۱. چنانکه سینگر می‌گوید، کولچین در سخنرانی خود در کنگرهٔ بین‌المللی ریاضیدانان در مسکو، به سال ۱۹۹۶، اینها را مسائل برجستهٔ عمده در نظریهٔ پیکارـوسیو اعلام کرد. متن این سخنرانی در این کتاب چاپ شده است.

2. faithful

جبری خطیِ $G$ روی $C$ یک گروه گالوای دیفرانسیلی روی $F$ است خود نتیجه‌ای از قضایای وجودی تحلیلی است. برای انواع خاصی از گروههای همبند، برهانهای جبری ارائه شده‌اند، و سینگر و میتشی [۱۳] برهانی جبری برای همهٔ گروههای همبند ارائه کردند. ولی برای گروههای دلخواه به‌ویژه برای گروههای متناهی، مسأله هنوز حل نشده است: در اینجا مسأله عیناً مانند این است که آنها را به‌صورت گروههای گالوای معمولی روی $F$ نمایش دهیم. این نکته‌ای است که از منظر تحلیلی معلوم است (قضیهٔ وجودی ریمان) ولی کلاً از منظر جبری چنین نیست.

برای بعضی دیگر از میدانهای پایه نیز مسألهٔ معکوس قابل حل است. به عنوان مثال، رمیس [۱۲] راه حلی برای سریهای توانی همگرا روی $C$ ارائه کرده است.

همان‌طور که قبلاً گفته شد، هنگامی‌که کولچین شروع به بررسی نظریهٔ گالوا برای توسیعهای میدانی دیفرانسیلی $F \supset E$ کرد که در آن لزوماً $G(E/F)$ خطی نیست، ناچار بود که یک توسیع دیفرانسیلی عمومی $F \supset U$ را به‌کار بگیرد. $U$ شامل نسخه‌های یکریختِ همهٔ توسیعهای دیفرانسیلی متناهیْ‌مولدِ توسیعهای دیفرانسیلی متناهیْ‌مولد $F$ در داخل $U$ است. این مطلب به‌طور طبیعی منجر به مفهوم میدان «دیفرانسیلی بسته» و «بستار دیفرانسیلی» می‌شود.

دو مفهوم اخیر علاوه بر جبر دیفرانسیلی در منطق هم مفید واقع شده است. مقالهٔ پوازا [۱۴] در بخش «شرح و حواشی» این موضوع را به‌خوبی شرح می‌دهد. همچنین در مقالهٔ بیویوم و کاسیدی [۳] در بخش «شرح و حواشی»، قسمتهای مربوط به میدانهای دیفرانسیلی عمومی، میدانهای دیفرانسیلی بسته، و بستار دیفرانسیلی نیز درآمد خوبی به این موضوع است. البته خواننده می‌تواند مطالعه را با مقالهٔ کولچین [۱۰] شروع کند که در مقدمه‌اش این هدف مطرح می‌شود: «توصیفی یکپارچه از تمامی موضوع در حال و هوا و به زبان جبر دیفرانسیلی با هدف آشکارکردن هر چه صریحتر ایده‌های مربوط و روابط میان این ایده‌ها.»

مقالهٔ بیویوم و کاسیدی از این نیز فراتر می‌رود و حاوی درآمدی بر دیگر درونمایهٔ اصلی آثار کولچین یعنی نظریهٔ گروههای جبری دیفرانسیلی و هندسهٔ جبری دیفرانسیلی است. این مجلد همچنین شامل خلاصهٔ «سخنرانیهای سمینار» کولچین در انجمن ریاضی آمریکا در سال ۱۹۷۵ در بارهٔ گروههای جبری دیفرانسیلی و مقالهٔ ارائه‌شدهٔ او به سال ۱۹۷۸ در آکادمی علوم شوروی و نیز یادداشتهای درسهای او در پکن در سال ۱۹۸۴ در همان موضوع می‌باشد. صورت قطعی و نهایی کارهای او در این زمینه در کتاب گروههای جبری دیفرانسیلی[۱] [۷] به چاپ رسیده است. در مقالهٔ بیویوم و کاسیدی، درآمد بسیار خوبی به این موضوع، و نیز به موضوع وسیعتر هندسهٔ جبری دیفرانسیلی، عرضه شده است، آن هم به زبانی که خوانندهٔ ناآشنا به جزئیاتِ نظریهٔ کولچین می‌تواند آن را بفهمد. مقالهٔ فوق، علاوه بر فواید دیگر، خواننده را قادر می‌سازد که درک کند چرا کولچین فکر می‌کرد که این جنبه از موضوع باید مهم باشد و همچنین پی به اهمیت دستاوردهای تکمیل‌کننده و کاربردی بیویوم و کاسیدی ببرد.

۱. این ناقد که کتاب مزبور را هم در منمتیکال ریویوز و هم در بستن‌البلات فورماتماتیک نقد کرد، نوشت که «این کتابِ مشکلی است»؛ و هنوز هم چنین است.

جبر دیفرانسیلی که شامل نظریهٔ گالوای دیفرانسیلی است، در سالهای اخیر رشد یافته و وسیعتر شده است. حتی برشمردن تحقیقات فعلی در مباحث خاصِ مورد علاقهٔ کولچین نیز فراتر از مجال نقد حاضر است. خوشبختانه مقالات سینگر، پوزا، بیویوم و کاسیدی در «شرح و حواشی» به همراه مجموعاً ۳۴۰ عنوان در بخش کتابشناسی، این کار را به‌خوبی انجام می‌دهند و البته کارهای بسیار دیگری نیز هست که به آنها ارجاع نشده است. با این همه، مایلم مطلب را با ذکر دو نکته در بارهٔ رویکردهای مختلف به نظریهٔ گالوای دیفرانسیلی خاتمه دهم.

به‌نظر می‌رسد توافقی عام در بین محققان این رشته وجود دارد که نویدبخش‌ترین روش بررسی نظریهٔ پیکار-وسیو عبارت است از روش دُلینی، از طریق نظریهٔ رسته‌های تاناکایی. در اینجا فکر اصلی عبارت است از ساختن یک توسیع پیکار-وسیو از میدان $F$ برای عملگر دیفرانسیلی $L$ روی $F$ به این طریق: ابتدا در تناظر با $F$ و $L$، یک $F$-فضای برداری متناهی‌بعد $W$ با یک هموستار $\bigtriangledown$ در نظرگرفته می‌شود، به طوری که هستهٔ $\bigtriangledown$ مجموعهٔ پُری از جوابهای $L = O$ برای یک پایه باشد. سپس زیررستهٔ تانسوردارِ رستهٔ تانسوردارِ $F$-فضاهای برداری متناهی‌بعد با هموستارهای تولیدشده توسط $(W, \bigtriangledown)$، در نظر گرفته می‌شود. این زیررسته، تاناکایی است و بنابراین یک گروه جبری خطی وجود دارد که به‌طور متعارف با آن متناظر است، و معلوم شده که گروه گالوای دیفرانسیلی است و توسیع پیکار-وسیو میدان تابعی این گروه است. حتی می‌توان عملگر $L$ را حذف نمود و تنها با $F$-فضایی برداری همراه با هموستار شروع کرد. مقالهٔ سینگر در این مجلد شامل یک نمای کلی از این نظریه است. در نقد برتران [۱] هم چشم‌اندازی کلی آمده است. خواننده‌ای که زبان و اصطلاحات طرحها[۱] برایش راحت‌تر است، می‌تواند به مقالهٔ اولیهٔ دُلینی [۴] مراجعه کند. با این حال، در بیشتر شروحِ مبسوطِ نظریهٔ پیکار-وسیو که برای دانشجویان نگاشته می‌شوند، مانند [۱۱]، از روشهای ماقبل تاناکایی استفاده می‌شود.

گروههای جبریِ متناظر با رسته‌های تاناکایی، خطی‌اند و بنابراین رویکرد فوق به نظریهٔ پیکار-وسیو محدود می‌شود.

موضوعی که تا به امروز کمتر کاویده شده، عبارت است از ارتباط مابین نظریهٔ عام گالوا در رسته‌ها و نظریهٔ گالوای دیفرانسیلی. چند سال قبل، یانه‌لیدزه [۵] نشان داد که نظریهٔ گالوای دیفرانسیلی پیکار-وسیو حالت خاصی از نظریهٔ گالوا در رسته‌هاست و با استفاده از آن روش ساختی برای گروه گالوای دیفرانسیلیِ در حالت پیکار-وسیو ارائه کرد. محتمل است که این رویکرد نظریهٔ قویاً نرمال را نیز شامل شود.

## مراجع


1. Bertrand, D., Review of Lectures on Differential Galois Theory by A. Magid, Bull. Amer. Math. Soc., vol. 33, 1996, pp. 289-294.

2. Borel, A., Algebraic groups and Galois theory in the work of Ellis R. Kolchin, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 505-526. CMP 99:09


1. schemes

pp. 141-170. MR **49**:4982

11. Magid, A., Lectures on differential Galois theory, University Lecture Series, 7. Second printing with corrections, American Mathematical Society, Providence, RI, 1994. MR **95j**:12008

12. Martinet, J. and Ramis, J.-P., Théorie de Galois différentielle et resommation, in Computer Algebra and Differential Equations, ed. Tournier, E., Academic Press, New York, 1990, pp. 117-214. MR **91d**:12014

13. Mitschi, C. and Singer, M., Connected linear groups as differential Galois groups, J. Algebra, vol. **184**, 1996, pp. 333-361. MR **97g**:12004

14. Poizat, B., Les corps différentiellment clos, compagnons de route de la théorie des modéles, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 555-566. CMP 99:09

15. Singer, M., Direct and inverse problems in differntial Galois theory, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 527-554. CMP 99:09

16. Tretkoff, C. and Tretkoff, M., Solution of the inverse problem of differential Galois theory in the classical case, Amer. J. Math. vol. 101, 1979, pp. 1327-1332. MR **80k**:12033

17. Weil, A., Foundations of Algebraic Geometry, A.M.S. Colloquium Publ. 29, American Mathematical Society, New York (1946). MR **9**:303c

******

• Andy R. Magid, *Bull. Amer. Math. Soc. (N.S.)*, (3) **37** (2000) 337-342.

* اندی مجید، دانشگاه اُکلاهما، آمریکا

amagid@ou.edu

3. Buium, A. and Cassidy, P., Differential algebraic geometry and differential algebraic groups: From algebraic differential equations to Diophantine geometry, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society. Providence, 1999, pp. 567-636. CMP 99:09

4. Deligne, P., Catégories tannakiennes, in The Grothendieck Festschrift Volume II, Cartier, P., et al., Birkhäuser, Boston, 1990, pp. 111-196. MR **92d**:14002

5. Janelidzhe, G., Galois theory in categories: the new example of differential fields, in Categorical Topology and Its Relation to Analysis Algebra and Combinatorics: Prague, Czechoslovakia, 22-27 August 1988, eds. Adamek, J. and Mac Lane, S., World Scientific (1989), pp. 369-380.

6. Kolchin, E., Differential Algebra and Algebraic Groups, Academic Press, New York (1973). MR **58**:27929

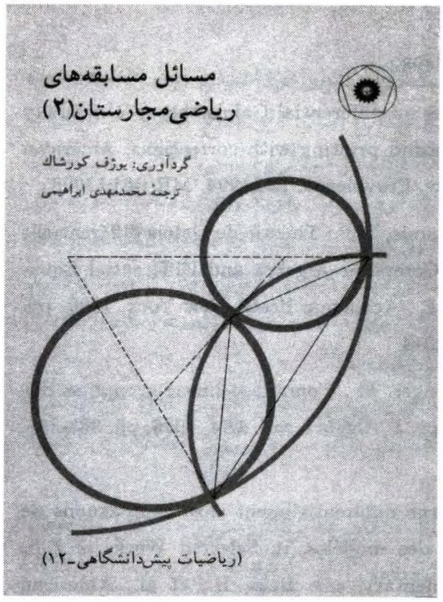
7. ———, Differential Algebraic Groups, Academic Press, New York (1985). MR **87i**:12016

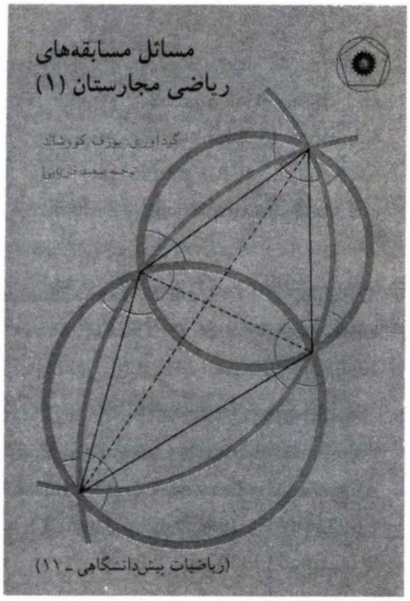
8. ———, Algebraic matric groups and the Picard-Vessiot theory of homogeneous ordinary differential equations, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 505-526. Originally published in Ann. of Math. vol. 49, 1948, pp. 1-42. MR **9**:561c

9. ———, Galois theory of differential fields, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 505-526. Originally published in Amer. J. Math. vol. 75, 1953, pp. 753-824. MR **15**:394a

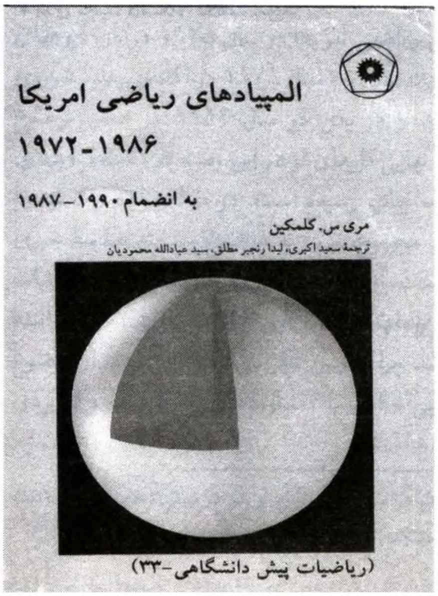
10. ———, Constrained extensions of differential fields, in Selected Works of Ellis Kolchin with Commentary, eds. Bass, H., et al., American Mathematical Society, Providence, 1999, pp. 505-526. Originally published in Adv. in Math. vol. 12, 1974,

# از مجموعهٔ ریاضیات پیش‌دانشگاهی






# NASHR-E RIYĀZI

Volume 12, Numbers 1,2, October 2001



*Nashr-e Riyāzi* is a Persian-language, expository mathematics journal published biannually (in April and October) by Iran University Press.

Annual subscription rates (including airmail postage) are: Middle East £ 19, Europe & Asia £ 21, North America & Far East £ 26.

For more information write to Iran University Press, 85 Park Avenue, Tehrān 15134, Iran, or contact ⟨nashriaz@vax.ipm.ac.ir⟩.

## CONTENTS



* An asterisk indicates that the article was originally published elsewhere. Complete address of the original article appears at the end of the article.

ISSN: 1015-2857

## از کتابهای مرکز نشر دانشگاهی