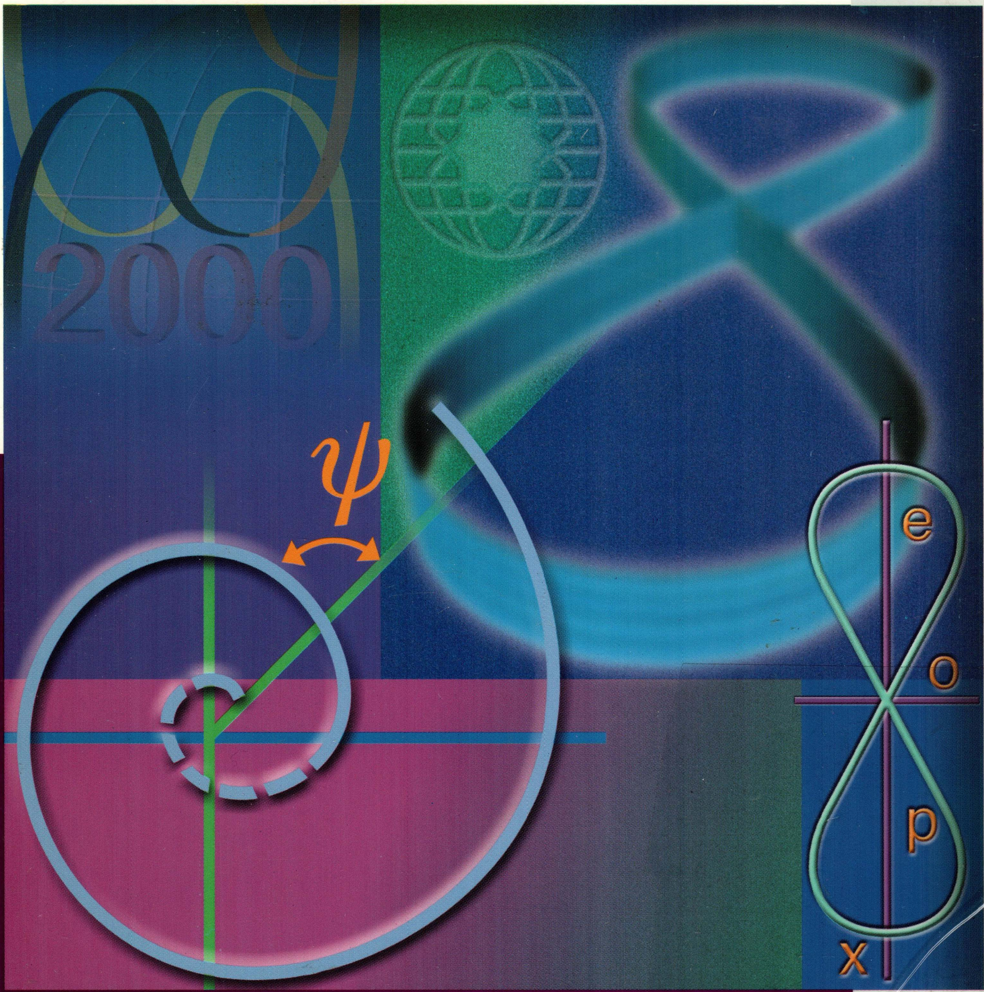
شرکت انتشارات فنی ایران

جوهر، روش و کارایی ریاضیات ۳

(به مناسبت سال جهانی ریاضیات)



تألیف: گروه مؤلفین انستیتوی ریاضی استکلووای شوروی
ترجمه: پرویز شهریاری

جوهر، روش و کارایی ریاضیات ۳

تألیف: گروه مؤلفین انستیتوی ریاضی استکلووای شوروی

ترجمه: پرویز شهریاری

ویرایش: دکتر شهریار شهریاری

شرکت انتشارات فنی ایران

فهرستنویسی پیش از انتشار کتابخانهٔ ملی جمهوری اسلامی ایران

جوهر، روش و کارایی ریاضیات / تألیف گروه مؤلفین انستیتوی ریاضی استکلووای شوروی ؛ ترجمه پرویز شهریاری ؛ ویرایش شهریار شهریاری . ـ تهران: شرکت انتشارات فنی ایران ، ۱۳۸۰ .

۳ ج.

ISBN 964-6232-59-0 (ج . ۱) - ISBN 964-6232-58-2 (دوره) . -

ISBN 964-6232-61-2 (ج . ۳) - ISBN 964-6232-60-4 (ج . ۲)

فهرستنویسی بر اساس اطلاعات فیپا.

۱. ریاضیات. الف. انستیتوی ریاضی استکلووا Matematicheskii institut im . V. A . Steklova ب . شهریاری ، پرویز ، ۱۳۰۵ ـ ، مترجم . ج . عنوان .

QA ۳۷/۲/ ۹ج   ۵۱۰

۱۳۷۹

کتابخانه ملی ایران   ۱۴۶۵۱-۷۸ م

شرکت انتشارات فنی ایران

جوهر، روش و کارایی ریاضیات ۳

تألیف : گروه مؤلفین انستیتوی ریاضی استکلووای شوروی

ترجمه : پرویز شهریاری

ویرایش : دکتر شهریار شهریاری

ناشر : شرکت انتشارات فنی ایران ؛ **تلفن** : ۶۴۹۰۱۴۶، ۶۴۶۲۲۱۸، ۸۷۵۰۴۴۷

نوبت چاپ : اول ۱۳۸۰

تیراژ : ۳۰۰۰ نسخه

امور فنی : محمدعلی رزاقی

حروفچینی : عبدی، شهیر

لیتوگرافی : کورش

چاپ : سعید نو

شابک (دورهٔ ۳ جلدی) : ۹۶۴-۶۲۳۲-۵۸-۲   ISBN: 964-6232-58-2 (3 Vol. Set)

شابک (جلد سوم) : ۹۶۴-۶۲۳۲-۶۱-۲   ISBN: 964- 6232 -61-2 (Vol. 3)

علم چندانکه بیشتر خوانی　　چون عمل در تو نیست نادانی

نه محقق بود نه دانشمند　　چارپایی بر او کتابی چند

آن تهی مغز را چه علم و خبر　　که بر او هیزمست یا دفتر

سعدی

پیش‌گفتار

این کتاب مجموعه‌ای است از آگاهی‌های ریاضی، که آشنایی با آن‌ها، برای هر درس‌خوانده‌ای لازم است. با این که کتاب در سال‌های پنجاه سدهٔ بیستم تنظیم شده و به‌طور طبیعی شامل پیشرفت‌هایی که در دهه‌های بعد پدید آمده، نیست، ارزش خود را در زمان ما هم حفظ کرده است.

از ویژگی‌های کتاب، جامع بودن و انسانی بودن آن است؛ نویسندگان کتاب، همه جا به تاریخ ریاضیات، فلسفهٔ ریاضیات و کاربردهای ریاضیات توجه داشته‌اند و «ریاضیات» را به‌عنوان مجموعه‌ای یکپارچه و یگانه، و تا آن‌جا که ممکن بوده است با زبانی ساده و غیرتخصصی، بررسی کرده‌اند. همان‌طور که در پیش‌گفتار هیأت تحریریه آمده، کتاب در انستیتوی ریاضیات اتحاد شوروی و با شرکت و بحث دسته‌جمعی بزرگترین ریاضی‌دانان آن زمان شوروی تنظیم شده است و، به‌این ترتیب، از غربال نقد ریاضی‌دانانِ با صلاحیت گذشته است.

تنها دربارهٔ فصل چهاردهم (جلد دوم)، با توجه به پیشرفت‌های حیرت‌آوری که در دهه‌های اخیر در ساختمان رایانه‌ها به‌وجودآمده، باید گفت، تنها ارزش تاریخی دارد و می‌تواند خواننده را با موقعیت کار رایانه‌ها در آغاز پیدایش خود آشنا کند که، به نوبهٔ خود، برای هر کسی جالب است. دیگر فصل‌ها به‌طور کامل قابل استفاده است و اگر لازم به‌یادآوری نکته یا نکته‌هایی بوده، ویراستار محترم ترجمهٔ فارسی، به آن‌ها اشاره کرده است.

کتاب بعد از انتشار، بلافاصله به بیشتر زبان‌های زندهٔ دنیا برگردانده شد و به‌عنوان یکی از کتاب‌های مرجع در اختیار دانش‌پژوهان قرار گرفت. در ایران، برای نخستین‌بار، فصل‌های

اول و دوم جلد اول در سال‌های اول دههٔ چهل خورشیدی و سپس ترجمهٔ جلدهای اول و دوم کتاب در سال ۱۳۵۶ منتشر شد و اینک به همت شرکت انتشارات فنی ایران، هر سه جلد در اختیار خوانندهٔ علاقه‌مند قرارمی‌گیرد.

از پسرم شهریار که با دقت وسواس‌گونه‌ای تمامی ترجمه را، ضمن تطبیق آن با اصل کتاب، خوانده‌اند و هرجا لازم بوده است، یادداشتی در پاورقی افزوده‌اند، بسیار سپاس دارم؛ بی‌تردید بدون همراهی او، خواننده نمی‌توانست کتابی به این صورت در اختیار داشته باشد.

پرویز شهریاری

فهرست

عنوان	صفحه

عنوان	صفحه

پیش‌گفتار

ریاضیات، در دوران باستان در بستگی با نیازهای زندگی پدید آمد، و به تدریج به دستگاهی از دانش‌های گوناگون تبدیل شد. ریاضیات نیز، همچون سایر دانش‌ها بازتابی از قانون‌های طبیعت است و به‌عنوان سلاح نیرومندی برای شناخت طبیعت و پیروزی بر آن به‌کار می‌رود. ولی از آن‌جا که ریاضیات بیش از اندازه انتزاعی و ذهنی است، رشته‌های جدید آن برای کسانی که ویژه کار نیستند تا اندازهٔ زیادی قابل دسترس نیست. همین ویژگی انتزاعی بودن ریاضیات، از روزگاران باستان، پندارهای ذهن‌گرایانه دربارهٔ بی‌ارتباطی آن با طبیعت به‌وجود آورد.

هدف نویسندگان کتاب حاضر این بوده است که گروه گسترده‌ای از روشنفکران را با جوهر و روش بخشهای ریاضیات و پایه‌های عملی و شیوه‌های پیشرفت آن آشنا سازند.

کمترین آگاهی‌های ریاضی که برای خوانندهٔ این کتاب فرض‌شده، ریاضیات دورهٔ دبیرستانی است. با وجود این، نوشته‌های این کتاب یک‌دست نیست. کسانی که می‌خواهند به مقدمه‌های ریاضیات عالی آشنا شوند، می‌توانند از چند بخش اول کتاب استفاده کنند، ولی برای فهم دقیق بخش‌های بعد، باید با کتابهای درسی مربوط آشنا باشند. مجموعهٔ کتاب برای خوانندگانی قابل استفاده است که به‌کاربرد روش‌های آنالیز ریاضی (محاسبهٔ دیفرانسیلی و انتگرالی) آشنا باشند. برای کسانی که با دانش‌های طبیعی سر و کار دارند (مهندسان رشته‌های مختلف و معلمان ریاضی) بخش‌های ویژه‌ای وجود دارد که آنان را با رشته‌های جدیدتر ریاضی آشنا می‌کند.

طبیعی است در چارچوب یک کتاب نمی‌توان حتی همهٔ روش‌های مختلف ریاضی را بررسی کرد و بنابراین برای گزینش مطلب‌های کتاب، نوعی آزادی عمل لازم بوده است. ولی در خط‌های کلی‌تر، این کتاب روی‌هم اندیشهٔ درستی دربارهٔ وضع کنونی ریاضی،

پیدایش و پیشرفت آن به دست می‌دهد. کتاب تا اندازه‌ای برای کسانی هم در نظر گرفته شده است که به موضوع‌های اساسی کتاب آشنایی کامل داشته باشند. همچنین کتاب باید بتواند بعضی از تنگ‌نظری‌هایی را هم که گاهی بعضی از ریاضی‌دان‌های جوان ما از خود نشان می‌دهند، برطرف کند.

بخش‌های جداگانهٔ کتاب به وسیلهٔ نویسندگان مختلف نوشته شده است. با وجود این مجموعهٔ کتاب نتیجهٔ کوشش دسته‌جمعی است. طرح کلی و گزینش موضوع‌های بخش‌های گوناگون آن در معرض دید و داوری دسته‌جمعی قرار گرفته و براساس تبادل زندهٔ اندیشه‌ها تکمیل شده است. ریاضی‌دان‌های بسیاری از شهرهای اتحاد شوروی، در بحثی که به وسیلهٔ انستیتوی ریاضی ترتیب داده شده بود شرکت کردند و دربارهٔ متن مقدماتی آن یادآوری‌های پر ارزشی کردند و البته این یادآوری‌ها و پیشنهادها، از طرف نویسندگان در نظر گرفته شده است.

همچنین برخی از نویسندگان در تهیهٔ متن آخر بخش‌هایی که به وسیلهٔ نویسندگان دیگر نوشته شده است به طور مستقیم شرکت کردند. بخش اساسی مقدمهٔ بخش دوم را ***ب. ن. دِلون*** (*B. N. Delon*) نوشته و ***د. ک. فادیف*** (*D.K. Fadeev*) در تهیهٔ بخش‌های چهارم و بیستم کوشش فراوان کرده است.

همچنین عده‌ای از کسانی که بخش جداگانه‌ای تهیه نکرده‌اند، در کار تهیهٔ کتاب مشارکت داشته‌اند: ***ل. ب. کانتارویچ*** (*L. B. Kontorovitch*) بند چهارم از بخش چهاردهم و ***ا. آ. لادیژنسکایا*** (*O. A. Ladigenskaia*) بند ششم از بخش ششم و ***آ. گ. پاستنیکوف*** (*A. G. Postnikov*) بند پنجم از بخش دهم را نوشته‌اند، و ***ا. آ. اُله نیک*** (*O. A. Oleinik*) در تهیهٔ متن بخش پنجم و ***یو. و. پروخوروف*** (*U. V. Prokhorov*) در آخرین بررسی متن بخش یازدهم شرکت داشته‌اند.

و. آ. زالگالر (*V. A. Zalgaler*) بعضی از بخش‌های اول و دوم و هفتم و هفدهم را نوشته است و آخرین بررسی متن کتاب به وسیلهٔ ***زالگالر*** و ***ویدنسکی*** (*Videnski*) و با شرکت ***ت. و. راگازکینا*** (*T. V. Rogozkina*) و ***آ. پ. لئونوا*** (*A. P. Leonova*) انجام گرفته است. بخش اساسی تصویرها به وسیلهٔ ***ا. پ. سنکینی*** (*A. P. Senkini*) رسم شده است.

هیأت تحریریه

بخش پانزدهم

نظریهٔ تابع‌های با متغیر حقیقی

س. ب. سِتچکین

۱. ورود به مطلب

در پایان سدهٔ هجدهم و آغاز سدهٔ نوزدهم ، ساختمان حساب دیفرانسیل و انتگرال (حسابان، کم و بیش آماده شده بود؛ پیش از آن (و درواقع، در تمامی سدهٔ هجدهم)، دانشمندان به ساختن شاخه‌های تازه‌ای از آن مشغول بودند، حقیقت‌های تازه‌ای کشف می‌کردند، جنبه‌های تازه‌تری از کاربرد محاسبهٔ دیفرانسیلی و انتگرالی را در مسأله‌های گوناگونِ مکانیک، اخترشناسی و صنعت جست‌وجو می‌کردند و تکامل می‌دادند. به‌تدریج این امکان پدید آمد که به بررسی نتیجه‌گیری‌ها بپردازند، آن‌ها را منظم کنند و بررسی دقیق مفهوم‌های اساسی آنالیز را در برنامهٔ کار خود قرار دهند؛ خیلی زود روشن شد، پایه‌ها و بنیان‌های آنالیز ریاضی، چندان استوار نیست.

در سدهٔ هجدهم، بزرگترین دانشمندانِ زمان، حتی در این باره که «تابع چیست»، هم‌رأی نبودند. این وضع، بحثی دراز را دربارهٔ درستی یا نادرستیِ این یا آن راه‌حل مسأله و درستی یا نادرستیِ این یا آن نتیجه‌گیری مشخص ریاضی به‌دنبال داشت. به‌تدریج معلوم شد، باید برای دقیق‌تر کردن مفهوم‌های اساسی آنالیز، تلاش کرد. نارسایی درک به‌ظاهر روشن و شهودی از مفهوم «پیوستگی» و «ویژگی‌های تابع پیوسته»، موجب تصورهای نادرستی از این‌گونه شده بود که: تابع‌های پیوسته همیشه دیفرانسیل‌پذیرند. ریاضیات، با چنان تابع‌های پیچیده‌ای کار می‌کرد که دیگر نمی‌شد بر روشنی مطلب و درک شهودی تکیه کرد. فراهم کردن نظم و ترتیبی در مفهوم‌های اساسی آنالیز، ضرورت جدی پیدا کرده بود.

در این مسیر، نخستین گام جدی را ***لاگرانژ*** برداشت و **کوشی** راه او را ادامه داد. کوشی تعریف مفهوم‌های حد، پیوستگی و انتگرال را دقیق کرد، به‌نحوی که تا امروز هم درستیِ خود را حفظ کرده‌اند. کم و بیش در همان زمان، ***بُولتسانو*** ریاضی‌دان چک، به بررسی ژرفی

در ویژگی‌های اساسیِ تابع‌های پیوسته پرداخت.

دربارهٔ این ویژگی‌های تابع‌های پیوسته، بیشتر صحبت می‌کنیم. تابع پیوسته $f(x)$ را در نظر بگیرید که در بازهٔ بستهٔ $[a, b]$، یعنی به‌ازای همهٔ عددهایی که در نابرابری $a \le x \le b$ صدق می‌کنند، داده شده باشد. پیشتر، این حکم را روشن می‌پنداشتند: اگر تابع در دو انتهای دامنهٔ تعریف خود، یعنی برای $x = a$ و $x = b$، مقدارهایی را بپذیرد که علامت‌های مختلفی داشته باشند، در نقطه‌ای بین دو مقدار a و b برابر صفر می‌شود. اکنون، این حقیقت بُرهانی دقیق پیدا کرد؛ درضمن، این حکم هم بادقت ثابت شد که، اگر تابعی در یک بازهٔ بسته پیوسته باشد، در نقطه‌هایی از این بازه، به بیشترین و کمترین مقدار خود می‌رسد.

بررسی این ویژگی‌های تابع‌های پیوسته، ناگزیر منجر به بررسی دقیق‌تری از طبیعتِ عددهای حقیقی می‌شد. از این‌جا بود که نظریهٔ عددهای حقیقی شکل گرفت و ویژگی‌های اساسی محور عددی، تنظیم و روشن شد.

تکامل بعدیِ آنالیز ریاضی، ضرورت بررسی «ناهنجارترین» تابع‌ها، به‌ویژه تابع‌های ناپایدار و ناپیوسته را به‌دنبال داشت. تابع‌های ناپیوسته، از جمله به‌عنوان حد تابع‌های پیوسته پدید می‌آیند؛ درضمن از قبل معلوم نیست، آیا تابعِ حدی پیوسته است یا ناپیوسته. همچنین در روندی، که شامل تغییری تند و ناگهانی باشد، با تابع ناپیوسته سروکار داریم. مسألهٔ تازه‌ای در برابر قرار می‌گیرد: تعمیم دستگاه آنالیز به تابع‌های ناپایدار و ناپیوسته.

ریمان این موضوع را بررسی کرد که مفهوم انتگرال را، دربارهٔ چه گروهی از تابع‌های ناپیوسته می‌توان به‌کار برد. نتیجهٔ همهٔ این تلاش‌ها در پایه‌گذاری آنالیز، به این‌جا رسید که نظام ریاضی تازه‌ای پدید آمد: نظریهٔ تابع‌های با متغیر حقیقی.

اگر آنالیز ریاضی سنتی، در اساس بر تابع‌های «به‌هنجار» (مثل تابع‌های پیوسته، دیفرانسیل‌پذیر) تکیه دارد، نظریهٔ تابع‌های با متغیر حقیقی، گروه بسیار گسترده‌تری از تابع‌ها را در بر می‌گیرد. اگر در آنالیز ریاضی، تعریف برخی عمل‌ها (مثل انتگرال‌گیری) برای تابع‌های پیوسته داده می‌شود، نظریهٔ تابع‌های با متغیر حقیقی، در این باره بررسی می‌کند که این تعریف را برای چه گروهی از تابع‌ها می‌توان به‌کار برد و چگونه می‌توان با بازسازی تعریف، صوت عام‌تر و گسترده‌تری به آن داد. به‌ویژه، نظریهٔ تابع‌های با متغیر حقیقی می‌تواند به این پرسش پاسخ دهد که «طول منحنی» یعنی چه و دربارهٔ کدام منحنی‌ها می‌توان از مفهوم «طول» استفاده کرد.

خودِ نظریهٔ تابع‌های با متغیرحقیقی، براساس و مبنای نظریهٔ مجموعه‌هاساخته‌شده‌است.

به همین دلیل، بحث خود را از بررسی عنصرهای نظریهٔ مجموعه‌ها آغاز می‌کنیم؛ سپس به بررسی مجموعهٔ نقطه‌ها می‌پردازیم و بخش را با یکی از مفهوم‌های اساسی نظریهٔ تابع‌های با متغیر حقیقی، یعنی «انتگرال *له‌بگ*» به پایان می‌بریم.

۲. مجموعه‌ها

انسان به‌طور دایم با گروه‌های گوناگونی از چیزها سر و کار دارد. همان‌طور که در بخش اول (جلد اول) روشن کردیم، همین امر موجب پیدایش مفهوم «عدد» و سپس مفهوم «مجموعه» شده که یکی از ساده‌ترین مفهوم‌های بنیانی ریاضیات است و به تعریف دقیق تن نمی‌دهد. هدفِ آن‌چه در این‌جا می‌آوریم، درک مفهوم مجموعه است، بدون این که مدعی تعریف آن باشیم.

مجموعه، به دسته، گروه و اجتماعی از چیزها گفته می‌شود که به‌وسیلهٔ نشانه یا قانونی به هم مربوط باشند. مجموعه، یک مفهوم انتزاعی است. وقتی دسته‌ای از چیزها را، به عنوان مجموعه در نظر می‌گیریم، همهٔ بستگی‌ها و رابطه‌های بین چیزهای مختلفی که مجموعه را تشکیل داده‌اند، کنار می‌گذاریم و خصلت‌های فردیِ آن‌ها را نگه می‌داریم. بنابراین مجموعه‌ای که از پنج سکه تشکیل شده و مجموعه‌ای که شامل پنج سیب است، دو مجموعهٔ مختلف‌اند. ولی مجموعهٔ پنج‌سکه‌ای، که دایره‌وار روی میز چیده شده باشند، با مجموعهٔ همین سکه‌ها که روی هم قرار گرفته باشند، دو مجموعهٔ متفاوت نیستند.

چند مثال برای مجموعه‌ها می‌آوریم. می‌توان دربارهٔ مجموعهٔ شن‌های یک تودهٔ ماسه، مجموعهٔ همهٔ سیاره‌های دستگاه خورشیدی، مجموعهٔ همهٔ کسانی که در لحظهٔ مفروض در خانه‌ای هستند و یا مجموعهٔ همهٔ صفحه‌های این کتاب صحبت کرد. در ریاضیات هم، همیشه با مجموعه‌های مختلفی روبه‌رو هستیم، مثل مجموعهٔ همهٔ ریشه‌های یک معادله، مجموعهٔ همهٔ عددهای طبیعی، مجموعهٔ همهٔ نقطه‌های روی یک خط راست وغیره.

آن بخش از ریاضی، که ویژگی‌های کلی مجموعه‌ها را بررسی می‌کند، یعنی ویژگی‌هایی از مجموعه را که ربطی به طبیعتِ عضوهای تشکیل‌دهندهٔ آن ندارد، **نظریهٔ مجموعه‌ها** نامیده می‌شود. پایان سدهٔ نوزدهم و آغاز سدهٔ بیستم، شاهد پیشرفت بی‌اندازهٔ این بخش ریاضیات بود. بنیان‌گذار نظریهٔ علمی مجموعه‌ها، **ژرژ کانتور** ریاضی‌دان آلمانی است.

کارهای ***کانتور*** در زمینهٔ نظریهٔ مجموعه‌ها، براساس مسأله‌های مربوط به هم‌گراییِ رشته‌های مثلثاتی رشد کرد. و این، پدیده‌ای عادی است: بسیار پیش آمده است که بررسیِ مسأله‌های مشخصی از ریاضیات، به تنظیم و آرایش نظریه‌های کلی و انتزاعی تازه‌ای منجر شده است. اهمیت این ساختمان‌های انتزاعی در این است که بستگی خود را با همان مسأله‌هایی که از آن‌ها برآمده‌اند، نگه می‌دارد، بلکه در ضمن کار خود را به یک‌رشته از موضوع‌های دیگر هم، سرایت می‌دهند. نظریهٔ مجموعه‌ها هم برکنار از این وضع نبود. اندیشه‌ها و مفهوم‌های نظریهٔ مجموعه‌ها، در همهٔ شاخه‌های ریاضیات نفوذ کرد و به‌واقع، چهرهٔ آن را دگرگون ساخت. به همین مناسبت، بدون آشنایی با عنصرهای نظریهٔ مجموعه‌ها، نمی‌توان دورنمایی درست از ریاضیات امروز نشان داد. نظریهٔ مجموعه‌ها، به‌ویژه، برای نظریهٔ تابع‌های با متغیر حقیقی، اهمیت زیادی دارد.

مجموعه را وقتی می‌توان مفروض دانست که دربارهٔ هر چیزی، بتوان وابستگی یا عدم وابستگی آن را به مجموعه مشخص کرد. به زبان دیگر، هر مجموعه، با معلوم بودن همهٔ عنصرها یا چیزهایی که به آن تعلق دارند («عضوهای» مجموعه) مشخص می‌شود. وقتی مجموعهٔ M از عضوهای a، b، c، ... ، و تنها از همین عضوها تشکیل شده باشد، می‌نویسند:

$$M = \{ a , b , c , \ldots \}$$

بنابراین، ***عضوهای یک مجموعه***، یعنی عنصرهایی که به این مجموعه بستگی دارند. این حقیقت را که عنصر m، عضوی از مجموعهٔ M است، این‌طور نشان می‌دهند:

$$m \in M$$

و می‌خوانند: «m متعلق به M است» یا «m عضوی از مجموعهٔ M است». ولی اگر عنصر n متعلق به مجموعهٔ M نباشد، می‌نویسند:

$$n \bar{\in} M \quad \text{یا} \quad n \notin M$$

هر عنصر n، می‌تواند تنها یکی از عضوهای مجموعهٔ M باشد؛ به‌زبان دیگر، همهٔ عضوهای یک مجموعه، متفاوت‌اند و از هم تمیز داده می‌شوند.

هر عضو مجموعهٔ M، خود می‌تواند یک مجموعه باشد، ولی برای پرهیز از تناقض،

می‌پذیریم که خودِ مجموعهٔ M، عضوی از خودش نیست:

$$M \bar{\in} M$$

مجموعه‌ای که هیچ عضوی نداشته باشد، **مجموعهٔ تهی** نامیده می‌شود. برای نمونه، مجموعهٔ همهٔ ریشه‌های حقیقی معادلهٔ

$$x^۲ + ۱ = ۰$$

یک مجموعهٔ تهی است. بعد از این، مجموعهٔ تهی را با نماد θ نشان می‌دهیم.

اگر برای دو مجموعهٔ M و N، هر عضو x از مجموعهٔ M، عضوی از مجموعهٔ N هم باشد، گویند: M در N جای دارد، یا M جزئی از مجموعهٔ N است و یا، آن‌گونه که معمول است: M زیرمجموعه‌ای از N است؛ این حقیقت را به این صورت نشان می‌دهند:

$$M \subseteq N \quad \text{یا} \quad N \supseteq M$$

به‌عنوان نمونه، مجموعهٔ $M = \{۱, ۲\}$، زیرمجموعه‌ای از مجموعهٔ $N = \{۱, ۲, ۳\}$ است.

روشن است، همیشه داریم: $M \subseteq M$. برای سادگی کار، مجموعهٔ تهی را، زیرمجموعهٔ هر مجموعهٔ دل‌خواه، به‌حساب می‌آوریم.

دو مجموعه را *برابر* گوییم، وقتی عضوهای مجموعهٔ اول، همان عضوهای مجموعهٔ دوم، و برعکس، عضوهای مجموعهٔ دوم همان عضوهای مجموعهٔ اول باشد. برای نمونه، مجموعهٔ ریشه‌های معادلهٔ $x^۲ - ۳x + ۲ = ۰$ و مجموعهٔ $M = \{۱, ۲\}$ برابرند.

عمل‌های مختلفی روی مجموعه‌ها تعریف می‌کنیم:

اجتماع یا مجموع. مجموعه‌های M، N، P، ... را در نظر می‌گیریم. اجتماع یا مجموع این مجموعه‌ها به مجموعه‌ای مثل X گفته می‌شود که هر عضو آن، دست‌کم در یکی از «جمله‌های جمع» M، N، P، ... وجود داشته باشد و به‌جز عضوهای «جمله‌های جمع» عضو دیگری نداشته باشد:

$$X = M \cup N \cup P \cup \cdots \quad \text{یا} \quad X = M + N + P + \cdots$$

درضمن، اگر عضو x متعلق به چند «جملهٔ جمع» باشد، تنها یک‌بار در مجموع X وارد

می‌شود. روشن است:

$$M + M = M$$

و اگر داشته باشیم: $M \subseteq N$، آن‌گاه

$$M + N = N$$

اشتراک یا فصل مشترک. اشتراک یا فصل مشترک مجموعه‌های M، N، P، ... به مجموعهٔ Y گفته می‌شود که هریک از عضوهای آن، درعین‌حال، به همهٔ مجموعه‌های M، N، P، ... تعلق داشته باشد.

اشتراک مجموعه‌ها را، گاه حاصل‌ضرب آن‌ها هم می‌گویند. روشن است که:

$$M \,.\, M = M$$

و اگر داشته باشیم: $M \subseteq N$، آن‌گاه

$$M \,.\, N = M$$

اگر اشتراک دو مجموعهٔ M و N تهی باشد: $M.N = \theta$، گویند دو مجموعه فصل مشترک ندارند، یا دو مجموعه متقاطع نیستند.

برای نشان دادن عمل‌های اجتماع و اشتراک مجموعه‌ها، نمادهای Σ و Π هم‌استفاده می‌کنند، به این ترتیب

$$E = \sum E_i$$

به‌معنای اجتماع مجموعه‌های E_i؛ و

$$F = \prod E_i$$

به‌معنای اشتراک مجموعه‌های E_i.

از خواننده می‌خواهیم ثابت کند، اجتماع و اشتراک مجموعه‌ها، با قانون عادی توزیع‌پذیری

$$M\,(N + P) = MN + MP$$

و همچنین با قانون

$$M + NP = (M + N)\,(M + P)$$

به‌هم مربوط‌اند.

تفاضل. تفاضل دو مجموعهٔ M و N، به مجموعهٔ Z گفته می‌شود که شامل همهٔ عضوهایی از M باشد که به N تعلق ندارند:

$$Z = M - N$$

اگر داشته باشیم: $N \subseteq M$، تفاضل $Z = M - N$ را **متمم مجموعهٔ** N نسبت به M گویند.

به‌سادگی می‌توان ثابت کرد که این دو برابری همیشه برقرار است:

$$M\,(N-P) = MN - MP$$

و

$$(M-N) + MN = M$$

به این ترتیب، قانون‌های عمل دربارهٔ مجموعه‌ها، با قانون عادی حساب، بسیار متفاوت است.

مجموعه‌های باپایان و مجموعه‌های بی‌پایان. مجموعه‌ای را که دارای تعداد محدود و معینی عضو باشد، مجموعهٔ باپایان (یا مجموعهٔ متناهی) گویند. اگر تعداد عضوهای مجموعه، نامحدود باشد، مجموعه را بی‌پایان (یا نامتناهی) گویند. برای نمونه، مجموعهٔ همهٔ عددهای طبیعی، مجموعه‌ای بی‌پایان است.

دو مجموعهٔ M و N را در نظر می‌گیریم و این پرسش را مطرح می‌کنیم: آیا تعداد عضوها در این دو مجموعه، یکی است؟

اگر M مجموعه‌ای با پایان باشد، تعداد (یا کمیت) عضوهای آن، با یک عدد طبیعی ـ تعداد عضوهای مجموعه ـ معین می‌شود. در این حالت، برای مقایسهٔ کمیت عضوهای دو مجموعهٔ M و N، کافی است تعداد عضوهای M و تعداد عضوهای N را معین و این دو عدد را با هم مقایسه کنیم. همچنین طبیعی است اگر یکی از دو مجموعهٔ M و N باپایان و دیگری بی‌پایان باشد، مجموعهٔ بی‌پایان را دارای عضوهای بیشتری نسبت به مجموعهٔ باپایان بدانیم.

ولی در حالتی که هر دو مجموعهٔ M و N بی‌پایان باشند، از راه شمارش سادهٔ عضوها به جایی نمی‌رسیم. بنابراین، بلافاصله، این پرسش پیش می‌آید: آیا تعداد عضوها، در همهٔ

مجموعه‌های بی‌پایان یکی است، یا مجموعه‌های بی‌پایانی وجود دارد که تعداد و عضوها در یکی بیشتر و در دیگری کمتر است؟ اگر حالت دوم درست است، با چه روشی می‌توان، کمیت عضوها را، در مجموعه‌های بی‌پایان، با هم مقایسه کرد؟ در این‌جا، به همین موضوع می‌پردازیم.

تناظر یک‌به‌یک. دوباره دو مجموعهٔ باپایانِ M و N را در نظر می‌گیریم. چگونه می‌توان بدون شمردن تعداد عضوها در هر مجموعه، متوجه شد، کدام‌یک از آن‌ها، تعداد بیشتری عضو دارد؟ برای این منظور می‌توان *زوج‌هایی* تشکیل داد، به‌نحوی که در هر زوج، یک عضو M و یک عضو N در کنار یکدیگر باشند. در این صورت اگر عضوی از مجموعهٔ M نتواند با عضوی از N جفت شود، به‌معنای آن است که تعداد عضوهای M بیشتر از تعداد عضوهای N است. این روش را با مثالی روشن می‌کنیم.

فرض کنید در یک سالن تعدادی افراد و تعدادی صندلی وجود داشته باشد. برای این که بدانیم تعداد افراد بیشتر است یا تعداد صندلی‌ها، کافی است از افراد بخواهیم هرکدام روی یک صندلی بنشینند. اگر کسی یا کسانی بدون صندلی بمانند به‌معنای آن است که تعداد افراد بیشتر است، اگر چند صندلی خالی بماند به این معناست که تعداد صندلی‌ها بیشتر است؛ ولی اگر همه نشسته باشند و همهٔ صندلی‌ها هم اشغال شده باشد، تعداد افراد با تعداد صندلی‌ها برابر است. این روشِ مقایسهٔ تعداد عضوها در مجموعه‌ها، در مقایسه با شمارش مستقیم عضوها، این برتری را دارد که بدون هیچ تفاوت خاصی، می‌توان از آن دربارهٔ مجموعه‌های بی‌پایان هم استفاده کرد.

مجموعهٔ همهٔ عددهای طبیعی

$$M = \{ ۱ , ۲ , ۳ , ۴ , \dots \}$$

و مجموعهٔ همهٔ عددهای زوج

$$N = \{ ۲ , ۴ , ۶ , ۸ , \dots \}$$

را در نظر می‌گیریم. کدام یک از دو مجموعهٔ M و N، عضوهای بیشتری دارد؟ در نگاه اول گمان می‌رود تعداد عضوهای مجموعهٔ اول بیشتر است. با وجود این می‌توانیم شبیه حالت مجموعه‌های پایان، از عضوهای این دو مجموعهٔ بی‌پایان هم، زوج‌هایی تشکیل دهیم.

جدول ۱

M	۱	۲	۳	۴	...
N	۲	۴	۶	۸	...

به این ترتیب، هیچ عضوی از مجموعهٔ M و هیچ عضوی از مجموعهٔ N، تنها نمی‌ماند. البته می‌توانستیم زوج‌ها را این‌طور منظم کنیم:

جدول ۲

M	۱	۲	۳	۴	۵	...
N	–	۲	–	۴	–	...

در این صورت، بسیاری از عضوهای مجموعهٔ M، تنها و بدون جفت می‌ماند. از طرف دیگر، می‌توانستیم زوج‌ها را به این صورت تنظیم کنیم:

جدول ۳

M	–	۱	–	۲	–	۳	–	...
N	۲	۴	۶	۸	۱۰	۱۲	۱۴	...

که در این حالت، بسیاری از عضوهای مجموعهٔ N، تنها و بدون جفت باقی می‌ماند.

به‌این ترتیب، وقتی مجموعه‌های A و B بی‌پایان باشند، با روش‌های مختلف تشکیل زوج‌ها، می‌توان به نتیجه‌های مختلفی رسید. اگر روشی برای تشکیل زوج‌ها وجود داشته باشد که برای آن، هر عضو مجموعهٔ A، جفتی از مجموعهٔ B؛ و برای هر عضو مجموعهٔ B، جفتی از مجموعهٔ A وجود داشته باشد، گویند بین دو مجموعهٔ A و B تناظر یک‌به‌یک وجود دارد، یا عضوهای دو مجموعهٔ A و B، در تناظر یک‌به‌یک هستند. بنابراین، برای دو مجموعهٔ M و N، همان‌طور که دیدیم (جدول ۱)، می‌توان این تناظر یک‌به‌یک را برقرار کرد.

اگر بتوان، بین عضوهای دو مجموعهٔ A و B تناظر یک‌به‌یک برقرار کرد، گویند

تعداد عضوهای دو مجموعه یکی است، یا **دو مجموعه *هم‌توان‌اند*.** ولی اگر با همهٔ روش‌هایِ ممکن برای تشکیل زوج‌ها، از جمله بعضی از عضوهای مجموعهٔ A همواره بدون‌جفتی از مجموعهٔ B بمانند، گویند تعداد عضوهای مجموعهٔ A بیشتر از تعداد عضوهای مجموعهٔ B است، یا مجموعهٔ A نسبت به مجموعهٔ B توان بیشتری دارد.

به این ترتیب، پاسخ یکی از پرسش‌های خود را پیدا کردیم: چگونه می‌توان کمیت عضوها را در مجموعه‌های بی‌پایان مقایسه کرد؟ ولی این پاسخ، ما را به پاسخ پرسش دوم نزدیک نمی‌کند: آیا به‌طور کلی، مجموعه‌های بی‌پایان توان‌های متفاوتی دارند؟ برای رسیدن به پاسخ، چند نمونه از ساده‌ترین مجموعه‌های بی‌پایان را بررسی می‌کنیم.

مجموعه‌های شمارا. اگر بتوان بین عضوهای مجموعهٔ A و عضوهای مجموعهٔ همهٔ عددهای طبیعیِ

$$Z = \{ ۱ , ۲ , ۳ , ۴ , \ldots \}$$

تناظر یک‌به‌یک برقرار کرد، آن‌وقت A را ***مجموعه‌ای شمارا*** می‌نامند. به زبان دیگر، مجموعهٔ A را وقتی شمارا گویند که بتوان عضوهای آن را به‌یاری دنبالهٔ عددهای طبیعی شماره‌گذاری کرد، یعنی بتوان آن را به این صورت نوشت:

$$a_۱ , a_۲ , a_۳ , \ldots , a_n , \ldots$$

جدول ۱ نشان می‌دهد، مجموعهٔ همهٔ عددهای زوج، مجموعه‌ای شمارا است (در این جدول، عدد بالا به‌عنوان شمارهٔ عدد متناظر آن در سطر پایین، در نظر گرفته می‌شود).

مجموعه‌های شمارا، در بین مجموعه‌های بی‌پایان، کوچکترین هستند: در هر مجموعهٔ بی‌پایان، می‌توان زیرمجموعه‌ای یافت که مجموعه‌ای شمارا باشد.

اگر دو مجموعهٔ باپایان، عضوهای مشترک نداشته باشند، مجموع آن‌ها بیشتر از هریک از دو مجموعه، دارای عضو است. ولی این قانون، ممکن است برای مجموعه‌های بی‌پایان درست نباشد. فرض کنید «ز» مجموعهٔ همهٔ عددهای زوج، «ف» مجموعهٔ همهٔ عددهای

فرد و Z مجموعهٔ همهٔ عددهای طبیعی باشد. جدول ۴ نشان می‌دهد مجموعه‌های «ف» و «ز» شمارا هستند؛ با وجود این، مجموعهٔ Z (= «ز» + «ف»)، باز هم مجموعه‌ای شمارا است:

جدول ۴

ز	۲	۴	۶	۸	...
ف	۱	۳	۵	۷	...
Z	۱	۲	۳	۴	...

«خراب شدن» اصل «کُل از جزء خود بزرگتر است» در مجموعه‌های بی‌پایان، نشان می‌دهد ویژگی‌های مجموعه‌های بی‌پایان، با ویژگی‌های مجموعه‌های باپایان، ***تفاوت کیفی*** دارد. گذار از محدود به نامحدود، از متناهی به نامتناهی، از نهایت به بی‌نهایت، گواهی است بر درستی اصل دیالکتیکی تغییر کیفی ویژگی‌ها.

ثابت می‌کنیم ***مجموعهٔ همهٔ عددهای گویا، مجموعه‌ای است شمارا***. برای این منظور، همهٔ عددهای گویا را در جدولی به این صورت تنظیم می‌کنیم:

(۱)	(۲)	(۳)	(۴)	(۵)	(۶)	
۱	۲	۳	۴	۵	۶	...
۰	−۱	−۲	−۳	−۴	−۵	...
$\frac{۱}{۲}$	$\frac{۳}{۲}$	$\frac{۵}{۲}$	$\frac{۷}{۲}$	$\frac{۹}{۲}$	$\frac{۱۱}{۲}$	...
$-\frac{۱}{۲}$	$-\frac{۳}{۲}$	$-\frac{۵}{۲}$	$-\frac{۷}{۲}$	$-\frac{۹}{۲}$	$-\frac{۱۱}{۲}$	...
$\frac{۱}{۳}$	$\frac{۲}{۳}$	$\frac{۴}{۳}$	$\frac{۵}{۳}$	$\frac{۷}{۳}$	$\frac{۸}{۳}$	...
$-\frac{۱}{۳}$	$-\frac{۲}{۳}$	$-\frac{۴}{۳}$	$-\frac{۵}{۳}$	$-\frac{۷}{۳}$	$-\frac{۸}{۳}$	...
...	...	...	...	...	...	...

در جدول، در سطر اول، همهٔ عددهای طبیعی را به ترتیب صعودی قرار داده‌ایم؛ در سطر دوم «صفر» و همهٔ عددهای درست منفی را به ترتیب نزولی؛ در سطر سوم، کسرهای

مثبت و ساده‌نشدنی با مخرج ۲ به ترتیب صعودی؛ در سطر چهارم کسرهای ساده‌نشدنی منفی با مخرج ۲ و به ترتیب نزولی وغیره. روشن است، هر عدد گویا یک‌بار (و تنها یک‌بار) در این جدول آمده است. اکنون همهٔ عددهای این جدول را، به ردیفی که با پیکان نشان داده شده است، شماره‌گذاری می‌کنیم. به این ترتیب، همهٔ عددهای گویا، به ردیف ویژه‌ای یک دنباله را تشکیل می‌دهند:

شمارهٔ عدد گویا	۱	۲	۳	۴	۵	۶	۷	۸	۹	...
عدد گویا	۱	۲	۰	۳	۱−	$\frac{۱}{۲}$	۴	۲−	$\frac{۳}{۲}$	...

به این ترتیب، بین همهٔ عددهای گویا و همهٔ عددهای طبیعی، تناظر یک‌به‌یک برقرار می‌شود؛ یعنی مجموعهٔ همهٔ عددهای گویا، مجموعه‌ای شمارا است.

مجموعه‌های با توان پیوستار. اگر بتوان بین عضوهای مجموعهٔ M و نقطه‌های پاره‌خط راست $۰ \le x \le ۱$ ، تناظر یک‌به‌یک برقرار کرد، گویند مجموعهٔ M دارای توان پیوستار است. به‌ویژه، بنابراین تعریف، خود مجموعهٔ نقطه‌های پاره‌خط راست $۰ \le x \le ۱$ ، توان پیوستار دارد.

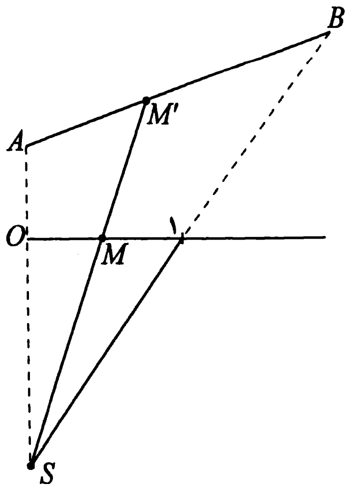
روی شکل ۱ دیده می‌شود، مجموعهٔ نقطه‌های هر پاره‌خط راست AB، دارای توان پیوستار است. در این‌جا تناظر یک‌به‌یک با روش هندسی و به یاری تصویر مرکزی برقرار شده است.

به‌سادگی می‌توان ثابت کرد، مجموعهٔ نقطه‌های هر فاصلهٔ باز $a < x < b$ و مجموعهٔ همهٔ نقطه‌های هر خط راست $-\infty < x < +\infty$ ، در تناظر یک‌به‌یک قرار دارند و بنابراین، مجموعهٔ نقطه‌های واقع بر هر خط راست، دارای توان پیوستار است.

این حکم بسیار جالب‌تر است: مجموعهٔ نقطه‌های واقع بر محیط و درون مربعِ

$$۰ \le x \le ۱ \quad , \quad ۰ \le y \le ۱$$

توان پیوستار دارد. به این ترتیب، اگر دقیق‌تر بگوییم، در مربع «همان‌قدر» نقطه وجود دارد که روی یک پاره‌خط راست.



شکل ۱

۳. عددهای حقیقی[۱]

تکامل مفهوم عدد را، به‌تفصیل در بخش اول (جلد اول) دیده‌ایم. در این‌جا تصور بیشتری از نظریهٔ عددهای حقیقی، که در سدهٔ نوزدهم و در بستگی با بنیان‌گذاری مفهوم‌های اصلی آنالیز پدید آمده است، به خواننده می‌دهیم.

عددهای گویا. فرض را بر این می‌گیریم که خواننده با ویژگی‌های اصلی عددهای گویا آشناست. بدون این که بحث را به‌درازا بکشانیم، این ویژگی‌ها را به‌یاد می‌آوریم. عددهای گویا، یعنی عددهای به‌صورت $\frac{m}{n}$ (m و n عددهایی درست‌اند و $n \neq 0$) به مجموعهٔ عددهایی گفته می‌شود که دربارهٔ آن، دو عمل (جمع و ضرب) تعریف شده باشد. این عمل‌ها از یک‌رشته قانون (اصل موضوع) پیروی می‌کنند. در این‌جا، a، b، c، ... را، عددهایی گویا گرفته‌ایم.

I. *اصل موضوع‌های جمع*

۱) $a + b = b + a$ (اصل جابه‌جایی)؛

---

۱. برای نوشتن این بند، از سفارش‌های آ. ن. کُولموگُورُوف سود برده‌ام.

۲) $a + (b + c) = (a + b) + c$ (اصل شرکت‌پذیری) ؛

۳) معادله

$$a + x = b$$

یک جواب منحصر دارد (وجود عمل عکس).

از این اصل موضوع‌ها، به‌طور مستقیم نتیجه می‌شود: عبارت $a + b + c$، مفهوم یگانه‌ای دارد. عدد گویای ۰ وجود دارد، به‌نحوی که، برای آن داریم: $a + 0 = a$؛ درضمن برای جمع، عمل عکس - تفریق - وجود دارد، یعنی می‌توان $b - a$ را محاسبه کرد.

به این ترتیب، از دیدگاه جبری می‌توان گفت: عددهای گویا، نسبت به عمل جمع، یک گروه جابه‌جایی تشکیل می‌دهند.

II. *اصل موضوع‌های ضرب*

۱) $ab = ba$ (اصل جابه‌جایی)؛

۲) $a(bc) = (ab)c$ (اصل شرکت‌پذیری)؛

۳) معادلهٔ

$$ay = b$$

به‌شرط $a \neq 0$، یک جواب منحصر دارد (یعنی عمل عکس وجود دارد).

از این اصل‌ها نتیجه می‌شود: عبارت abc معنا دارد؛ عدد ۱ وجود دارد که برای آن داریم: $a.1 = a$؛ درضمن، برای عددهای گویای مخالف صفر، عمل عکس - تقسیم - وجود دارد؛ مجموعهٔ همهٔ عددهای گویا - به‌جز صفر- نسبت به عمل ضرب، یک گروه جابه‌جایی تشکیل می‌دهند.

III. *اصل موضوع پخشی*

۱) $(a + b)c = ac + bc$

همهٔ اصل موضوع‌های I تا III نشان می‌دهند که: عددهای گویا، نسبت به عمل‌های جمع و ضرب، به‌اصطلاح تشکیل ***میدان*** (***یا هیأت***) ***جبری*** می‌دهند.

IV. *اصل موضوع‌های ترتیبی*

۱) برای هر دو عدد گویای a و b، یکی و تنها یکی از این سه رابطه برقرار است:

$$a < b \quad , \quad a > b \quad , \quad a = b$$

۲) اگر داشته باشیم: $a < b$ و $b < c$، آنگاه $a < c$،

۳) اگر داشته باشیم: $a < b$، آن‌گاه $a + c < b + c$ (ویژگی یکنوایی جمع)،

۴) اگر داشته باشیم: $a < b$ و $c > ۰$، آن‌گاه $ac < bc$ (یکنوایی ضرب برای $c > ۰$).

این اصل موضوع‌ها، روی‌هم، این امکان را پدید می‌آورد که، مجموعهٔ عددهای گویا را «*میدان مرتب*» بنامیم.

به‌جز عددهای گویا، دستگاه‌های دیگری هم وجود دارد که همین اصل‌ها سازگارند و بنابراین، می‌توان از *میدان‌های مرتب* صحبت کرد.

دو ویژگی دیگر عددهای گویا را یادآوری می‌کنیم.

تراکم. برای هر دو عدد گویای a و b به‌شرط $a < b$، می‌توان عدد گویای c را طوری پیدا کرد که داشته باشیم: $a < c < b$.

شمارا بودن. مجموعهٔ همهٔ عددهای گویا، مجموعه‌ای شمارا است (بند ۲ را در همین بخش ببینید).

اندازه‌گیری کمیت‌ها. ناکافی بودنِ عددهای گویا برای ریاضیات، در یک مسألهٔ بسیار مهم روشن می‌شود: مسألهٔ مربوط به اندازه‌گیری کمیت‌ها. این مسأله را روی ساده‌ترین نمونه، یعنی اندازه‌گیری طول پاره‌خط‌های راست، بررسی می‌کنیم.

خط راستی را در نظر می‌گیریم که روی آن، جهتی معین، نقطه‌ای برای مبداء (نقطهٔ ۰) و واحدی برای اندازه‌گیری، مشخص کرده باشیم. در این صورت، معنای پاره‌خط راست OA که انتهای آن در نقطهٔ $\frac{۱}{۲}$، $\frac{۱}{۳}$، $\frac{۲}{۳}$، $-\frac{۱}{۳}$ وغیره باشد، برای ما قابل درک است. به‌طور کلی، هر عدد گویای a را می‌توان، با نقطه‌ای مثل A روی خط راست، متناظر کرد، یعنی نقطه‌ای با مختص $x = a$. در چنین حالتی، عدد a، طول پاره‌خط راست جهت‌دار OA را معین می‌کند. ولی، با این طرح برای طول، هر پاره‌خط راستی را نمی‌توان با عددی (گویا) اندازه گرفت. برای نمونه، همان‌طور که برای یونانیان باستان هم روشن بود، طول قطرِ مربعِ به ضلع واحد، با هیچ عدد گویایی قابل اندازه‌گیری نیست. به زبان دیگر، تعداد نقطه‌های واقع بر خط راست، بیشتر از تعداد عددهای گویاست. طبیعی‌ترین راه خروج از این موقعیت، این است که بین عددها و طول‌ها، تناظر یک‌به‌یک برقرار کنیم، یعنی مفهوم عدد را گسترش دهیم.

عددهای حقیقی. دیدیم، عددهای گویا به‌تنهایی، برای اندازه‌گیری کمیت‌ها کافی نیستند؛ به این نتیجه رسیدیم که، باید مفهوم عدد را به‌نحوی گسترش دهیم که بین عددها و نقطه‌های

واقع بر خط راست، تناظر یک‌به‌یک برقرار شود. برای این منظور، تلاش می‌کنیم روشن کنیم: آیا نمی‌شود موقعیت هر نقطهٔ دل‌خواه بر خط راست را، به‌نحوی به‌یاری همان نقطه‌های گویا معین کنیم؟ چنین طرحی در حوزهٔ عددهای گویا، ما را به مفهوم عددهای حقیقی می‌رساند.

α را نقطهٔ دل‌خواهی از خط راست فرض می‌کنیم. همهٔ عددهای گویای a را می‌توان به دو بخش تقسیم کرد: در یک بخش، نقطه‌های a را قرار می‌دهیم که در سمت چپ α قرار گرفته باشند؛ و در بخش دیگر، آن نقطه‌های a که در سمت راست α واقع‌اند. آن‌چه به خودِ عدد α مربوط است، اگر به‌تصادف گویا باشد، می‌تواند متعلق به هرکدام از بخش‌ها باشد. چنین تقسیمی از عددهای گویا را **بُرش** گویند. دو برش را وقتی یکسان گوییم که دو مجموعهٔ همهٔ عددهای گویای واقع در بخش چپ این برش‌ها، و دو مجموعهٔ عددهای گویای واقع در بخش راست آن‌ها (با دقت تا یک نقطه)، بر هم منطبق باشند. اکنون به‌روشنی دیده می‌شود که، دو عدد مختلف α و β به‌یاری دو برش مختلف تعریف می‌شوند. درواقع، از آن‌جا که عددهای گویا، همه‌جا روی خط راست، به‌صورتی متراکم توزیع شده‌اند، می‌توان عددهای گویای r_1 و r_2 را طوری پیدا کرد که درست بین α و β قرار گرفته باشند. در این صورت، این دوعدد، برای یکی از برش‌ها، در بخش راست و برای برش دیگر، دربخش چپ واقع می‌شوند.

به این ترتیب، هر نقطهٔ واقع بر خط راست، برشی را در دامنهٔ نقطه‌های گویا تعریف می‌کند؛ در ضمن، نقطه‌های مختلف، متناظر با برش‌های مختلف‌اند. این مطلب مهم است که می‌توان برش را طور دیگری تعریف کرد، به‌نحوی که تا اندازه‌ای، با تعریف قبلی متفاوت باشد. در این‌جا، خودِ عدد α را وارد در تعریف نمی‌کنند: برش را در دامنهٔ عددهای گویا می‌نامیم، وقتی که همهٔ عددهای گویا را به دو بخش، یعنی به دو مجموعهٔ غیرتهی و بدون اشتراکِ A و B چنان تقسیم کرده باشد که برای هر $a \in A$ و $b \in B$ داشته باشیم $a < b$. به یاری چنین تعریفی از برش است که می‌توان، به‌صورت یک‌ارزشی، نقطه یا رخنه‌ای را که برش پدید آورده است، بازیافت. به زبان دیگر، به یاری برش در مجموعهٔ عددهای گویا، می‌توان هر نقطهٔ واقع بر خط راست را تعریف کرد. این طرح را ر. **دِدْکیند**، ریاضی‌دان آلمانی پیشنهاد کرد؛ به همین دلیل، آن را **برش دِدْکیند** می‌نامند.

برش، تنها روشِ ممکن، برای تعیین جای هر نقطه به‌وسیلهٔ عددهای گویا، نیست. روش **ژرژ کانتور** که در این‌جا می‌آوریم، به عمل عادیِ اندازه‌گیری نزدیک‌تر است. دوباره α را نقطهٔ دل‌خواهی از خط راست می‌گیریم. می‌توان دو نقطهٔ گویا و به‌دل‌خواه نزدیک به همِ a و

b را طوری پیدا کرد که نقطهٔ α، بین دو نقطهٔ a و b باشد. نقطه‌های a و b، به‌تقریب، جای نقطهٔ α را مشخص می‌کنند. این روند تعیین تقریبی نقطهٔ α را، به‌صورتی بی‌پایان و به‌ترتیبی ادامه می‌دهیم که، در هر گام، دقت آن نسبت به گام پیشین، بیشتر شود. در این صورت، دستگاهی از بازه‌های بسته $[a_n , b_n]$ که دو انتهای هریک از آن‌ها، در نقطه‌های گویا قرار دارد داریم و:

$$[a_{n+1} , b_{n+1}] \subseteq [a_n , b_n] \quad , \quad b_n - a_n \to 0 \quad (n \to \infty)$$

دستگاه بازه‌های بسته‌ای که با این شرط‌ها سازگار باشد، ***دستگاه تودرتوی بازه‌های بسته*** نامیده می‌شود. روشن است، چنین دستگاهی از بازه‌های بسته، به‌صورتی یک‌ارزشی، جای نقطهٔ α را معین می‌کند.

به‌یاری چنین طرح‌هایی در حوزهٔ عددهای گویا، می‌توان عددهای حقیقی را تعریف کرد؛ به‌دنبال این تعریف، تعریفِ عمل‌های مربوط به عددهای حقیقی می‌آید و ثابت می‌شود همان اصل موضوع‌های مربوط به عددهای گویا دربارهٔ عددهای حقیقی هم برقرارند. اکنون دیگر، هر نقطهٔ واقع بر خط راست، متناظر است با یک عدد حقیقی و برعکس، هر عدد حقیقی متناظر با نقطه‌ای منحصر از خط راست است. به همین دلیل است که اغلب، مجموعهٔ عددهای حقیقی را ***محور عددی*** (یا خط راست عددی) هم می‌نامند.

اصل‌های پیوستگی. بین مجموعهٔ عددهای گویا و مجموعهٔ عددهای حقیقی، تفاوتی اساسی وجود دارد. مجموعهٔ همهٔ عددهای حقیقی، ویژگی‌هایی دارد که مُعرف پیوستگیِ این مجموعه است، درحالی که مجموعهٔ همهٔ عددهای گویا، چنین ویژگی‌هایی ندارد. این ویژگی‌ها را، اغلب «اصل‌های پیوستگی[1]» می‌نامند که مهم‌ترینِ آن‌ها را می‌آوریم.

اصل دِدِکیند. اگر مجموعهٔ همهٔ عددهای حقیقی، به دو مجموعهٔ غیرتهی و بدون عضوهای مشترکِ X و Y، چنان تقسیم شود که، برای هر $x \in X$ و $y \in Y$، نابرابریِ $x < y$ برقرار باشد، آن‌وقت یک و تنها یک عدد ξ (رخنه) وجود دارد که برای آن، نابرابری‌های $x \leq \xi \leq y$، برای هر $x \in X$ و $y \in Y$ برقرار است. مجموعهٔ عددهای حقیقی x، که با شرط $a \leq x \leq b$ سازگار باشند، ***فاصله*** یا ***بازهٔ*** (بسته) محور عددی نامیده می‌شود که آن را با نماد

---

۱. در حال حاضر، مرسوم است که این ویژگی‌ها را «اصل‌های کامل‌بودن» بنامند (ویراستار).

$[a, b]$ نشان می‌دهند. دستگاه بازه‌های $[a_n, b_n]$ را **تودرتو** (یا منقبض‌شونده) گویند، به‌شرطی که داشته باشیم:

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \quad \text{و} \quad b_n - a_n \to 0 \quad (n \to \infty)$$

اصل کانتور. برای هر دستگاه تودرتوی بازه‌های $[a_n, b_n]$، یک و تنها یک عدد حقیقی ξ وجود دارد که متعلق به همهٔ این بازه‌هاست.

اصل ویرشتراس. هر دنبالهٔ غیرنزولیِ عددهای حقیقی، که از بالا کران‌دار باشد، هم‌گراست.

دنبالهٔ عددهای حقیقی $\{x_n\}$ را **بنیادی**[1] گویند، وقتی برای هر $\varepsilon > 0$، عدد طبیعی N وجود داشته باشد، به‌نحوی که برای هر $n > N$ و هر عدد طبیعی p، داشته باشیم:

$$\left| x_{n+p} - x_n \right| < \varepsilon$$

اصل کوشی. هر دنبالهٔ بنیادی از عددهای حقیقی هم‌گراست.

از آن‌جا که به ساختار دقیق و منظم عددهای حقیقی نپرداختیم، نمی‌توانیم این موضوع را بررسی کنیم که، این اصل‌ها، دربارهٔ مجموعهٔ عددهای حقیقی صادق‌اند. نزدیک‌ترین هدف ما در این‌جا، بررسی بستگی متقابلی است که این اصل‌ها، با یکدیگر دارند. فرض کنیم، یکی از اصل‌های پیوستگی برای عددهای حقیقی برقرار باشد، ببینیم در این صورت، کدام اصل دیگر پیوستگی را می‌توان از آن نتیجه گرفت.

نتیجهٔ کلی در این باره این است که، همهٔ اصل‌های پیوستگی هم‌ارزند.

عدد b را کوچکترین کران بالای مجموعهٔ E گوییم

$$b = \sup E$$

وقتی که: ۱) برای هر $x \in E$ داشته باشیم: $x \leq b$؛ ۲) عدد $b' < b$ وجود نداشته باشد که دارای همین ویژگی باشد.

درضمن، با توجه به اصل دِدکیند ثابت می‌کنیم: هر مجموعهٔ غیرتهی E از عددها، که کران بالا داشته باشد، دارای کوچکترین کران بالاست. همهٔ عددهای حقیقی را، با توجه به

---

۱. به این گونه دنباله‌ها، دنباله‌های کوشی هم می‌گویند (ویراستار).

معیار زیر به دو دستهٔ X و Y تقسیم می‌کنیم: x را عضو X می‌گیریم، اگر عدد $a \in E$ وجود داشته باشد به‌نحوی که داشته باشیم $a \geq x$؛ همچنین y را عضو Y می‌گیریم، اگر برای هر $a \in E$ داشته باشیم $a < y$. به‌سادگی می‌توان روشن کرد که این، یک برش را تعریف می‌کند. بنا بر اصل ددکیند، این برش، دارای رخنهٔ ξ است؛ و همین رخنه، کوچکترین کران بالای مجموعهٔ E است.

اکنون ثابت می‌کنیم، اصل ویرشتراس را می‌توان از اصل ددکیند، نتیجه گرفت. $\{x_n\}$ را، دنباله‌ای غیرنزولی و کران‌دار، با کران از بالا، از عددهای حقیقی فرض می‌کنیم. هم‌اکنون ثابت کردیم، چنین دنباله‌ای، دارای کوچکترین کران بالای ξ است. بنا به تعریف کوچکترین کران بالا $x_n \leq \xi$ $(n \in N)$؛ برای هر $\varepsilon > ۰$، شماره‌ای مثل $n_۰$ پیدا می‌شود، به‌نحوی که داشته باشیم $x_{n_۰} > \xi - \varepsilon$. با توجه به یکنوایی دنبالهٔ $\{x_n\}$، از این‌جا نتیجه می‌شود: $\xi - \varepsilon < x_n \leq \xi$ (برای هر $n > n_۰$)، یعنی دنبالهٔ $\{x_n\}$ همگرا و دارای حد ξ است.

برای اثبات بستگی عکس، بین اصل‌های ددکیند و ویرشتراس، یادآوری می‌کنیم که از اصل ویرشتراس، می‌توان اصل ارشمیدس را نتیجه گرفت.

اصل ارشمیدس. دو عدد حقیقی $a > ۰$ و b هرچه باشند، می‌توان عدد طبیعی n را پیدا کرد، به‌نحوی که داشته باشیم: $na > b$.

این اصل، به‌معنای آن است که، برای هر عدد حقیقی b، دنبالهٔ $\left\{ \frac{b}{n} \right\}$ $(n \in N)$، به‌سمت صفر همگراست.

فرض کنیم، اصل ویرشتراس برقرار باشد، ولی اصل ارشمیدس برقرار نباشد. این حکم به‌معنای آن است که عدد $a > ۰$ وجود دارد، به‌نحوی که دنبالهٔ $x_n = na$ کران‌دار باشد. به‌جز آن، این دنباله صعودی است و بنابراین، بنا بر اصل ویرشتراس، دارای حدی مثل ξ است. از این‌جا نتیجه می‌شود، در فاصلهٔ $\left] \xi - \frac{a}{۲} , \xi \right[$، نقطهٔ $x_n = na$ از دنبالهٔ ما وجود دارد. ولی در این صورت

$$x_{n+۱} = (n+۱)\, a > \xi$$

و این، متناقض با آن است که ξ را کوچکترین کران بالای $\{x_n\}$ گرفته‌ایم.

از اصل ویرشتراس، اصل ددکیند به‌دست می‌آید. فرض کنید مجموعهٔ همهٔ عددهای حقیقی، به مجموعِ دو مجموعهٔ جدا از هم X و Y طوری تقسیم شده باشد که برای هر $x \in X$ و $y \in Y$ داشته باشیم: $x < y$. ثابت می‌کنیم، این برش، دارای رخنهٔ منحصر ξ

است. m را عددی درست و n را عددی طبیعی می‌گیریم. فرض می‌کنیم، x_n بزرگترین عضو به‌صورت $\frac{m}{2^n} \in X$ باشد، به‌نحوی که داشته باشیم: $x_n + \frac{1}{2^n} \in Y$. چون مجموعهٔ عددهای به‌صورت $\frac{m}{2^n}$ زیرمجموعه‌ای از مجموعهٔ عضوهای به‌صورت $\frac{m}{2^{n+1}}$ است، پس $x_n \leq x_{n+1}$. به‌جز این، دنبالهٔ $\{x_n\}$ کران‌دار است (از جمله به‌وسیلهٔ عدد $(x_1 + \frac{1}{2})$). بنابراین، بنا بر اصل ویرشتراس، دارای حد ξ است. ثابت می‌کنیم، همین ξ، رخنهٔ برشِ ما است. درواقع، اگر $x < \xi$، آن‌وقت $x \in X$. اگر هم $y > \xi$، آن‌گاه $y \in Y$، زیرا بنا به اصل ارشمیدس، می‌توان شماره‌ای مثل n پیدا کرد به‌نحوی که داشته باشیم: $\frac{1}{2^n} < y - \xi = a$. ولی $x_n < \xi$ ، $x_n + \frac{1}{2^n} \in Y$، در این صورت $y = \xi + a > x_n + \frac{1}{2^n}$ و بنابراین $y \in Y$.

همچنین می‌توان ثابت کرد، اصل‌های کانتور و کوشی هم‌ارزند. با وجود این، از جمله از برقراری اصل کوشی، نمی‌توان اصل دِدکیند را نتیجه گرفت. این حکم را باید به این معنا فهمید: میدان مرتبی وجود دارد که در آن، اصل کوشی برقرار است، ولی اصل دِدکیند صدق نمی‌کند. ولی اگر اصل ارشمیدس را درست بدانیم، هر چهار اصل هم‌ارز می‌شوند.

ناشمارا بودن پیوستار. ثابت می‌کنیم، مجموعهٔ همهٔ نقطه‌های پاره‌خط راست $0 \leq x \leq 1$ ، مجموعه‌ای شمارا نیست. اثبات را با برهان خُلف می‌دهیم. فرض کنید، مجموعهٔ همهٔ نقطه‌های پاره‌خط راست $0 \leq x \leq 1$ ، مجموعه‌ای شمارا باشد. در چنین صورتی می‌توان همهٔ نقطه‌های x از این پاره‌خط راست را، به‌یاری عددهای طبیعی شماره‌گذاری کرد:

$$x_1 , x_2 , x_3 , \dots , x_n , \dots \qquad (1)$$

بازهٔ (بستهٔ) σ_1 را در بازهٔ $[0 , 1]$ طوری انتخاب می‌کنیم که طول آن از ۱ کوچکتر باشد؛ درضمن، شامل نقطهٔ x_1 نباشد. بی‌تردید، چنین بازه‌ای را می‌توان پیدا کرد. سپس بازهٔ σ_2 را در درون بازهٔ σ_1 طوری انتخاب می‌کنیم که طول σ_2 کمتر از $\frac{1}{2}$ باشد و درضمن، شامل نقطه‌های x_1 و x_2 نباشد. به‌طور کلی، با در دست داشتن بازهٔ σ_{n-1}، بازهٔ σ_n را در درون آن طوری انتخاب می‌کنیم که طولی کمتر از $\frac{1}{n}$ داشته باشد و درضمن، شامل نقطه‌های x_1، x_2، ... ، x_n نباشد. به این ترتیب، دنباله‌ای بی‌پایان از بازه‌های

$$\sigma_1 , \sigma_2 , \sigma_3 , \dots , \sigma_n , \dots$$

به‌دست می‌آید. در این دنباله، هر بازه جزئی از بازه پیش از خود است؛ درضمن طول σ_n ، با

بزرگ شدن n، به‌سمت صفر میل می‌کند. بنابراین، با توجه به اصل کانتور، نقطهٔ منحصر x از بازه [۰, ۱] وجود دارد که به همهٔ پاره‌خط‌های راست σ_n متعلق است. از آن‌جا که، بنا به فرض، همهٔ نقطه‌های بازه [۰, ۱]، در دنبالهٔ (۱) آمده است، بنابراین، نقطهٔ x هم (یعنی نقطهٔ مشترک همهٔ بازه‌های σ_n) بر نقطه‌ای مثل x_m از این دنباله، منطبق است. از طرف دیگر، بنابر ساختمان بازه‌ها، σ_m نباید شامل نقطهٔ x_m باشد، یعنی $x \neq x_m$. به این ترتیب، به‌تناقض رسیدیم. درنتیجه، فرض نخستین ما، دربارهٔ شمارا بودن مجموعهٔ همهٔ نقطه‌های بازهٔ [۰, ۱] درست نیست، یعنی مجموعهٔ همهٔ نقطه‌های بازهٔ [۰, ۱]، مجموعه‌ای ناشمارا است و این، همان چیزی است که می‌خواستیم ثابت کنیم.

این قضیه روشن می‌کند، مجموعه‌های بی‌پایان، توان‌های مختلفی دارند و، بنابراین، برای پرسشی که در بند ۱، زیر عنوان مجموعه‌های باپایان و مجموعه‌های بی‌پایان مطرح کردیم، به پاسخ رسیدیم: تعداد عضوهای مجموعه‌های بی‌پایان، همیشه یکی نیست.

۴. مجموعه‌های نقطه‌ای

در بند پیش به مجموعه‌هایی برخوردیم که عضوهای آن‌ها را ***نقطه‌ها*** تشکیل می‌دادند. به‌ویژه، مجموعهٔ همهٔ نقطه‌های واقع بر یک خط راست و مجموعهٔ همهٔ نقطه‌های (x,y) را در مربع $0 \leq x \leq 1$ و $0 \leq y \leq 1$ بررسی کردیم. اکنون اندکی بیشتر به بررسی این مجموعه‌ها می‌پردازیم.

مجموعه‌ای که هر عضو آن یک نقطه باشد، ***مجموعهٔ نقطه‌ای*** نامیده می‌شود. بنابراین، می‌توان از مجموعه‌های نقطه‌ای روی خط راست، روی صفحه و در هرگونه فضا صحبت کرد. برای ساده‌تر شدن مطلب، در این‌جا، تنها دربارهٔ مجموعه‌های نقطه‌ای ***روی خط راست*** بحث می‌کنیم.

بین عددهای حقیقی و نقطه‌های واقع بر یک خط راست، رابطه‌ای تنگاتنگ وجود دارد: هر عدد حقیقی را می‌توان متناظر با نقطه‌ای از خط راست و، برعکس، هر نقطهٔ واقع بر خط راست را می‌توان متناظر با یک عدد حقیقی قرار داد. بنابراین، وقتی از مجموعه‌های نقطه‌ای صحبت می‌کنیم، مجموعه‌هایی را هم‌که از عددهای‌حقیقی تشکیل شده‌اند ــ مجموعه‌های واقع بر محور عددی ــ جزو آن‌ها به‌حساب می‌آوریم. برعکس، برای این که یک مجموعهٔ

نقطه‌ای را مطرح کنیم، اغلب مختص همهٔ نقطه‌های مجموعهٔ خود را، در نظر می‌گیریم.

مجموعه‌های نقطه‌ای (و در حالت خاص، مجموعه‌های نقطه‌ای روی خط راست)، ویژگی‌هایی دارند که آن‌ها را از مجموعه‌های غیرمشخص جدا می‌کند و نظریهٔ مجموعه‌های نقطه‌ای را، به‌صورت نظام ریاضی مستقلی درمی‌آورد. پیش از همه، در این‌جا می‌توان دربارهٔ ***فاصلهٔ*** بین دو نقطه سخن گفت. سپس، بین نقطه‌های واقع بر خط راست، می‌توان نسبت به هم، ردیفی قایل شد (به‌یاری واژه‌هایی مثل «در سمت چپ»، «در سمت راست»)؛ بر همین اساس است که می‌گویند: مجموعهٔ نقطه‌ای روی خط راست، یک **مجموعهٔ مرتب** است. سرانجام، همان‌طور که پیش از این هم گفتیم، اصل کانتور دربارهٔ خط راست درست است؛ این ویژگی خط راست را، به‌عنوان ***کامل بودن*** خط راست می‌شناسند.

نمادهایی را که برای ساده‌ترین مجموعه‌های روی خط راست وجود دارد، می‌آوریم.

بازهٔ $[a,b]$، یعنی مجموعهٔ نقطه‌هایی که مختص آن‌ها در نابرابری‌های $a \le x \le b$ صدق کنند.

بازهٔ (a,b)، یعنی مجموعهٔ نقطه‌هایی که مختص آن‌ها در نابرابری‌های $a < x < b$ صدق می‌کنند.

نیم‌بازه‌های $[a,b)$ و $(a,b]$ که، به‌ترتیب، با شرط‌های $a \le x < b$ و $a < x \le b$ تعریف می‌شوند.

بازه‌ها و نیم‌بازه‌ها ممکن است ناویژه باشند. از جمله بازه $(-\infty, +\infty)$ به‌معنای تمامی خط راست و بازهٔ $(-\infty, b]$ به‌معنای مجموعهٔ همهٔ نقطه‌هایی است که، برای آن‌ها، داشته باشیم: $x \le b$.

از بررسی امکان‌های مختلف استقرارِ مجموعه بر خط راست، آغاز می‌کنیم.

مجموعه‌های کران‌دار و مجموعه‌های بی‌کران. مجموعهٔ E، نقطه‌های روی خط راست، یا از نقطه‌هایی تشکیل شده است که فاصلهٔ آن‌ها از مبداء مختصات، از عدد مثبت مشخصی تجاوز نمی‌کند ویا شامل نقطه‌هایی است که به‌دل‌خواه، دور از مبداء مختصات قرار دارند. مجموعهٔ E را در حالت اول ***کران‌دار*** و در حالت دوم ***بی‌کران*** گویند. مجموعهٔ همهٔ نقطه‌های بازهٔ $[0,1]$ نمونه‌ای از مجموعه‌های کران‌دار و مجموعهٔ همهٔ نقطه‌های با مختص درست، نمونه‌ای از مجموعه‌های بی‌کران است.

روشن است، اگر a نقطه‌ای ثابت بر خط راست باشد، مجموعهٔ E، تنها در حالتی کران‌دار

است که فاصلهٔ از نقطهٔ a تا هر نقطهٔ دل‌خواه $x\in E$، از عدد مثبت معینی تجاوز نکند.

مجموعه‌ای که از بالا یا از پایین کران‌دار است. E را مجموعه‌ای از نقطه‌های واقع بر خط راست می‌گیریم. اگر روی خط راست، نقطه‌ای مثل A وجود داشته باشد، به‌نحوی که هر نقطهٔ دل‌خواه $x\in E$، در سمت چپ نقطهٔ A باشد، آن‌وقت گویند: مجموعهٔ E ***از بالا کران‌دار*** است. به همین ترتیب، اگر روی خط راست، نقطه‌ای مثل A پیدا شود که هر نقطهٔ دل‌خواه $x\in E$ در سمت راست آن باشد، گویند مجموعهٔ E ***از پایین کران‌دار*** است. برای نمونه، مجموعهٔ همهٔ نقطه‌های با مختص مثبت از پایین، و مجموعهٔ همهٔ نقطه‌های خط راست با مختص منفی از بالا کران‌دار است.

تعریفی که برای مجموعهٔ کران‌دار آوردیم، با این تعریف هم‌ارز است: مجموعهٔ E از نقطه‌های واقع بر خط راست را کران‌دار گوییم، وقتی که از بالا و پایین محدود (یعنی دارای کران) باشد. با آن که دو تعریف شباهت زیادی با هم دارند، بین آن‌ها اختلافی اساسی دیده می‌شود: تعریف اول، بر فاصلهٔ بین نقطه‌های واقع بر خط راست تکیه می‌کند و تعریف دوم، بر مرتب بودن مجموعه.

همچنین می‌توان گفت: یک مجموعه وقتی کران‌دار است که، به‌طور کامل، در درون بازه‌ای مثل $[a ,b]$ واقع باشد.

کوچکترین کران بالا و بزرگترین کران پایین مجموعه. فرض کنید، مجموعهٔ E، کرانِ بالا داشته باشد. در این صورت، روی خط راست، نقطه‌های A وجود دارند که، در سمت راست آن‌ها، یک نقطه از مجموعهٔ E هم، پیدا نمی‌شود. با استفاده از اصل کانتور، می‌توان ثابت کرد، بین همهٔ نقطه‌های A (که دارای این ویژگی‌اند)، نقطه‌ای وجود دارد که در سمت چپ همهٔ دیگران قرار گرفته است. این نقطه را، ***کوچکترین کران بالای*** مجموعهٔ E گویند. به همین ترتیب، ***بزرگترین کران پایین*** یک مجموعهٔ نقطه‌ای تعریف می‌شود.

اگر در خودِ مجموعهٔ E، سمت‌راست‌ترین نقطه وجود داشته باشد، روشن است که همین نقطه، کوچکترین کران بالای مجموعهٔ E است. ولی ممکن است در مجموعهٔ E، نتوان نقطه‌ای پیدا کرد که سمت‌راست‌ترین نقطه باشد. از جمله، مجموعهٔ نقطه‌های

$$\frac{۰}{۱},\ \frac{۱}{۲},\ \frac{۲}{۳},\ \frac{۳}{۴},\ \frac{۴}{۵},\ \dots$$

کران بالا دارد، ولی در این مجموعه، نقطه‌ای پیدا نمی‌شود که در سمت راست همهٔ نقطه‌های دیگر باشد. در چنین حالتی، کوچکترین کران بالای a، متعلق به مجموعهٔ E نیست، ولی می‌توان نقطه‌هایی از مجموعهٔ E را پیدا کرد که به‌دل‌خواه، نزدیک به a باشند. در مثالی که آوردیم، داریم $a = ۱$.

جاگرفتن مجموعهٔ نقطه‌ای، در نزدیکی نقطه‌ای از خط راست. E را مجموعه‌ای نقطه‌ای و x را، نقطه‌ای از خط راست می‌گیریم. امکان‌های گوناگون استقرار مجموعهٔ E در نزدیکی x را بررسی می‌کنیم. این حالت‌ها ممکن است پدید آید:

۱. نه خودِ نقطهٔ x و نه نقطه‌های به‌اندازهٔ کافی نزدیک به آن، عضو مجموعهٔ E نیستند.

۲. نقطهٔ x، متعلق به مجموعهٔ E نیست، ولی نقطه‌های به‌قدر کافی نزدیک به آن، در مجموعهٔ E وجود دارد.

۳. نقطهٔ x به مجموعهٔ E تعلق دارد، ولی همهٔ نقطه‌های به‌اندازهٔ کافی نزدیک به آن، متعلق به مجموعهٔ E نیستند.

۴. نقطهٔ x متعلق به مجموعهٔ E است؛ درضمن، نقطه‌های دیگری هم از مجموعهٔ E وجود دارند که به‌اندازهٔ کافی به x نزدیک‌اند.

نقطهٔ x را در حالت اول ***نقطهٔ بیرونیِ مجموعهٔ*** E؛ در حالت سوم ***نقطهٔ تنهایی مجموعهٔ*** E، و در حالت‌های دوم و چهارم ***نقطهٔ حدی مجموعهٔ*** E گویند.

به این ترتیب، اگر x عضوی از مجموعهٔ E نباشد، یا نقطهٔ بیرونی و یا نقطهٔ حدی برای مجموعهٔ E است؛ ولی اگر x عضوی از مجموعهٔ E باشد، آن‌وقت یا نقطهٔ تنهایی از مجموعهٔ E و یا نقطهٔ حدی آن است.

نقطهٔ حدی می‌تواند عضو مجموعهٔ E باشد یا نباشد و با این شرط مشخص می‌شود: نقطه‌هایی از مجموعهٔ E، که به‌اندازهٔ کافی به آن نزدیک باشند، وجود داشته باشد. به‌زبان دیگر، وقتی x را، نقطهٔ حدیِ مجموعهٔ E گویند که هر فاصلهٔ δ، که نقطهٔ x را در بر می‌گیرد، بی‌نهایت نقطه از مجموعهٔ E را در بر بگیرد. مفهوم نقطهٔ حدی، یکی از مهم‌ترین مفهوم‌های نظریهٔ مجموعه‌های نقطه‌ای است.

اگر نقطهٔ x و همهٔ نقطه‌های به‌اندازهٔ کافی نزدیک به آن، متعلق به مجموعهٔ E باشند، آن‌وقت x را، ***نقطهٔ درونی*** E گویند. هر نقطهٔ x که، برای مجموعهٔ E، نه بیرونی باشد و نه درونی، ***نقطهٔ مرزی مجموعهٔ*** E نامیده می‌شود.

مثال‌هایی برای روشن‌تر شدن این مفهوم‌ها، می‌آوریم.

مثال ۱. فرض کنید، مجموعهٔ $E_۱$، شامل نقطه‌هایی با این مختص‌ها باشد:

$$۱ , \frac{۱}{۲} , \frac{۱}{۳} , \dots , \frac{۱}{n} , \dots$$

در این مجموعه، هر نقطه، نقطه‌ای تنها است؛ نقطهٔ ۰ نقطهٔ حدی آن است (که عضوی از مجموعهٔ $E_۱$ نیست)؛ همهٔ نقطه‌های دیگر خط راست، نسبت به مجموعهٔ $E_۱$، نقطه‌های بیرونی‌اند.

مثال ۲. فرض کنید، مجموعهٔ $E_۲$، شامل همهٔ نقطه‌های گویای بازهٔ $[۰ , ۱]$ باشد. این مجموعه، نقطه‌های تنها ندارد؛ هر نقطه از بازهٔ $[۰ , ۱]$ نقطهٔ حدی برای $E_۲$ است؛ همهٔ نقطه‌های دیگر خط راست، نسبت به مجموعهٔ $E_۲$، نقطه‌های بیرونی‌اند. روشن است، در میان نقطه‌های حدی مجموعهٔ $E_۲$، هم نقطه‌های متعلق به آن و هم نقطه‌هایی که به آن تعلق ندارند، پیدا می‌شود.

مثال ۳. فرض کنید، مجموعهٔ $E_۳$، شامل همهٔ نقطه‌های پاره‌خط راست $[۰ , ۱]$ باشد. مثل مثال قبل، مجموعهٔ $E_۳$ هم، نقطهٔ تنها ندارد و هر نقطهٔ پاره‌خط راست $[۰ , ۱]$ نقطهٔ حدی آن است. با وجود این، برخلاف مثال قبل، همهٔ نقطه‌های مجموعهٔ $E_۳$، متعلق به این مجموعه‌اند.

مثال ۴. فرض کنید، مجموعهٔ $E_۴$، شامل همهٔ نقطه‌های با مختص درست روی خط راست باشد. هر نقطه از مجموعهٔ $E_۴$، نقطهٔ تنهای آن است؛ مجموعهٔ $E_۴$، نقطهٔ حدی ندارد.

درضمن یادآوری می‌کنیم، در مثال ۳، هر نقطهٔ بازهٔ $(۰ , ۱)$، نقطه‌ای درونی از $E_۳$ و در مثال ۲، هر نقطه از بازهٔ $[۰ , ۱]$، نقطهٔ مرزی $E_۲$ است.

از این مثال‌ها دیده می‌شود، یک مجموعهٔ بی‌پایان از نقطه‌های واقع بر خط راست، ممکن است نقطه‌های تنها داشته باشد ($E_۱ , E_۴$) و ممکن است نقطهٔ تنهایی نداشته باشد ($E_۲ , E_۳$)؛ به همین ترتیب، چنین مجموعه‌ای، ممکن است نقطه‌های درونی داشته باشد ($E_۳$) و ممکن است شامل چنین نقطه‌هایی نباشد ($E_۱ , E_۲ , E_۴$). دربارهٔ نقطه‌های حدی،

تنها مجموعهٔ $E_۴$ بود که نقطهٔ حدی نداشت. همان‌طور که قضیهٔ مهم زیر هم ثابت می‌کند، این مطلب، به آن‌جا مربوط می‌شود که مجموعهٔ $E_۴$ بی‌کران است.

قضیهٔ بولتسانو ـ ویرشتراس. ***هر مجموعهٔ کران‌دار واقع بر خط راست، دست‌کم یک نقطهٔ حدی دارد.***

این قضیه را ثابت می‌کنیم. E را مجموعه‌ای بی‌پایان و کران‌دار از نقطه‌هایی واقع بر خط راست، فرض می‌کنیم. چون مجموعهٔ E کران‌دار است، بنابراین به‌طور کامل، در بازه‌ای مثل $[a, b]$ قرار دارد. این بازه را نصف می‌کنیم. از آن‌جا که مجموعهٔ E بی‌پایان است، دست‌کم در یکی از بازه‌های حاصل، بی‌نهایت نقطه از مجموعهٔ E واقع است. این بازه را $\sigma_۱$ می‌نامیم (در حالتی که هر دو نیمهٔ بازهٔ $[a, b]$، شامل بی‌نهایت نقطه از مجموعهٔ E باشد، یکی از دو نیمه و در مثل نیمهٔ سمت چپ را به‌عنوان $\sigma_۱$ انتخاب می‌کنیم). اکنون، بازهٔ $\sigma_۱$ را به دو نیمهٔ برابر بخش می‌کنیم. چون بخشی از مجموعهٔ E که در بازهٔ $\sigma_۱$ واقع است، مجموعه‌ای بی‌پایان است، بنابراین دست‌کم در یکی از دو نیمهٔ آن، بی‌نهایت نقطه از مجموعهٔ E وجود دارد. این بازه را $\sigma_۲$ می‌نامیم. روند تقسیم بازه‌ها به دو نیمهٔ برابر را، به‌طور نامحدود ادامه می‌دهیم و هر بار، نیمه‌ای را انتخاب می‌کنیم که شامل بی‌نهایت نقطه از مجموعهٔ E باشد. دنبالهٔ بازه‌های

$$\sigma_۱ , \sigma_۲ , \dots , \sigma_n , \dots$$

به‌دست می‌آید. این دنبالهٔ بازه‌ها (یعنی پاره‌خط‌های راست) ویژگی‌هایی دارد: هر بازهٔ $\sigma_{n+۱}$، بخشی از بازهٔ قبلیِ σ_n است؛ هر بازهٔ σ_n شامل بی‌نهایت نقطه از مجموعهٔ E است؛ طول بازه‌های σ_n به‌سمت صفر میل می‌کنند. دو ویژگی نخست، بلافاصله و از شیوهٔ ساختمان این دنباله نتیجه می‌شود؛ برای اثبات ویژگی سوم، کافی است یادآوری کنیم، اگر طول پاره‌خط راست $[a, b]$ برابر l باشد، طول پاره‌خط راست σ_n برابر $\frac{l}{۲^n}$ می‌شود. بنا بر اصل کانتور، نقطهٔ منحصر x وجود دارد که متعلق به همهٔ بازه‌های σ_n است. ثابت می‌کنیم، این نقطهٔ x، یک نقطهٔ حدی از مجموعهٔ E است. برای این منظور، کافی است ثابت کنیم، اگر δ بازه‌ای شامل نقطهٔ x باشد، شامل بی‌نهایت نقطه از مجموعهٔ E است. چون هر بازهٔ σ_n شامل x است، درضمن، طول بازه‌های σ_n به سمت صفر میل می‌کند، بنابراین برای مقدارهای به‌قدر کافی بزرگ n، بازهٔ σ_n به‌طور کامل، در بازهٔ δ قرار می‌گیرد. ولی بنا بر شرط، σ_n شامل بی‌نهایت نقطه از مجموعهٔ E است. بنابراین در δ هم، بی‌نهایت نقطه از

مجموعهٔ E وجود دارد. به این ترتیب، نقطهٔ x، درواقع یک نقطهٔ حدی از مجموعه است و قضیه ثابت می‌شود.

تمرین. ثابت کنید، اگر مجموعهٔ E از بالا کران‌دار باشد و نقطهٔ سمت راست نداشته باشد، کوچکترین کران بالای آن، یک نقطهٔ حدی از مجموعهٔ E است (که البته، به خودِ E تعلق ندارد).

مجموعه‌های بسته و مجموعه‌های باز. یکی از مسأله‌های اساسیِ نظریهٔ مجموعه‌های نقطه‌ای، بررسی ویژگی‌های گونه‌های مختلف این مجموعه‌هاست. خواننده را روی دو نمونه، با این نظریه آشنا می‌کنیم. درواقع در این‌جا، به ویژگی به‌اصطلاح مجموعه‌های بسته و مجموعه‌های باز می‌پردازیم.

مجموعه را بسته گویند، وقتی شامل همهٔ نقطه‌های حدی خود باشد. در حالتی هم که مجموعه دارای نقطهٔ حدی نیست، آن را جزو مجموعه‌های بسته به‌حساب می‌آورند. مجموعهٔ بسته، به‌جز نقطه‌های حدی، می‌تواند نقطه‌های تنها هم داشته باشد. مجموعه را *باز* گویند، وقتی که هر نقطهٔ آن، یک نقطهٔ درونی برای مجموعه باشد.

نمونه‌هایی از مجموعه‌های بسته و باز می‌آوریم. هر بازهٔ $[a, b]$، مجموعه‌ای بسته و هر بازهٔ (a, b)، مجموعه‌ای باز است. نیم‌بازه‌های ناویژهٔ $(-\infty, b]$ و $[a, +\infty)$، مجموعه‌های بسته و بازه‌های ناویژهٔ $(-\infty, b)$ و $(a, +\infty)$، مجموعه‌هایی باز هستند. هر خط راست، درعین‌حال، هم مجموعه‌ای بسته و هم مجموعه‌ای باز است. بهتر است، مجموعهٔ تهی را هم، درعین‌حال، مجموعه‌ای بسته و باز در نظر بگیریم. هر مجموعهٔ با پایان از نقطه‌های واقع بر خط راست، بسته است، زیرا چنین مجموعه‌ای نقطهٔ حدی ندارد. مجموعه‌ای که از نقطه‌های

$$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots$$

تشکیل شده باشد، مجموعه‌ای بسته است؛ این مجموعه دارای یک نقطهٔ حدی $x = 0$ است که به مجموعه تعلق دارد.

مسأله این است: می‌خواهیم روشن کنیم، چگونه می‌توان یک مجموعهٔ دل‌خواه بسته یا باز را تشکیل داد؟ برای این منظور، به چند گزارهٔ کمکی نیاز داریم، که آن‌ها را بدون اثبات می‌آوریم:

۱. اشتراک چند مجموعهٔ بسته، خود یک مجموعهٔ بسته است.

۲. اجتماع چند مجموعهٔ باز، خود یک مجموعهٔ باز است.

۳. اگر یک مجموعهٔ بسته کران بالا داشته باشد، در آن صورت کوچکترین کران بالا دارد. به همین ترتیب، اگر مجموعهٔ بسته از پایین کران‌دار باشد، آنگاه دارای بزرگترین کران پایین است.

اگر E مجموعهٔ دل‌خواهی از نقطه‌های واقع بر خط راست باشد، آن‌وقت، مجموعهٔ همهٔ نقطه‌هایی از خط راست را که به مجموعهٔ E تعلق ندارند، **مجموعهٔ متمم** E می‌نامند و با CE (یا EC) نشان می‌دهند. روشن است، اگر x برای E، نقطهٔ بیرونی باشد، برای مجموعهٔ CE نقطهٔ درونی است و برعکس.

۴. اگر F مجموعه‌ای بسته باشد، متمم آن CF مجموعه‌ای باز است و برعکس.

گزارهٔ ۴ نشان می‌دهد، بین مجموعه‌های بسته و باز، رابطه‌ای جدی وجود دارد. یکی از آن‌ها متمم دیگری است. به‌اعتبار این گزاره، کافی است، تنها به بررسی مجموعه‌های بسته ویا تنها مجموعه‌های باز بپردازیم. آگاهی از ویژگی‌های مجموعه‌های یک گونه، به‌طور مستقیم ویژگی‌های مجموعه‌های گونه دیگر را، روشن می‌کند. از جمله، هر مجموعهٔ باز را می‌توان با جدا کردن یک مجموعهٔ بسته از خط راست، به‌دست آورد.

به بررسیِ ویژگی‌های مجموعه‌های بسته می‌پردازیم. در آغاز تعریفی می‌آوریم. F را مجموعه‌ای بسته فرض کنید. اگر بازهٔ (a, b) دارای این ویژگی باشد که، هیچ کدام از نقطه‌های آن، متعلق به مجموعهٔ F نباشد، ولی خود نقطه‌های a و b عضو مجموعهٔ F باشند، در این صورت (a, b) را ***بازهٔ مجاور*** **مجموعهٔ** F می‌نامند. بازه‌های ناویژهٔ (a, ∞) یا $(-\infty, b)$ را هم از جملهٔ بازه‌های مجاور به‌حساب می‌آوریم، به‌شرطی که نقطهٔ a یا نقطهٔ b عضوی از مجموعهٔ F باشد، ولی خود بازه، اشتراکی با F نداشته باشد. ثابت می‌کنیم، اگر نقطهٔ x عضو مجموعهٔ بستهٔ F نباشد، عضو یکی از بازه‌های مجاور آن خواهد بود.

بخشی از مجموعهٔ F را که در سمت راست x قرار دارد، با نماد F_x نشان می‌دهیم. چون خود نقطهٔ x متعلق به مجموعهٔ F نیست، F_x را می‌توان به‌صورت این اشتراک نشان داد:

$$F_x = F \,.\, [x, +\infty)$$

هریک از دو مجموعهٔ F و $[x, +\infty)$، بسته است. بنابراین، با توجه به گزارهٔ ۱، F_x هم مجموعه‌ای بسته است. اگر مجموعهٔ F_x تهی باشد، به‌معنای آن است که هیچ بخشی از

نیم‌بازه $[x, \infty)$، متعلق به F نیست. فرض می‌کنیم، F_x مجموعه‌ای تهی نباشد. چون این مجموعه به‌طور کامل در نیم‌بازهٔ $[x, \infty)$ قرار دارد، بنابراین دارای کرانِ پایین است. بزرگترین کران پایین آن را b می‌نامیم. بنا بر گزارهٔ ۳ داریم: $b \in F_x$، و این به‌معنای آن است که $b \in F$. چون b بزرگترین کران پایین مجموعهٔ F_x است، بنابراین نیم‌بازهٔ $[x, b)$، که در سمت چپ نقطهٔ b واقع است، شامل نقطه‌ای از مجموعهٔ F_x نیست و درنتیجه، شامل نقطه‌ای از مجموعهٔ F هم نیست. به این ترتیب، نیم‌بازهٔ $[x, b)$ را ساختیم که شامل نقطه‌ای از مجموعهٔ F نیست و درضمن، یا $b = \infty$ ویا b متعلق به مجموعهٔ F است. به همین ترتیب، نیم‌بازهٔ $(a, x]$ ساخته می‌شود که شامل نقطه‌ای از مجموعهٔ F نیست و درضمن، یا $a = -\infty$ ویا $a \in F$. اکنون دیگر روشن است، بازهٔ (a, b)، که شامل نقطهٔ x است، یک بازهٔ مجاور برای مجموعهٔ F است. به‌سادگی دیده می‌شود، اگر (a_1, b_1) و (a_2, b_2)، دو بازهٔ مجاور مجموعهٔ F باشند، آن‌وقت این دو بازه یا بر هم منطبق‌اند و یا اشتراکی ندارند.

از آن‌چه گفتیم نتیجه می‌شود: هر مجموعهٔ بستهٔ واقع بر خط راست را، می‌توان از راه جدا کردن تعدادی بازه از خط راست، یعنی بازه‌های مجاور مجموعهٔ F، به‌دست آورد. از آن‌جا که در هر بازه، دست‌کم یک نقطهٔ گویا وجود دارد و درضمن، همهٔ نقطه‌های گویای واقع بر خط راست، مجموعه‌ای شمارا تشکیل می‌دهند، بنابراین به‌سادگی روشن می‌شود، تعداد همهٔ بازه‌های مجاور، از شمارا بودن تجاوز نمی‌کند. از این‌جا، نتیجهٔ نهایی به‌دست می‌آید: هر مجموعهٔ بستهٔ واقع بر خط راست را، می‌توان از راه جدا کردن یک مجموعهٔ شمارا، از بازه‌های مجزا، به‌دست آورد.

از این‌جا، و با توجه به گزارهٔ ۴، بلافاصله نتیجه می‌شود: هر مجموعهٔ باز واقع بر خط راست، عبارت است از اجتماع تعداد شمارایی از بازه‌های مجزا. همچنین با توجه به گزاره‌های ۱ و ۲، روشن است: هر مجموعه‌ای که با این روش ساخته می‌شود، درواقع یک مجموعهٔ بسته (باز) است.

همان‌طور که از نمونهٔ زیر دیده می‌شود، مجموعه‌های بسته می‌توانند ساختمانی بغرنج داشته باشند.

مجموعهٔ تام کانتوری. مجموعهٔ بستهٔ خاصی می‌سازیم که ویژگی‌های جالبی دارد. پیش از همه، بازه‌های ناویژهٔ $(-\infty, ۰)$ و $(۱, \infty)$ را از خط راست جدا می‌کنیم. بازهٔ $[۰, ۱]$ باقی می‌ماند. سپس، از این بازه، بازهٔ $\left(\frac{۱}{۳}, \frac{۲}{۳}\right)$ راکه شامل یک‌سوم میانیِ آن است، جدا می‌کنیم. از دو بازهٔ

باقی‌مانده، یعنی $\left[\circ , \frac{۱}{۳}\right]$ و $\left[\frac{۲}{۳} , ۱\right]$ ، یک‌سوم میانیِ هرکدام را جدا می‌کنیم. این روند جدا کردنِ یک‌سومِ میانیِ بازه‌های باقی‌مانده را، به‌طور نامحدود، ادامه می‌دهیم. مجموعهٔ نقطه‌هایی از خط راست را، که بعد از جدا کردنِ همهٔ این بازه‌ها به‌دست می‌آید، مجموعهٔ تام کانتوری گویند. مجموعهٔ تام کانتوری را با حرف P نشان می‌دهیم.

برخی از ویژگی‌های این مجموعه را بررسی می‌کنیم. مجموعهٔ P بسته است، زیرا از راه جدا کردن مجموعه‌ای از بازه‌های باز مجزا، از خط راست به‌دست آمده است. مجموعهٔ P هم تهی نیست، زیرا درهرحال، شامل نقطه‌های انتهایی همهٔ بازه‌های انتخابی است.

مجموعهٔ بستهٔ F را *تام* گویند، وقتی که شامل نقطه‌های تنها نباشد، یعنی وقتی که هریک از نقطه‌های آن، نقطه‌ای حدی باشد. ثابت می‌کنیم، مجموعهٔ P تام است. درواقع، اگر x، نقطهٔ تنهایی از مجموعهٔ P باشد، آن‌وقت باید انتهای مشترک دو بازهٔ مجاور این مجموعه باشد. ولی بنا بر ساختمان این مجموعه، بازه‌های مجاور در مجموعهٔ P، دارای انتهای مشترک نیستند.

مجموعهٔ P، شامل ولو یک بازه باز نیست. فرض کنیم بازه بازی همچون δ، به‌طور کامل، متعلق به مجموعهٔ P باشد. در آن صورت، این بازه، به‌طور کامل متعلق به یکی از بازه‌هایی خواهد بود که، در nامین گام ساختن مجموعهٔ P، به‌دست می‌آید. ولی این، ممکن نیست، زیرا وقتی $n \to \infty$، طول هریک از این بازه‌ها به‌سمت صفر میل می‌کند.

می‌توان ثابت کرد، مجموعهٔ P دارای توان پیوستار است. از این‌جا، به‌ویژه نتیجه می‌شود: مجموعهٔ تام کانتوری، به‌جز نقطه‌های انتهایی بازه‌های مجاور، نقطه‌های دیگری را هم در بر می‌گیرد. درواقع، اگر تنها دو انتهای بازه‌های مجاور را در نظر بگیریم، با یک مجموعهٔ شمارا سر و کار پیدا می‌کنیم.

در بسیاری شاخه‌های ریاضیات، مرتب با گونه‌های مختلف مجموعه‌های نقطه‌ای برمی‌خوریم و آگاهیِ بر ویژگی‌های این مجموعه‌ها، برای بررسی اغلب این مسأله‌های ریاضی، ضرورت دارد. به‌ویژه، نظریهٔ مجموعه‌های نقطه‌ای، برای آنالیز ریاضی و مکان‌شناسی (توپولوژی) اهمیت زیادی دارد.

نمونه‌هایی از دخالت مجموعه‌های نقطه‌ای را، در شاخه‌های سنتی آنالیز ریاضی، می‌آوریم. $f(x)$ را تابع پیوسته‌ای فرض کنید که روی پاره‌خط راست $[a ,b]$ داده شده باشد. عدد α را ثابت می‌گیریم و مجموعهٔ همهٔ نقطه‌های x را که برای آن‌ها داشته باشیم: $f(x) \geq \alpha$، در نظر می‌گیریم. به‌سادگی ثابت می‌شود، این مجموعه، می‌تواند مجموعهٔ بستهٔ

دل‌خواهی واقع بر بازهٔ $[a, b]$ باشد. به همین ترتیب، مجموعه‌ای از نقطه‌های x که برای آن‌ها داشته باشیم: $f(x) > \alpha$، می‌تواند مجموعهٔ باز و دل‌خواه $G \subset [a, b]$ باشد. اگر

$$f_۱(x) \ , f_۲(x) \ , \ \dots \ , f_n(x) \ , \ \dots$$

دنباله‌ای از تابع‌های پیوسته و مفروض، روی پاره‌خط راست $[a, b]$ باشد، مجموعهٔ نقطه‌های x، که در آن‌جا دنبالهٔ مفروض هم‌گرا (متقارب) است، نمی‌تواند دل‌خواه باشد و به گونهٔ معینی مربوط است.

نظامی از ریاضیات، که به بررسی ساختمان مجموعه‌های نقطه‌ای می‌پردازد، ***نظریهٔ توصیفی مجموعه‌ها*** نام دارد. در زمینه پیشبرد نظریهٔ توصیفی مجموعه‌ها و تکامل آن، پیش از هرکس، باید از ***ن. ن. لوزین*** ریاضی‌دان شوروی و شاگردان او ***پ. س. آلکساندروف***، ***م. یا. سوسلین***، ***آ. ن. گولموگُوروف***، ***م. آ. لاورنتیف***، ***پ. س. نُوویکوف***، ***ل. وِگِلدیش***، ***آ. آ. لیاپونوف*** و دیگران نام برد.

بررسی‌های لوزین و شاگردان او نشان داد، بین نظریهٔ توصیفی مجموعه‌ها و منطق ریاضی، رابطه‌ای تنگاتنگ وجود دارد. دشواری‌هایی که ضمن بررسی مسأله‌های نظریهٔ توصیفی مجموعه‌ها پدید می‌آید (از جمله، مسألهٔ مربوط به توان یک مجموعه)، درواقع، چنان دشواری‌هایی‌اند که سرشت منطقی دارند. برعکس، روش‌های منطق ریاضی، امکان نفوذ بیشتری را در برخی از مسأله‌های مربوط به نظریهٔ توصیفی مجموعه‌ها، به‌وجود آورده است.

۵. اندازهٔ مجموعه‌ها

مفهوم اندازهٔ مجموعه‌ها را، باید تعمیم مفهوم طول پاره‌خط راست (و البته، تعمیمی از مرحلهٔ بالای انتزاع) دانست. در ساده‌ترین حالت (که ما هم، تنها به آن می‌پردازیم)، مسأله عبارت از این است که، نه‌تنها برای طول پاره‌خط‌های راست (بازه‌های بسته)، بلکه برای هر مجموعهٔ پیچیدهٔ نقطه‌ای واقع بر خط راست، تعریفی پیدا کنیم.

بازهٔ $[۰, ۱]$ را، به‌عنوان واحد اندازه‌گیری انتخاب می‌کنیم. در این صورت روشن است، طول بازهٔ $[a, b]$ برابر $b - a$ می‌شود. به همین ترتیب، طبیعی است، اگر دو بازهٔ بدون

اشتراکِ $[a_1, b_1]$ و $[a_2, b_2]$ را در نظر بگیریم، برای طول مجموعهٔ E، که از این دو بازه تشکیل شده است، عدد $(b_1 - a_1) + (b_2 - a_2)$ را بپذیریم. ولی گام‌های بعدی را نمی‌توان به همین سادگی برداشت: طول مجموعه‌ای را که روی خط راست است و سرشت بغرنج‌تری دارد، به چه معنا باید گرفت؟ از جمله، طول مجموعهٔ کانتوری P، که در بند ۴ این بخش بررسی کردیم، برابر با چیست و چگونه باید محاسبه کرد؟ از این‌جا روشن می‌شود، مفهوم طول در مجموعه‌های واقع بر خط راست، نیاز به تعریف دقیق ریاضی دارد.

مسألهٔ تعیین طول مجموعه‌ها، یا آن‌طور که اغلب می‌گویند، مسألهٔ اندازه‌گیری مجموعه‌ها، اهمیت زیادی دارد، زیرا برای تعمیم مفهوم انتگرال، لازم است. مفهوم اندازهٔ مجموعه‌ها در سایر مسأله‌های مربوط به نظریهٔ تابع و همچنین در نظریهٔ احتمال، در مکان‌شناسی (توپولوژی)، در آنالیز تابعی (فوکسیونِل) وغیره هم، کاربرد دارد.

در این‌جا، تعریف *آ. لِه‌بِگ* ریاضی‌دان فرانسوی را، دربارهٔ اندازهٔ مجموعه‌ها می‌آوریم. همین تعریف، پایهٔ تعریف انتگرالی را تشکیل می‌دهد که به‌وسیلهٔ خود له‌بِگ طرح‌شده است.

اندازهٔ مجموعه‌های باز و بسته. از تعریف اندازهٔ یک مجموعهٔ دل‌خواه باز یا بسته آغاز می‌کنیم. همان‌طور که در بند ۴ یادآوری کردیم، هر مجموعهٔ باز واقع بر خط راست، عبارت است از اجتماع تعدادی محدود یا شمارا، از بازه‌هایی که دوبه‌دو بدون اشتراک‌اند.

اندازهٔ یک مجموعهٔ باز، به مجموع طول‌های بازه‌های تشکیل‌دهندهٔ آن گفته می‌شود. بنابراین اگر

$$G = \sum (a_i, b_i)$$

به این شرط که بازه‌های (a_i, b_i) دوبه‌دو بدون اشتراک باشند، آن‌وقت اندازهٔ G برابر است با $\sum (b_i - a_i)$. اگر در حالت کلی، اندازهٔ مجموعهٔ E را با نماد μE نشان دهیم، آن‌وقت

$$\mu G = \sum (b_i - a_i)$$

در حالت خاص، اندازهٔ یک بازه برابر با طول آن می‌شود:

$$\mu (a, b) = b - a$$

هر مجموعهٔ بسته F روی بازهٔ بستهٔ $[a, b]$، به‌نحوی که دو انتهای بازهٔ $[a, b]$ متعلق به F

باشد، از بازهٔ $[a ,b]$ و با کنار گذاشتن مجموعهٔ بازی مثل G از آن، به‌دست می‌آید. با توجه به این مطلب، اندازهٔ مجموعهٔ بستهٔ $F \subseteq [a ,b]$، که در آن a و b عضوهایی از F هستند، برابر است با تفاضل طول بازهٔ $[a ,b]$ و اندازهٔ مجموعهٔ باز G که متمم مجموعهٔ F نسبت به $[a ,b]$ است.

$$\mu F = (b - a) - \mu G \tag{۲}$$

با این تعریف، اندازهٔ هر بازهٔ بستهٔ $[a ,b]$، برابر است با طول آن:

$$\mu [a ,b] = b - a$$

و اندازهٔ مجموعه‌ای که شامل تعداد محدودی نقطه باشد، برابر است با صفر.

تعریف کلی اندازه. برای این که بتوانیم اندازهٔ مجموعه‌ها را، در حالتی که سرشتی کلی‌تر از مجموعه‌های باز و بسته دارند، تعریف کنیم، به یک مفهوم کمکی نیاز داریم. E را مجموعه‌ای واقع بر بازهٔ بسته $[a ,b]$ فرض می‌کنیم. همهٔ گونه‌های ممکنِ پوشش‌های مجموعهٔ E را در نظر می‌گیریم، یعنی همهٔ مجموعه‌های باز $V(E)$ که شامل E هستند. اندازهٔ هریک از مجموعه‌های $V(E)$ را معین می‌کنیم. این اندازه‌ها، مجموعه‌ای از عددهای مثبت را تشکیل می‌دهند. این مجموعهٔ عددها، کران پایین دارد (دست‌کم، عدد ۰ را می‌توان به‌عنوان کران پایین این مجموعه در نظر گرفت)؛ بنابراین، مجموعه دارای بزرگترین کران پایین است که آن را با $\mu_e E$ نشان می‌دهیم. $\mu_e E$ را ***اندازهٔ بیرونی مجموعهٔ*** E گویند.

$\mu_e E$ را اندازهٔ بیرونی مجموعهٔ E و $\mu_e CE$ را اندازهٔ بیرونیِ متمم آن، نسبت به بازهٔ $[a ,b]$ می‌گیریم.

اگر این رابطه برقرار باشد:

$$\mu_e E + \mu_e CE = b - a \tag{۳}$$

آن‌وقت، مجموعهٔ E را ***اندازه‌پذیر*** و عدد $\mu_e E$ را ***اندازهٔ*** آن گویند: $\mu E = \mu_e E$. در حالتی که برابری (۳) برقرار نباشد، گویند مجموعهٔ E اندازه‌ناپذیر است؛ مجموعهٔ اندازه‌ناپذیر، دارای اندازه‌ای نیست.

توجه کنیم که همیشه داریم:

$$\mu_e E + \mu_e CE \geq b - a \tag{۴}$$

بیشتر توضیح می‌دهیم. طول ساده‌ترین مجموعه‌ها (مثل بازه‌های باز و بسته)، ویژگی‌هایی دارند. مهم‌ترین آن‌ها را می‌آوریم:

۱. اگر مجموعه‌های $E_۱$ و E اندازه‌پذیر باشند و درضمن $E_۱ \subseteq E$، آنگاه

$$\mu E_۱ \leq \mu E$$

یعنی اندازهٔ زیرمجموعه‌ای از مجموعهٔ E، از اندازهٔ خودِ مجموعهٔ E تجاوز نمی‌کند.

۲. اگر مجموعه‌های $E_۱$ و $E_۲$ اندازه‌پذیر باشند، آن‌وقت مجموعهٔ $E = E_۱ + E_۲$ هم اندازه‌پذیر است و داریم:

$$\mu (E_۱ + E_۲) \leq \mu E_۱ + \mu E_۲$$

یعنی، اندازهٔ مجموع، از مجموع اندازه‌های جمله‌های جمع، تجاوز نمی‌کند.

۳. اگر مجموعه‌های E_i $(i = ۱, ۲, ...)$ اندازه‌پذیر و دوبه‌دو بدون اشتراک باشند: $E_i E_j = \theta$ $(i \neq j)$، آن‌وقت مجموع آن‌ها $E = \sum E_i$ اندازه‌پذیر است و داریم:

$$\mu (\sum E_i) = \sum \mu E_i$$

یعنی، اندازهٔ تعدادی محدود یا شمارا، از مجموعه‌هایی که دوبه‌دو بدون اشتراک‌اند، برابر است با مجموع اندازه‌های جمله‌های جمع.

این ویژگی اندازه را، **ویژگی جمع‌پذیری** آن گویند.

۴. اگر مجموعه را، همچون یک جسم صُلب حرکت دهیم، اندازهٔ آن تغییر نمی‌کند.

جالب است، ویژگی‌های اصلی طول، برای مفهوم عام‌تر اندازهٔ مجموعه‌ها هم، قابل قبول است. ولی اگر اندازه را، به هر مجموعهٔ دل‌خواهی از نقطه‌های واقع بر خط راست، نسبت دهیم، می‌توان بادقت ثابت کرد، چنین وضعی ممکن نیست. به همین دلیل، در تعریف‌ها، هم مجموعه‌های اندازه‌پذیر را مطرح کردیم، و هم مجموعه‌هایی که اندازه‌ای ندارند یا اندازه‌پذیر نیستند. با وجود این، گروه مجموعه‌های اندازه‌پذیر چنان گسترده است که این موقعیت، مشکلی جدی پدید نمی‌آورد. حتی ساختن مجموعه‌های اندازه‌ناپذیر، دشواری‌های معینی را به‌همراه دارد.

چند نمونه از مجموعه‌های اندازه‌پذیر می‌آوریم.

مثال ۱. *اندازهٔ مجموعهٔ تام کانتوریِ* P (بند ۴ را ببینید). برای ساختن مجموعهٔ P از بازهٔ بسته $[۰ , ۱]$، در آغاز، بازهٔ مجاوری به‌طول $\frac{۱}{۳}$ را کنار می‌گذاریم، بعد دو بازهٔ مجاور به‌طول $\frac{۱}{۹}$، سپس، چهار بازهٔ مجاور به‌طول $\frac{۱}{۲۷}$ وغیره. به‌طور کلی، در گام nام $۲^{n-۱}$ بازهٔ مجاور به‌طول $\frac{۱}{۳^n}$ را کنار می‌گذاریم. بنابراین، مجموع طول‌های همهٔ بازه‌هایی که کنار گذاشته‌ایم، برابر است با

$$S = \frac{۱}{۳} + \frac{۲}{۹} + \frac{۴}{۲۷} + \cdots + \frac{۲^{n-۱}}{۳^n} + \cdots$$

جمله‌های این مجموع، یک تصاعد هندسی نزولی، با جملهٔ اول $\frac{۱}{۳}$ و قدرنسبت $\frac{۲}{۳}$ تشکیل می‌دهند. بنابراین، برای مجموع آن‌ها داریم:

$$S = \frac{\frac{۱}{۳}}{۱ - \frac{۲}{۳}} = ۱$$

به این ترتیب، مجموع طول‌های همهٔ بازه‌های مجاورِ مجموعهٔ کانتوری برابر است با ۱. به‌زبان دیگر اندازهٔ مجموعهٔ باز G، یعنی اندازهٔ مجموعهٔ متمم P، برابر ۱ می‌شود و برای اندازهٔ خودِ مجموعهٔ P به‌دست می‌آید:

$$\mu P = ۱ - \mu G = ۱ - ۱ = ۰$$

همان‌طور که این مثال نشان می‌دهد، مجموعه‌ای می‌تواند توان پیوستار داشته باشد و، درعین‌حال، اندازهٔ آن برابر صفر باشد.

مثال ۲. *اندازهٔ مجموعهٔ* R، *همهٔ نقطه‌های گویای بازهٔ* $[۰ , ۱]$. در آغاز ثابت می‌کنیم: $\mu_e R = ۰$. در بند ۲ ثابت کردیم، مجموعهٔ R مجموعه‌ای شمارا است. نقطه‌های مجموعهٔ R را به‌صورت دنباله‌ای می‌نویسیم:

$$r_۱ , r_۲ , \ldots , r_n , \ldots$$

سپس $\varepsilon > ۰$ را در نظر می‌گیریم و نقطهٔ r_n را به‌وسیلهٔ بازهٔ δ_n به‌طول $\frac{\varepsilon}{۲^n}$ محاصره می‌کنیم. مجموع $\delta = \sum \delta_n$، مجموعه‌ای باز است که R را پوشانده است. بازه‌های δ_n می‌توانند دارای

اشتراک باشند. بنابراین

$$\mu(\delta) = \mu(\sum \delta_n) \le \sum \mu\, \delta_n = \sum \frac{\varepsilon}{2^n} = \varepsilon$$

از آن‌جا که ε را می‌توان به‌دل‌خواه کوچک گرفت، بنابراین $\mu_e R = 0$.

بعد، با توجه به (۳) داریم:

$$\mu_e R + \mu_e CR \ge 1$$

یعنی $\mu_e CR \ge 1$. چون CR، بخشی از پاره‌خط راست $[0, 1]$ است، بنابراین $\mu_e CR \le 1$. به این ترتیب

$$\mu_e R + \mu_e CR = 1$$

و از آن‌جا[1]

$$\mu R = 0 \quad , \quad \mu CR = 1 \qquad (5)$$

این مثال نشان می‌دهد، یک مجموعه می‌تواند همه‌جا روی خط راست فشرده باشد و، درعین‌حال، اندازه‌ای برابر صفر داشته باشد.

مجموعه‌های با اندازهٔ صفر، در بسیاری از مسأله‌های مربوط به نظریهٔ تابع، نقشی ندارند و، بنابراین، دارای اهمیت خاصی نیستند. از جمله، تابع $f(x)$، تنها وقتی به‌صورت ریمانی انتگرال‌پذیر است که کران‌دار باشد و مجموعهٔ نقطه‌های ناپیوستگی آن، اندازه‌ای برابر صفر داشته باشد. از این‌گونه مثال‌ها، می‌توان به هر تعداد لازم آورد.

تابع‌های اندازه‌پذیر. به یکی از مهم‌ترین کاربردهای مفهوم اندازه در مجموعه‌ها، می‌پردازیم. درواقع، می‌خواهیم دربارهٔ طبقه‌ای از تابع‌ها صحبت کنیم که درعمل، ضمن کار بـا آنـالیز ریاضی و نظریهٔ تابع‌ها، با آن سر و کار داریم. طرح درست مسأله چنین است. دنبالهٔ تابع‌های $\{f_n(x)\}$ که در مجموعهٔ E داده شده است، در نظر می‌گیریم و فرض می‌کنیم در هر نقطهٔ E (به‌جز به‌احتمالی در نقطه‌های مجموعهٔ N که اندازه‌ای برابر صفر دارند) هم‌گرا باشد. در

---

۱. با استدلالی مشابه، می‌توان ثابت کرد، اندازهٔ هر مجموعهٔ شمارا، که بر خط راست واقع باشد، برابـر صـفر است.

این صورت، گویند دنبالهٔ $\{f_n(x)\}$، **به‌تقریب در همه‌جا**، هم‌گراست.

اگر دربارهٔ تابع‌های پیوسته، عمل ساختن حد دنبالهٔ تابع‌های «به‌تقریب در همه‌جا هم‌گرا» و عمل‌های جبری را به‌تکرار به‌کار بریم، چه تابع‌هایی به‌دست می‌آید؟

برای پاسخ به این پرسش، به چند مفهوم تازه نیاز داریم.

$f(x)$ روی مجموعهٔ E تعریف شده است و α را عددی دل‌خواه و حقیقی فرض می‌کنیم. مجموعهٔ همهٔ نقطه‌های E را، که برای آن‌ها داریم $f(x) > \alpha$، به‌صورت

$$E\,[\,f(x) > \alpha\,]$$

نشان می‌دهیم. به‌عنوان نمونه، اگر تابع $f(x)$ روی بازهٔ $[0\,,1]$ تعریف شده باشد، و در این بازه داشته باشیم $f(x)=x$، آنگاه مجموعهٔ $E\,[f(x) > \alpha]$، برای $\alpha < 0$ برابر $[0\,,1]$، برای $0 \le \alpha < 1$ برابر $(\alpha\,,1]$ و برای $\alpha \ge 1$ تهی است.

تابع $f(x)$ را که روی مجموعهٔ E تعریف شده است، اندازه‌پذیر گویند، وقتی که خودِ مجموعهٔ E اندازه‌پذیر و، درضمن، مجموعهٔ $E\,[f(x) > \alpha]$ هم، برای هر عدد حقیقی α، اندازه‌پذیر باشد.

می‌توان ثابت کرد، که هر تابع پیوسته‌ای که روی یک بازه تعریف شده باشد، اندازه‌پذیر است. ولی بسیاری از تابع‌های ناپیوسته هم، اندازه‌پذیرند؛ از جمله تابع دیریلکه، که برای عددهای گنگ بازهٔ $[0\,,1]$ برابر ۱، و برای سایر عددهای این بازه، برابر ۰ است.

بدون اثبات یادآوری می‌کنیم، تابع‌های اندازه‌پذیر دارای این ویژگی‌ها هستند:

۱. اگر $f(x)$ و $\varphi(x)$ در یک مجموعهٔ E، تابع‌هایی اندازه‌پذیر باشند، آن‌وقت تابع‌های

$$f+\varphi \;\;,\;\; f-\varphi \;\;,\;\; f.\varphi \;\; \text{و} \;\; \frac{f}{\varphi}$$

هم اندازه‌پذیرند (دربارهٔ $\frac{f}{\varphi}$، با شرط $\varphi \ne 0$).

این ویژگی نشان می‌دهد، انجام عمل‌های جبری روی تابع‌های اندازه‌پذیر، منجر به تابع‌هایی اندازه‌پذیر می‌شود.

۲. اگر دنبالهٔ تابع‌های اندازه‌پذیر $\{f_n(x)\}$، که در مجموعهٔ E تعریف شده‌اند، به‌تقریب همه‌جا به‌سمت تابع $f(x)$ هم‌گرا باشد، آن‌وقت، این تابع هم اندازه‌پذیر است.

به این ترتیب، عمل ساختن حدِ به‌تقریب همه‌جا هم‌گرا، از دنبالهٔ تابع‌های اندازه‌پذیر،

منجر به تابعی اندازه‌پذیر می‌شود.

این ویژگی‌های تابع‌های اندازه‌پذیر را، لِه‌بِگ آورده است. د. ف. یه‌گوروف و ن. ن. لوزین، ریاضی‌دانان شوروی، بررسی‌های ژرفی دربارهٔ تابع‌های اندازه‌پذیر انجام داده‌اند. به‌ویژه لوزین ثابت کرد: هر تابع اندازه‌پذیر را، که روی بازه‌ای داده شده باشد، با جابه‌جا کردن مقدارهای آن روی مجموعه‌ای با اندازه‌ای به‌قدر دل‌خواه کوچک، می‌توان به تابعی پیوسته تبدیل کرد.

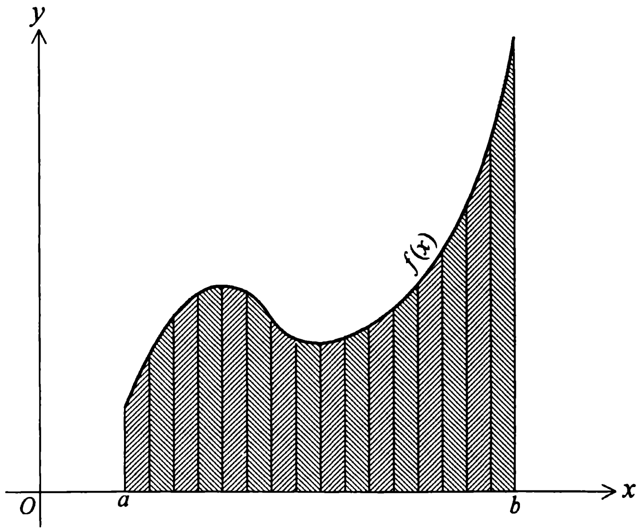
این نتیجه‌گیری لوزین و ویژگی‌های دیگری که دربارهٔ تابع‌های اندازه‌پذیر آوردیم، نشان می‌دهند، تابع‌های اندازه‌پذیر، جزو همان طبقه از تابع‌ها هستند که در آغاز این بحث، از آن صحبت کردیم. تابع‌های اندازه‌پذیر، درضمن، اهمیت زیادی در نظریهٔ انتگرال‌گیری دارند و، از این راه، می‌توان مفهوم انتگرال را به این صورت تعمیم داد که، هر تابع اندازه‌پذیر کران‌دار، انتگرال‌پذیر است. و این همان چیزی است که در بند بعد دربارهٔ آن صحبت می‌کنیم.

۶. انتگرال لِه‌بِگ

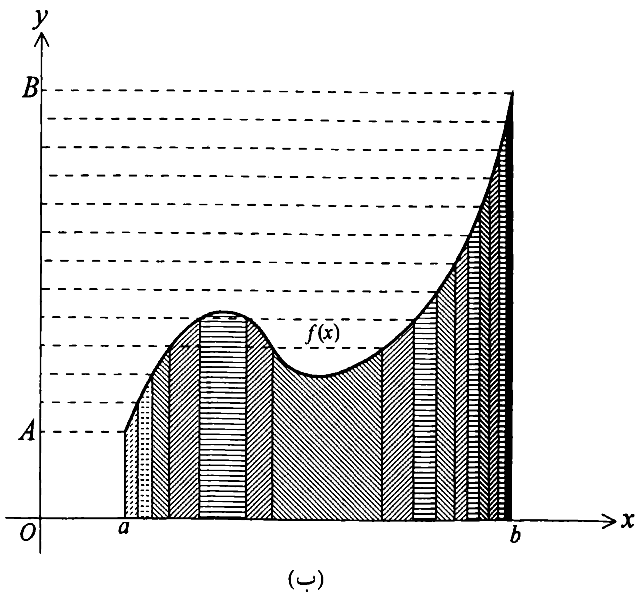
به بحث مرکزی این بخش می‌رسیم: تعریف و توضیح ویژگی‌های انتگرال له‌بِگ.

برای آشنایی با ساختمان انتگرال له‌بگ، مثالی می‌آوریم. فرض کنید با تعداد زیادی سکه، که ارزش‌های مختلفی دارند، سر و کار داشته باشیم و بخواهیم جمع کل مبلغ این سکه‌ها را پیدا کنیم. برای رسیدن به این هدف، با یکی از دو روش می‌توان عمل کرد. می‌توان سکه‌ها را یکی یکی کنار گذاشت و ارزش هر سکهٔ جدید را، به مجموع مبلغ سکه‌های قبلی افزود. ولی به ترتیب دیگری هم می‌توان عمل کرد: سکه‌ها را در دسته‌های جداگانه‌ای قرار می‌دهیم، به‌نحوی که سکه‌های دسته، ارزشی برابر داشته باشند؛ بعد تعداد هر دسته را می‌شماریم و عدد حاصل را در ارزش سکهٔ متناظر آن ضرب می‌کنیم؛ سپس، عددهایی را که به این ترتیب به‌دست می‌آید، با هم جمع می‌کنیم. روند اول محاسبه، متناظر با انتگرال ریمان و روش دوم متناظر با انتگرال له‌بِگ است.

از سکه‌ها بگذریم و به تابع‌ها بپردازیم. برای محاسبهٔ انتگرال ریمان، دامنهٔ تعریف تابع (یعنی محور طول؛ شکل ۲ـالف) را به بخش‌های کوچک تقسیم می‌کنیم، درحالی که برای



(الف)



(ب)

شکل ۲

محاسبهٔ انتگرال له‌بگ از تقسیم بُرد یعنی حوزهٔ مقدارهای تابع (محور عرض؛ شکل ۲ـب) استفاده می‌کنیم. این روش، پیش از له‌بِگ هم، دربارهٔ محاسبهٔ انتگرال تابع‌هایی که خصلت نوسانی دارند، به‌کار می‌رفت؛ ولی له‌بگ، برای نخستین بار، توانست آن را در حالت کلی

بررسی کند و برپایهٔ نظریهٔ اندازه‌ها، بنیان‌های آن را به‌طور دقیق آماده سازد.

ببینیم چه رابطه‌ای بین اندازهٔ مجموعه‌ها و انتگرال له‌بِگ وجود دارد؟ E را مجموعه‌ای اندازه‌پذیر واقع بر بازه‌ای مثل $[a, b]$ می‌گیریم. تابع $\varphi(x)$ را، با این شرط انتخاب می‌کنیم: به‌ازای مقدارهایی از x که متعلق به E باشد، برابر واحد، و برای مقدارهایی از x که به E تعلق ندارد، برابر صفر شود. به زبان دیگر فرض می‌کنیم:

$$\varphi(x) = \begin{cases} ۱ & (x \in E) \\ ۰ & (x \bar{\in} E) \end{cases}$$

در این صورت، $\varphi(x)$ را **تابعِ مشخصه مجموعهٔ** E گویند. اکنون، این انتگرال را در نظر می‌گیریم:

$$I = \int_a^b \varphi(x)\, dx$$

عادت کرده‌ایم، انتگرال را برابر مساحت شکلی مثل D به‌حساب آوریم که به‌وسیلهٔ محور طول، خط‌های راست $x = a$ و $x = b$ و منحنی $y = \varphi(x)$ محصور شده باشد (بخش دوم را در جلد اول ببینید). چون در این‌جا، «ارتفاع» شکل D، برای نقطه‌های $x \in E$، و تنها برای همین نقطه‌ها، مخالف صفر و برابر واحد است، بنابراین (بنا بر رابطهٔ مساحت که برابر است با حاصل‌ضرب طول در عرض)، مساحت آن از نظر عددی، باید برابر باشد با طول (اندازهٔ) مجموعهٔ E. به این ترتیب، I برابر است با اندازهٔ مجموعهٔ E:

$$I = \mu E \tag{۶}$$

که له‌بِگ هم، انتگرال از تابع $\varphi(x)$ را، به همین صورت تعریف کرده است.

خواننده باید به این مطلب توجه کند که، برابری (۶) عبارت است از **تعریف** $\int_a^b \varphi(x)\, dx$، به‌عنوان انتگرال لِه‌بگ. ممکن است با حالتی روبه‌رو باشیم که انتگرال I، به مفهومی که در بخش دوم (جلد اول) می‌شناسیم، یعنی به‌عنوان حد مجموع‌های انتگرالی وجود نداشته باشد. ولی در چنین وضعی هم، انتگرال I، به‌عنوان انتگرال له‌بِگ وجود دارد و برابر است با μE.

برای نمونه، انتگرال تابع دیریکله، یعنی $\Phi(x)$ را می‌آوریم. این تابع، در نقطه‌های گویای پاره‌خط راست $[۰, ۱]$ برابر صفر و در نقطه‌های گنگ این پاره‌خط راست، برابر واحد است.

از آن‌جا که بنابر (۵)، اندازهٔ مجموعهٔ نقطه‌های گنگِ پاره‌خط راست [۰ , ۱] برابر واحد است، بنابراین انتگرال له‌بِگ

$$\int_{۰}^{۱} \Phi(x)\, dx$$

برابر واحد می‌شود. به‌سادگی قابل تحقیق است که، برای این تابع، انتگرالِ ریمان وجود ندارد.

یک گزارهٔ کمکی. اکنون فرض کنید $f(x)$ تابعی کران‌دار و اندازه‌پذیر روی بازهٔ $[a\ ,b]$ باشد. ثابت می‌کنیم، هر تابعی از این‌گونه را، می‌توان با هر دقت لازم، به‌صورت ترکیبی خطی از تابع‌های مفسر مجموعه‌ها درآورد. برای این منظور، بازه‌ای روی محور عرض بین بزرگترین کران پایین A و کوچکترین کران بالا B از تابع، به‌وسیلهٔ نقطه‌های $y_۰=A$، $y_۱$، $y_۲$، ...، $y_n=B$، به بازه‌های به‌طول کمتر از ε بخش می‌کنیم که در آن، ε عدد ثابت دل‌خواهی است. سپس، اگر در نقطهٔ $x \in [a\ ,b]$ داشته باشیم:

$$y_۰ = A\ , y_۱\ , \ldots\ , y_n = B \quad , \quad y_i \le f(x) < y_{i+۱} \quad , \quad (i = ۰\ , ۱\ , \ldots\ , n-۱)$$

آن‌وقت، در این نقطه فرض می‌کنیم:

$$\varphi(x) = y_i$$

و اگر در نقطهٔ x، داشته باشیم:

$$f(x) = y_n = B$$

آن‌وقت فرض می‌کنیم:

$$\varphi(x) = y_n$$

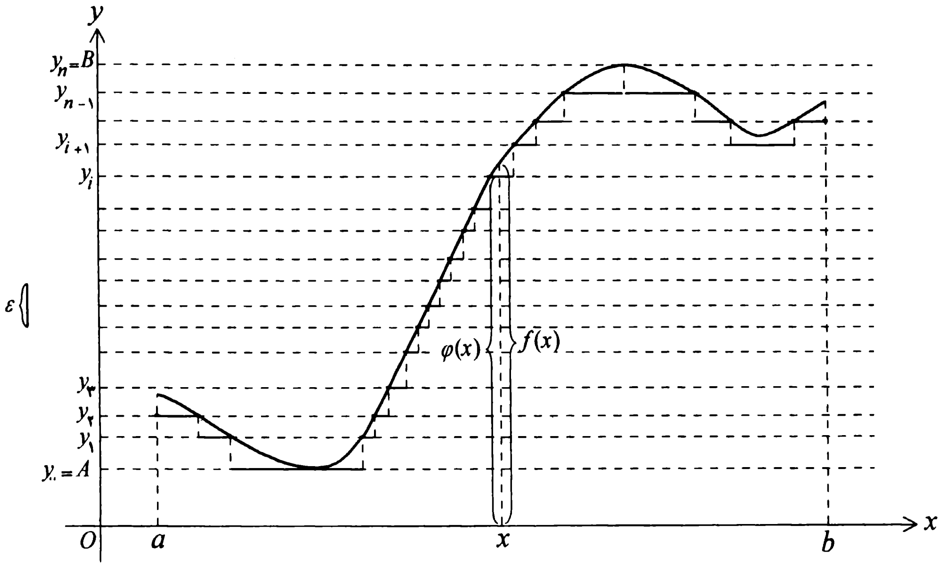
ساختمان تابع $\varphi(x)$ در شکل ۳ داده شده است.

بنابر ساختمان تابع $\varphi(x)$، در هر نقطهٔ بازهٔ $[a\ ,b]$ داریم:

$$|f(x) - \varphi(x)| < \varepsilon$$

به‌جز این، چون تابع $\varphi(x)$، تنها تعداد محدودی مقدارهای $y_۰$، $y_۱$، ...، y_n را قبول می‌کند، می‌توان آن را به این صورت نوشت:

$$\varphi(x) = y_۰ \cdot \varphi_۰(x) + y_۱ \cdot \varphi_۱(x) + \cdots + y_n \cdot \varphi_n(x) \qquad (۷)$$



شکل ۳

که در آن، $\varphi_i(x)$ عبارت است از تابع مشخصه مجموعه‌ای که برای آن $y_i = \varphi(x)$، یعنی $y_i \le f(x) < y_{i+۱}$ (در هر نقطهٔ $x \in [a, b]$، تنها یک جمله در سمت راست رابطهٔ (۷) مخالف صفر است). و این، همان گزاره‌ای است که لازم داریم.

تعریف انتگرال لِه‌بِگ. به تعریف انتگرال لهِبگ از یک تابع کران‌دار و اندازه‌پذیرِ دل‌خواه برمی‌گردیم. چون تابع $\varphi(x)$ خیلی کم با تابع $f(x)$ فرق دارد، بنابراین به‌عنوان مقدار تقریبی تابع $f(x)$، می‌توان انتگرال تابع $\varphi(x)$ را پذیرفت. ولی با توجه به این که تابع‌های $\varphi_i(x)$، تابع‌های مشخصه مجموعه‌ها هستند، با استفاده از قاعده‌های معمولی محاسبهٔ انتگرال، به‌دست می‌آید:

$$\int_a^b \varphi(x)\,dx = \int_a^b \{y_۰\varphi_۰(x) + y_۱\varphi_۱(x) + \cdots + y_n\varphi_n(x)\}dx =$$

$$= y_۰ \int_a^b \varphi_۰(x)\,dx + y_۱ \int_a^b \varphi_۱(x)\,dx + \cdots + y_n \int_a^b \varphi_n(x)\,dx =$$

$$= y_۰ \cdot \mu e_۰ + y_۱ \cdot \mu e_۱ + \cdots + y_n \cdot \mu e_n$$

که در آن μe_i عبارت است از اندازهٔ مجموعهٔ e_i از xهایی که برای آن‌ها داریم:

$$y_i \leq f(x) < y_{i+1}$$

به این ترتیب، مقدار تقریبی انتگرال لهِبگ از تابع $f(x)$ عبارت است از «مجموع انتگرالی لهِبگ»:

$$S = y_0 \cdot \mu e_0 + y_1 \cdot \mu e_1 + \cdots + y_n \cdot \mu e_n$$

به این ترتیب، انتگرال لهِبگ، به‌عنوان حد مجموع انتگرالی لهِبگ (S)، تعریف می‌شود، وقتی که

$$\max|y_{i+1} - y_i| \to 0$$

که متناظر است با هم‌گرایی یکنواختِ تابع $\varphi(x)$، به‌سمت تابع $f(x)$.

می‌توان ثابت کرد: مجموع‌های انتگرالی لهِبگ، برای هر تابع کران‌دار اندازه‌پذیر، دارای حدند، یعنی هر تابع کران‌دار اندازه‌پذیر، به مفهوم لهِبگ، انتگرال‌پذیر است. انتگرال لهبگ را دربارهٔ گروه‌هایی از تابع‌های اندازه‌پذیر بی‌کران هم می‌توان تعمیم داد. ولی در این‌جا به آن نمی‌پردازیم.

ویژگی‌های انتگرال لهِبگ. انتگرال لهِبگ همهٔ ویژگی‌های خوب انتگرال معمولی را دارد: انتگرال مجموع برابر است با مجموع انتگرال‌ها؛ ضریب ثابت را می‌توان از زیر نماد انتگرال بیرون آورد وغیره. ولی انتگرال لهِبگ، ویژگی جالب دیگری هم دارد که به انتگرال معمولی مربوط نمی‌شود. اگر تابع‌های اندازه‌پذیر $f_n(x)$، برای هر $n \in N$ و هر x از بازهٔ $[a, b]$، کران یکنواختی داشته باشند:

$$|f_n(x)| < K$$

و دنبالهٔ $\{f_n(x)\}$، به‌تقریب در همه‌جا، به‌سمت تابع $f(x)$ هم‌گرا باشد، آن‌وقت

$$\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx$$

به‌زبان دیگر، انتگرال لهِبگ اجازه می‌دهد، بی‌وقفه به‌سمت حد عبور کنیم. همین ویژگی، انتگرال لهِبگ را به‌صورتی بی‌اندازه مناسب و راحت درآورده است؛ و در برخی حالت‌ها، خود را به‌عنوان تنها وسیلهٔ بررسی نشان می‌دهد. انتگرال لهِبگ، به‌ویژه در نظریهٔ رشته‌های

مثلثاتی، در نظریهٔ فضاهای تابعی (بخش نوزدهم را ببینید) و دیگر شاخه‌های ریاضیات، کاربرد فراوان دارد.

مثالی می‌آوریم. $f(x)$ را تابعی متناوب با دورهٔ تناوب ۲π و

$$\frac{a_۰}{۲}+\sum_{n=۱}^{\infty}(a_n \cos nx + b_n \sin nx)$$

را، رشتهٔ فوریهٔ نظیر آن فرض کنید. اگر در مثل، تابع $f(x)$ پیوسته باشد، به‌سادگی می‌توان ثابت کرد:

$$\frac{۱}{\pi}\int_{۰}^{۲\pi} f^{۲}(x)\,dx = \frac{a_۰^{۲}}{۲}+\sum_{n=۱}^{\infty}(a_n^{۲}+b_n^{۲}) \qquad (۸)$$

این اتحاد را، برابری «پارسه‌وال» می‌نامند. این پرسش پیش می‌آید: برای چه گروهی از تابع‌های متناوب، برابری پارسه‌والِ (۸) برقرار است؟ پاسخ به این پرسش چنین است: برابری پارسه‌والِ (۸)، وقتی و تنها وقتی برقرار است که تابع $f(x)$ در بازهٔ $[۰ , ۲\pi]$ اندازه‌پذیر باشد و تابع $f^{۲}(x)$ در این بازه، دارای انتگرال به‌مفهوم لِه‌بِگ باشد.

بخش شانزدهم

جبر خطی

د. ک. فادِیف

۱. موضوع جبر خطی و دستگاه آن

تابع‌های خطی و ماتریس‌ها[1]. ساده‌ترین تابع‌های با یک متغیر، ***تابع خطی*** $l(x) = ax + b$ است؛ و همان‌طور که می‌دانیم، نمودار این تابع، ساده‌ترینِ منحنی‌ها، یعنی خط راست است.

با وجود این، تابع خطی یکی از مهم‌ترین تابع‌هاست. مطلب بر سر این است که هر منحنی «هموار» در بخش کوچکی از خود، به خط راست شباهت دارد؛ هرچه این بخش منحنی کوچکتر باشد، به خط راست بیشتر و نزدیک‌تر می‌چسبد. این موضوع، به زبان نظریهٔ تابع‌ها به این معنی است که هر تابع «هموار» (پیوسته ـ مشتق‌پذیر)، به‌ازای تغییر کوچک متغیر مستقل، به یک تابع خطی نزدیک است. تابع خطی را می‌توان با این رفتار مشخص کرد که نمو آن، متناسب است با نمو متغیر مستقل. درواقع داریم:

$$\Delta l(x) = l(x_{\circ} + \Delta x) - l(x_{\circ}) = a(x_{\circ} + \Delta x) + b - (ax_{\circ} + b) = a\Delta x$$

و برعکس، اگر داشته باشیم $\Delta l(x) = a\Delta x$، در آن صورت داریم:

$$l(x) - l(x_{\circ}) = a(x - x_{\circ})$$

و درنتیجه

$$l(x) = ax + l(x_{\circ}) - ax_{\circ} = ax + b$$

که در آن $b = l(x_{\circ}) - ax_{\circ}$. ولی از محاسبهٔ دیفرانسیلی می‌دانیم، در نمو هر تابع مشتق‌پذیر،

---

۱. ماتریس را، در زبان فارسی «آرایه» و «زهدان» هم گفته‌اند. ـ م.

به‌طور طبیعی، بخش اصلی آن به‌نام دیفرانسیل تابع جدا می‌شود که با نمو متغیر مستقل متناسب است؛ اختلاف نمو تابع از دیفرانسیل آن، بی‌نهایت کوچکی است که نسبت به نمو متغیر مستقل، از مرتبه‌ای بالاتر است. بنابراین تابع مشتق‌پذیر، به‌ازای تغییر بی‌نهایت کوچک متغیر مستقل، درواقع، با تقریب بی‌نهایت کوچکی از مرتبهٔ بالاتر، به تابع خطی نزدیک است.

وضع دربارهٔ تابع‌های با چند متغیر هم، به همین ترتیب است.

تابع خطی با چند متغیر به تابعی به‌صورت

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + b$$

گفته می‌شود. در حالت $b = 0$، تابع خطی را همگِن (متجانس) گویند. دو ویژگی مشخص‌کنندهٔ تابع‌های خطی با چند متغیر را می‌آوریم:

۱. نمو تابع خطی، با این فرض که تنها یکی از متغیرهای مستقل، نموی اختیار کند و مقدار بقیهٔ متغیرهای مستقل، بی‌تغییر بماند، متناسب است با نمو این متغیر مستقل.

۲. نمو تابع خطی، با این فرض که همهٔ متغیرهای مستقل دارای نموی باشند، برابر است با مجموع جبری نموهایی که از راه تغییر هریک از متغیرها به‌طور جداگانه، به‌دست می‌آید.

نقش تابع خطی با چند متغیر، بین همهٔ تابع‌هایی که چند متغیر دارند، شبیه نقشی است که تابع خطی با یک متغیر در میان همهٔ تابع‌های با یک متغیر به‌عهده گرفته است. هر تابع «هموار» (یعنی تابعی که دارای مشتق‌های نسبی پیوسته نسبت به همه متغیرها باشد)، به‌ازای تغییر کوچک متغیرهای مستقل، به یک تابع خطی نزدیک می‌شود. درواقع، نمو تابعی مثل

$$w = f(x_1, x_2, \ldots, x_n)$$

با تقریب بی‌نهایت کوچک‌های از مرتبهٔ بالاتر، برابر است با دیفرانسیل کلی

$$\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

که تابعی است خطی و همگِن از نمو متغیرهای مستقل dx_1، dx_2، ...، dx_n. از این‌جا نتیجه می‌شود که، خودِ تابع w، که برابر است با مجموع مقدار نخستینِ آن و نموش، را

می‌توان برحسب متغیرهای مستقل، به‌ازای تغییر کوچک آن‌ها بیان کرد. این رابطه به‌صورت تابعی خطی و ناهمگن با تقریب بی‌نهایت کوچک‌های از مرتبه‌های بالاتر می‌باشد.

مسأله‌هایی که، برای حل آن‌ها، به تابع‌های با چند متغیر نیاز داریم، ضمن بررسی کمیتی که مقدار آن به چند عامل بستگی دارد، به‌وجود می‌آید. اگر این بستگی ـ یعنی بستگی یک کمیت با چند عامل دیگر ـ خطی باشد، مسألهٔ ناشی از آن را خطی گویند. با توجه به آن‌چه دربارهٔ ویژگی‌های تابع‌های خطی گفتیم، می‌توان مسألهٔ خطی را، با این ویژگی‌ها مشخص کرد:

۱. *ویژگی متناسب بودن.* نتیجهٔ عمل هر عامل جداگانه، متناسب با مقدار آن است.

۲. *ویژگی استقلال.* نتیجهٔ کلی عمل، برابر است با مجموع نتیجه‌هایی که از عمل عامل‌های جداگانه به‌دست آمده است.

این که هر تابع همواری را، به‌ازای تغییرهای کوچک متغیرها، می‌توان در تقریب اول، با یک تابع خطی عوض کرد، بازتابی از یک اصل کلی است: هر مسأله‌ای را که در بارهٔ تغییر کمی در نتیجهٔ‌عمل چندعامل‌بحث‌می‌کند، می‌توان در تقریب‌اول و به‌ازای تأثیرهای کوچک، همچون یک مسألهٔ خطی، یعنی مسأله‌ای که دارای ویژگی‌های تناسب و استقلال است، در نظر گرفت. اغلب بررسی مسأله به این صورت، می‌تواند درعمل، نتیجه‌ای رضایت‌بخش داشته باشد (از جمله، در نظریهٔ سنتی کشسانی، نظریهٔ نوسان‌های کوچک وغیره).

خودِ کمیت‌های فیزیکی هم، اغلب به یاری چند عدد مشخص می‌شوند (نیرو، به‌یاری سه تصویری که روی محورهای مختصات دارد؛ حالت فشردهٔ جسم کشسان در نقطهٔ مفروض، به‌یاری شش مؤلفه به‌نام تانسور تنش وغیره). به این ترتیب، ضرورت بررسی هم‌زمان چند تابع از چند متغیر و در تقریب نخست، چند تابع خطی پیش می‌آید.

تابع خطیِ با یک متغیر، از نظر ویژگی‌های خود، آن‌قدر ساده است که نیازی به بررسی خاص ندارد. ولی دربارهٔ تابع‌های خطی با چند متغیر چنین نیست؛ وجود متغیرهای بیشتر، ویژگی‌های خاصی را به‌دنبال دارد که باید به بررسی آن‌ها پرداخت. اگر از حالت یک تابع از متغیرهای x_1، x_2، ... ، x_n، به‌سراغ مجموعه‌ای از چند تابع y_1، y_2، ... ، y_m، نسبت به همین متغیرها، برویم، باز هم کار پیچیده‌تر می‌شود. در این‌جا به‌عنوان تقریب اول، با

مجموعه‌ای از چند تابع خطی سر و کار داریم:

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n + b_1 \ , \\ y_2 &= a_{21}x_1 + \cdots + a_{2n}x_n + b_2 \ , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n + b_m \end{aligned}$$

مجموعهٔ تابع‌های خطی، موضوع ریاضی بسیار بغرنجی را تشکیل می‌دهد، ولی درعوض، بررسی آن مضمونی بی‌اندازه جالب و غنی دارد.

بررسی تابع‌های خطی و مجموعه‌های مربوط به آن‌ها، در درجهٔ اول، موضوع شاخه‌ای از جبر، به‌نام **جبر خطی** است.

از نظر تاریخی، نخستین مسألهٔ جبر خطی، عبارت است از حل دستگاه معادله‌های خطیِ

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \ , \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \ , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

ساده‌ترین حالت این مسأله، در دورهٔ دبیرستانی جبر مقدماتی بررسی می‌شود. مسألهٔ مربوط به جست‌وجوی روش‌های ساده‌تر حل عددی دستگاه‌ها، وقتی n عددی بزرگ باشد، توجه خاص بسیاری از دانشمندان را تا به امروز، به خود جلب کرده است، چرا که حل عددی دستگاه‌ها، به‌خاطر محاسبه‌های طولانی و بررسی‌های مفصل، نیاز به صرف وقت زیادی دارد.

تابع‌های خطی همگن (متجانس) را، ***صورت‌های خطی*** (یا فرم‌های خطی) هم می‌گویند. دستگاه مفروض صورت‌های خطی

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \ , \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n \end{aligned}$$

به‌یاری دستگاه ضریب‌های آن شرح داده می‌شود، زیرا ویژگی‌های این دستگاه صورت‌ها،

تنها به مقدار عددی ضریب‌های بستگی دارد، و نام متغیرها اهمیتی جدی ندارد. به‌عنوان نمونه، دو دستگاه صورت‌های

$$\begin{aligned} &3x_1 + x_2 - x_3, \qquad & 3t_1 + t_2 - t_3, \\ &2x_1 + x_2 + 3x_3, \qquad \text{و} \qquad & 2t_1 + t_2 + 3t_3, \\ &x_1 - x_2 - x_3, \qquad & t_1 - t_2 - t_3, \end{aligned}$$

ویژگی‌های یکسانی دارند و تفاوت متغیرهای آن‌ها را می‌توان بی‌اهمیت دانست.

مجموعهٔ ضریب‌های دستگاه صورت‌های خطی را، به‌طور طبیعی، به‌صورت یک جدول مستطیلی نشان می‌دهند:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

این‌گونه جدول‌ها را ***ماتریس*** می‌نامند. عددهای a_{ij} ***درایه‌های*** ماتریس را تشکیل می‌دهند (گاه، درایه را ***عضو*** ماتریس یا ***عنصر*** ماتریس هم گفته‌اند). نیاز به بررسی ماتریس‌ها، ناشی از نیازی است که از خود موضوع جبر خطی سرچشمه می‌گیرد.

مهم‌ترین حالت‌های خاص ماتریس، عبارت‌اند از: ۱) ماتریسی که تنها شامل یک ستون باشد که آن را، به‌طور ساده، ***ماتریس ستونی*** می‌نامند؛ ۲) ماتریسی که تنها از یک سطر تشکیل شده باشد که ***ماتریس سطری*** نام دارد؛ و سرانجام ۳) ***ماتریس مربعی***، یعنی ماتریسی که در آن، تعداد سطرها برابر با تعداد ستون‌ها باشد. تعداد سطرها (یا تعداد ستون‌ها) را در ماتریس مربعی، **مرتبهٔ** آن گویند. ماتریس (a)، که تنها شامل یک عدد (یعنی یک سطر و یک ستون) است، با این عدد شناسایی می‌شود.

طبیعی است، در رابطه با عمل‌های ساده روی مجموعهٔ صورت‌های خطی، باید تعریف‌هایی برای این عمل‌ها روی ماتریس پیدا کنیم.

دو دستگاه شامل صورت‌های خطی را درنظر می‌گیریم:

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n \end{aligned}$$

و

$$z_1 = b_{11}x_1 + \cdots + b_{1n}x_n \ ,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$z_m = b_{m1}x_1 + \cdots + b_{mn}x_n$$

این صورت‌ها را، به‌ترتیب، با هم جمع می‌کنیم:

$$y_1 + z_1 = (a_{11}+b_{11})\,x_1 + \cdots + (a_{1n}+b_{1n})\,x_n \ ,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$y_m + z_m = (a_{m1}+b_{m1})\,x_1 + \cdots + (a_{mn}+b_{mn})\,x_n$$

طبیعی است، ماتریسی را که از این دستگاه صورت‌ها به‌دست می‌آید، یعنی ماتریس

$$\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

را مجموع ماتریس‌های $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ و $\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$ به‌حساب آوریم.

به همین ترتیب، حاصل‌ضرب ماتریس $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ در عدد c، به‌عنوان ماتریس ضریب‌ها در دستگاه صورت‌های cy_1، cy_2، ...، cy_m تعریف می‌شود که در آن، y_1، y_2، ...، y_m عبارت‌اند از صورت‌هایی که ضریب‌های آنها، ماتریس $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ را تشکیل می‌دهند. از این تعریف نتیجه می‌شود:

$$c\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \cdots & \cdots & \cdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix}$$

سرانجام، عمل ضرب یک ماتریس در ماتریس دیگر، به این ترتیب تعریف می‌شود. فرض کنید داشته باشیم:

$$\left.\begin{array}{l} z_1 = a_{11}y_1 + \cdots + a_{1m}y_m \ , \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ z_k = a_{k1}y_1 + \cdots + a_{km}y_m \end{array}\right\} \qquad (1)$$

و

$$y_1 = b_{11}x_1 + \cdots + b_{1n}x_n \quad ,$$
$$\dotfill$$
$$y_m = b_{m1}x_1 + \cdots + b_{mn}x_n$$

اگر عبارت‌های y_1، y_2، ... ، y_m که برحسب x_1، x_2، ... ، x_n داده شده‌اند، در (۱) قرار دهیم، آن‌وقت z_1، z_2، ... ، z_k هم برحسب x_1، x_2، ... ، x_n و به‌صورت فرم‌های خطی بیان می‌شوند:

$$z_1 = c_{11}x_1 + \cdots + c_{1n}x_n \quad ,$$
$$\dotfill$$
$$z_k = c_{k1}x_1 + \cdots + c_{kn}x_n$$

ماتریس ضریب‌ها، یعنی $\begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{k1} & \cdots & c_{kn} \end{pmatrix}$، حاصل‌ضرب ماتریس‌های $\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{km} \end{pmatrix}$ و $\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$ نامیده می‌شود و آن را به این صورت می‌نویسند:

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{km} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

به‌سادگی می‌توان درایه‌های حاصل‌ضرب دو ماتریس را، برحسب درایه‌های عامل‌های ضرب، بیان کرد. درایهٔ c_{ij} عبارت است از ضریب x_j در عبارت z_i برحسب x_1، ... ، x_n.
ولی $z_i = a_{i1}y_1 + \cdots + a_{im}y_m$ و

$$y_1 = \cdots + b_{1j}x_j + \cdots \quad ,$$
$$\dotfill$$
$$y_m = \cdots + b_{mj}x_j + \cdots$$

و بنابراین

$$z_i = \cdots + (a_{i1}b_{1j} + \cdots + a_{im}b_{mj})x_j + \cdots \quad ,$$

از آن‌جا

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{im}b_{mj}$$

به‌این ترتیب، درایهٔ سطر iام و ستون jام حاصل‌ضرب دو ماتریس، برابر است با مجموع حاصل‌ضرب‌های درایه‌های سطر iام عامل اول در درایه‌های متناظر ستون jام عامل دومِ ضرب. مثال

$$\begin{pmatrix} ۲ & ۱ & ۳ \\ ۳ & ۱ & ۱ \end{pmatrix} \begin{pmatrix} ۳ & ۱ \\ ۱ & ۲ \\ ۲ & ۴ \end{pmatrix} =$$

$$= \begin{pmatrix} ۲\times ۳+ ۱\times ۱+ ۳\times ۲ & ۲\times ۱+ ۱\times ۲+ ۳\times ۴ \\ ۳\times ۳+ ۱\times ۱+ ۱\times ۲ & ۳\times ۱+ ۱\times ۲+ ۱\times ۴ \end{pmatrix} = \begin{pmatrix} ۱۳ & ۱۶ \\ ۱۲ & ۹ \end{pmatrix}$$

با آن که ماتریس موضوعی به‌اصطلاح «مرکب» است و در ترکیب آن درایه‌های زیادی وجود دارد، بهتر و ساده‌تر آن است که آن را با یک حرف نشان دهیم و، درضمن، برای جمع و ضرب از همان نمادهای عادی استفاده کنیم. برای نشان دادن ماتریس‌ها، حرف‌های بزرگ لاتینی را به‌کار می‌بریم. به‌کار بردن این نمادهای کوتاه، کار توضیح را، در نظریهٔ ماتریس‌ها، ساده‌تر می‌کند؛ این رابطه‌های کوتاه، که یادآور دستورهای معمولی است، رابطه‌های پیچیده‌ای را معرفی می‌کنند که به مجموعهٔ عددها (یعنی درایه‌های ماتریس‌هایی که در این رابطه‌ها شرکت دارند) مربوط می‌شود. برای نمونه، مجموعهٔ صورت‌های خطیِ

$$\begin{array}{c} a_{۱۱}x_{۱} + \cdots + a_{۱n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m۱}x_{۱} + \cdots + a_{mn}x_n \end{array}$$

در علامت‌گذاری ماتریسی، به‌صورت AX نشان داده می‌شود که در آن، A ماتریس ضریب‌ها است و X ستونی از متغیرهای $x_۱$، ... ، x_n. دستگاه معادله‌های خطی

$$\begin{array}{c} a_{۱۱}x_{۱} + \cdots + a_{۱n}x_n = b_۱ \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m۱}x_{۱} + \cdots + a_{mn}x_n = b_n \end{array}$$

را به این صورت می‌نویسند:

$$AX = B$$

که در آن، A عبارت است از ماتریس ضریب‌ها، X ماتریس ستونی از متغیرها و B ماتریس ستونی عددهای ثابت است.

البته، عمل‌های اصلی روی ماتریس‌ها ـ یعنی دو عمل جمع و ضرب ـ همیشه معین نیستند. عمل جمعِ ماتریس‌ها، وقتی معنا دارد که ***ساختمانی یکسان*** داشته باشند؛ یعنی هم تعداد سطرها و هم تعداد ستون‌ها در جمله‌های جمع، یکی باشد. درنتیجه عمل، ماتریسی با همان ساختمان به‌دست می‌آید. عمل ضرب، وقتی معنا دارد که تعداد ستون‌های ماتریس اول، برابرِ تعداد ستون‌های ماتریس دوم باشد. درنتیجه، ماتریسی به‌دست می‌آید که تعداد سطرهای آن برابر است با تعداد سطرهای عامل اول و تعداد ستون‌های آن، برابر است با تعداد ستون‌های عامل دوم ضرب.

عمل‌های روی ماتریس‌های مربعی، در بیشتر حالت‌ها، از همان قانون‌هایی پیروی می‌کنند، که دربارهٔ عملِ روی عددها صادق است؛ ولی برخی از قانون‌ها به‌هم می‌خورد.

ویژگی‌های اصلی عمل‌ها را، روی ماتریس‌ها، می‌آوریم:

۱. $A+B=B+A$ (قانون جابه‌جایی برای جمع).

۲. $(A+B)+C=A+(B+C)$ (قانون شرکت‌پذیری برای جمع).

۳. $c(A+B)=cA+cB$

۳′. $(c_1+c_2)A=c_1A+c_2A$

(قانون پخشی برای ضرب در عددها. در این‌جا c، c_1 و c_2 عددند نه ماتریس).

۴. $(c_1c_2)A=c_1(c_2A)$ (قانون شرکت‌پذیری ضربِ در عدد).

۵. ماتریس صفر $\begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}=O$ وجود دارد، به‌نحوی که برای هر ماتریس A داریم: $A+O=A$.

۶. $c\cdot O=O\cdot A=O$؛ برعکس، اگر $cA=O$ یا $c=0$ یا $A=O$ (در این‌جا، c، عدد است).

۷. برای هر ماتریس A، ماتریس متقابل $-A$ وجود دارد، به‌نحوی که $A+(-A)=O$.

۸. $(A+B)\cdot C=AC+BC$

۸′. $C(A+B)=CA+CB$

(قانون‌های پخشی، برای جمع‌وضرب ماتریس‌ها.)

۹. $(AB)C=A(BC)$ (قانون شرکت‌پذیری در ضرب).

۱۰. $(cA)B=A(cB)=c(AB)$.

این ویژگی‌ها، نه‌تنها دربارهٔ ماتریس‌های مربعی، بلکه برای هر ماتریس مستطیلی صادق‌اند، تنها با این شرط که انجام عمل‌ها ممکن باشد؛ و طبیعی است، دربارهٔ ماتریس‌های مربعیِ هم‌مرتبه، نیازی به‌این شرط نیست.

ویژگی‌هایی که آوردیم، شبیه ویژگی‌های عمل روی عددهاست.

اکنون به شرح دو عمل، که خاص ماتریس‌هاست، می‌پردازیم.

اول، در ضرب دو ماتریس، ولو در حالت مربعی، ممکن است قانون جابه‌جایی صدق نکند، یعنی AB همیشه برابر BA نیست. مثال:

$$\begin{pmatrix} ۳ & -۲ \\ -۱ & ۴ \end{pmatrix} \begin{pmatrix} ۱ & ۲ \\ ۳ & ۲ \end{pmatrix} = \begin{pmatrix} -۳ & ۲ \\ ۱۱ & ۶ \end{pmatrix} ,$$
$$\begin{pmatrix} ۱ & ۲ \\ ۳ & ۲ \end{pmatrix} \begin{pmatrix} ۳ & -۲ \\ -۱ & ۴ \end{pmatrix} = \begin{pmatrix} ۱ & ۶ \\ ۷ & ۲ \end{pmatrix}$$

دوم، می‌دانیم حاصل‌ضرب دو عدد، وقتی و تنها وقتی، برابر صفر می‌شود که، دست‌کم، یکی از عامل‌های ضرب برابر صفر باشد. و می‌دانیم، همین قضیه، مبنای نظریهٔ معادله‌های جبری است. این قضیه دربارهٔ ضرب ماتریس‌ها درست نیست؛ یعنی حاصل‌ضرب دو ماتریس می‌تواند برابر صفر شود، بی‌آنکه، ولو یکی از دو عامل ضرب، ماتریس صفر باشد:

$$\begin{pmatrix} ۱ & ۱ \\ ۱ & ۱ \end{pmatrix} \begin{pmatrix} ۱ & ۱ \\ -۱ & -۱ \end{pmatrix} = \begin{pmatrix} ۰ & ۰ \\ ۰ & ۰ \end{pmatrix}$$

ویژگی دیگری را هم، در ضرب ماتریس‌ها، یادآوری می‌کنیم. ماتریس $\bar{A}$ را ***ترانهادِ*** (ترانسپزه = جابه‌جاشدهٔ) ماتریس A گویند، وقتی در هر سطر $\bar{A}$، درایه‌های ستون متناظر A، با حفظ ترتیب آن، قرار گرفته باشد. از جمله برای ماتریس

$$A = \begin{pmatrix} ۱ & ۲ \\ ۳ & ۴ \\ ۵ & ۶ \end{pmatrix}$$

ماتریس ترانهاد چنین است:

$$\bar{A} = \begin{pmatrix} ۱ & ۳ & ۵ \\ ۲ & ۴ & ۶ \end{pmatrix}$$

عمل ضرب ماتریس‌ها با ماتریس ترانهاد با رابطهٔ

$$\overline{AB} = \bar{B}\,\bar{A}$$

مربوط می‌شود. این اتحاد را، با استفاده از قاعدهٔ ضرب ماتریس‌ها، می‌توان به سادگی ثابت کرد.

نظریهٔ ماتریس، بخش مهم و جدانشدنیِ جبر خطی را تشکیل می‌دهد و نقش ابزاری را دارد که، به‌یاری آن، می‌توان، مسأله‌های جبر خطی را عنوان و حل کرد.

قیاس‌ها و شباهت‌های هندسی در جبر خطی. به‌جز شرح تاریخی که دربارهٔ اندیشه‌ها و مسأله‌های جبر خطی آوردیم، نیاز آنالیز ریاضی و هندسه، و به‌ویژه هندسهٔ تحلیلی هم، ضرورت تکامل جبر خطی را ایجاب می‌کرد و درضمن، این شاخه‌های ریاضی، به‌نوبهٔ خود، موجب غنای اندیشه‌ها و مسأله‌های جبر خطی شد. روشن است، هندسهٔ تحلیلی در صفحه و تا حد زیادی در فضا، در بخشی که به نظریهٔ خط‌های راست و صفحه‌ها مربوط می‌شود، به‌صورت ساده‌ای، از امکان‌های جدید جبر خطی استفاده می‌کند. درواقع، خط راستِ واقع بر صفحه، به‌وسیلهٔ معادله‌ای خطی که شامل دو متغیر است داده می‌شود که معرف رابطهٔ بین دو مختص هر نقطه از خط راست است. در فضا، صفحه را به‌وسیلهٔ معادله‌ای خطی با سه متغیر (مختصات نقطهٔ دل‌خواهی از صفحه)، و خط راست را، به‌یاری دو معادلهٔ خطی نشان می‌دهند.

روشن است، اگر از مفهوم بُردار استفاده کنیم، در هندسهٔ تحلیلی و، بنابراین، در نظریهٔ ساده‌ترین نوع دستگاه‌های معادله‌های خطی؛ به سادگی و روشنی بیشتری دست می‌یابیم. به همین ترتیب، اگر در جبر خطی و به‌ویژه، در نظریهٔ عمومی دستگاه معادله‌های خطی، از مفهوم بُردار، به‌معنای نوعی تعمیم آن، استفاده کنیم، سادگی و روشنی بیشتری به‌دست می‌آید. راه این تعمیم، چنین است. بُردار (در فضا) به‌وسیلهٔ سه عدد داده می‌شود: سه تصویر آن بر محورهای مختصات. هر سه عدد حقیقی، به‌نوبهٔ خود، می‌تواند به‌صورت هندسی، یک بردار را مشخص کند(در فضا).

برای بُردارها، عمل جمع («بنا بر قاعدهٔ متوازی‌الاضلاع») و عمل ضرب در عدد، معلوم است. این عمل‌ها، شبیه عمل‌هایی است که روی نیروها، سرعت‌ها، شتاب‌ها و کمیت‌های فیزیکی دیگری که با بُردار مشخص می‌شوند، به‌انجام می‌رسد.

اگر بُردارها، به‌وسیلهٔ مختصات خود (یعنی تصویرهای روی محورهای مختصات) داده شده باشند، آن‌وقت عمل جمع و عمل ضرب در یک عدد (که باید دربارهٔ بُردارها انجام شود)، متناظر با همان عمل‌ها روی سطرها (یا ستون‌ها) از مختصات آن‌هاست.

به این ترتیب، سطر یا ستونی را که از سه عضو تشکیل شده باشد، به‌سادگی می‌توان، از نظر هندسی، به‌عنوان بُرداری در فضای سه‌بعدی تعبیر کرد؛ درضمن، عمل‌های اصلی روی

«سطرها» (یا «ستون‌ها»)، به‌صورت عمل‌های نظیر آن‌ها، روی بُردارهای فضایی تعبیر می‌شود؛ به‌نحوی که جبر سطرها یا ستون‌های شامل سه عضو، هیچ تفاوتی با جبر بُرداری در فضای سه‌بعدی ندارد. به این خاطر طبیعی است که اصطلاح‌های هندسی را در جبر خطی به‌کار بگیریم.

ستون یا سطر شامل n عدد، $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$، به‌عنوان یک «بُردار» در نظر گرفته می‌شود، یعنی به‌عنوان عنصری از یک «فضای بُرداری n بعدی». مجموع دو بُردارِ $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ و $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$، بُردار $\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \cdots \\ x_n+y_n \end{pmatrix}$ و حاصل‌ضرب بُردار $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ در عدد c، بُردار $\begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}$ به‌حساب می‌آید.

مجموعهٔ همهٔ بُردارها (ستون‌ها)، بنا به تعریف، فضای بُرداری حسابیِ nبُعدی نامیده می‌شود.

همراه با فضای بُرداری حسابی nبُعدی، می‌توان مفهوم فضای نقطه‌ای nبُعدی را هم وارد و هر ستونِ شامل n عدد حقیقی را، با یک نقطهٔ هندسی متناظر کرد. در این صورت، فضای بُرداری nبُعدی را، می‌توان این‌طور تعریف کرد: هر دو نقطهٔ A و B را، نظیر بُردار $\overrightarrow{AB}$ می‌گیریم که از A به B می‌رود و مختصات آن (تصویرهای روی محورهای مختصات)، بنابر تعریف، برابر است با تفاضل مختصات دو نقطهٔ متناظر B و A. دو بُردار را وقتی برابر می‌گیریم که، مختصات متناظر آن‌ها، برابر باشند، شبیه حالت فضای سه‌بُعدی که در آن‌جا، دو بردار را وقتی برابر می‌دانیم که بتوان یکی از آن‌ها را، با انتقال دیگری موازی با خود، به‌دست آورد.

بین بردارهای فضای بُرداری nبُعدی و نقطه‌های فضای نقطه‌ای nبُعدی، تناظر یک‌به‌یک برقرار است.

نقطهٔ $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ را مبدأ مختصات و هر نقطهٔ دیگر را، متناظر با برداری می‌گیرند که مبداء مختصات را به این نقطه وصل کرده است. در این صورت، هر بُردار متناظر با نقطهٔ انتهای

این بردار می‌شود، با این فرض که آغاز بُردار را، مبداء مختصات بگیریم. وارد کردن فضای نقطه‌ای، شباهت‌ها و قیاس‌های تازه‌ای را موجب می‌شود که به‌یاری آن‌ها می‌توان فضای برداری را «بهتر دید».

با این‌همه، در تعمیم‌ها، تعریف دقیق فضای نقطه‌ای، پیچیدگی‌های بیشتری پیدا می‌کند و، به همین مناسبت، در این‌جا از این مفهوم استفاده نمی‌کنیم. خواننده‌ای که می‌خواهد از شباهت‌های ناشی از بررسی فضای نقطه‌ای استفاده کند، باید عنصرهای فضای برداری را، همچون بردارهایی در نظر بگیرد که از مبداء مختصات آغاز شده‌اند.

وارد کردن اصطلاح‌های هندسی، به ما امکان می‌دهد تا در جبر خطی، از شباهت‌هایی استفاده کنیم که براساس معرفت شهودی هندسی و از راه بررسی هندسهٔ فضای سه‌بُعدی به‌دست می‌آیند. البته، در استفاده از این شباهت‌ها، باید محتاط بود. متوجه باشیم، برای هریک از «استدلال‌های» عینی ـ هندسی، باید امکان تحقیق دقیق منطقی با استفاده از تعریف نقطه‌ای مفهوم‌های «هندسی» وجود داشته باشد و بتوان قضیه‌ها را بادقت ثابت کرد.

ویژگی مشخص عنصرهای فضای برداریِ nبعدی، در وجود عمل‌های جمع و ضرب در عدد است، که عمل‌های مربوط به عددها را به‌یاد می‌آورد. همان‌طور که دربارهٔ عملِ روی ماتریس‌ها گفتیم، برای عمل جمع، قانون‌های جابه‌جایی و شرکت‌پذیری صدق می‌کنند، قانون‌های پخشی برای ضرب در عدد درست است، عمل جمع به‌صورت یک‌ارزشی برگشت‌پذیر است، حاصل‌ضرب بردار در عدد تنها وقتی برابر صفر می‌شود که یا بردار برابر با بردار صفر باشد ویا عدد برابر صفر شود.

وجود این ویژگی‌ها، تنها برای ماتریس‌های ستونی (یا سطری) نیست. مجموعه‌ای از ماتریس‌ها که ساختمانی یکسان داشته باشند، همچنین کمیت‌های برداری فیزیکی (نیرو، سرعت، شتاب وغیره) هم، همین ویژگی‌ها را دارند. برخی از موضوع‌های ریاضی دیگری هم، که سرشتی به‌کلی متفاوت دارند، دارای این ویژگی‌ها هستند، مثل مجموعهٔ همهٔ چندجمله‌ای‌های با یک متغیر، مجموعهٔ همهٔ تابع‌های پیوسته که روی پاره‌خط راستی مثل $[a, b]$ داده شده باشند، مجموعهٔ همهٔ جواب‌های معادلهٔ دیفرانسیلی خطی همگن وغیره.

این موقعیت، برای تعمیم بعدی فضای برداری و به‌ویژه، وارد کردن فضاهای خطی عمومی سودمند است. عنصرهای یک‌چنین فضای تعمیم‌یافته‌ای، می‌توانند هرگونه موضوع‌های ریاضی یا فیزیکی باشد که تعریف عمل جمع یا عمل ضربِ در یک عدد، برای آن‌ها به نحوی طبیعی تعریف شده باشد. این حرکت به‌سوی یک فضای خطی عمومی و

به‌کلی انتزاعی، همان‌طور که خواهیم دید، هیچ دشواری خاصی در نظریه پدید نمی‌آورد. هر فضای خطی (و البته nبعدی، که مفهوم آن را در بند بعد روشن می‌کنیم)، چه از نظر ساختار و چه از نظر ویژگی‌های خود، تفاوتی با فضای خطی حسابی ندارد، ولی زمینهٔ کاربرد آن بی‌اندازه گسترده‌تر است و امکان استفاده از روش‌های جبر خطی را در پهنهٔ بسیار گسترده‌ای از مسأله‌های نظریِ دانش‌های طبیعی فراهم می‌آورد.

۲. فضای خطی

تعریف فضای خطی. تعریفِ دقیقِ فضای خطی را می‌آوریم.

فضای خطی، به مجموعه‌ای از موضوع‌ها (با طبیعتی دل‌خواه) گفته می‌شود که، برای آن‌ها، مفهوم‌های جمع و ضربِ در عدد معنا داشته باشند، به‌نحوی که با شرط‌های زیر سازگار باشند.

۱. $(X+Y)+Z=X+(Y+Z)$؛

۲. عنصر «خنثای» 0 وجود داشته باشد، به‌نحوی که برای هر X داشته باشیم: $X+0=X$؛

۳. برای هر عنصر X، عنصر متقابل $-X$ وجود داشته باشد، به‌نحوی که داشته باشیم: $X+(-X)=0$؛

۴. $X+Y=Y+X$؛

۵. $1\cdot X=X$؛

۶. $c_1(c_2X)=c_1c_2X$؛

۷. $(c_1+c_2)X=c_1X+c_2X$؛

۸. $c(X+Y)=cX+cY$.

در این‌جا، X، Y و Z عنصرهایی از فضای خطی، و 1، c_1، c_2 و c عدد هستند.

این ویژگی‌ها (که آن‌ها را ***اصل موضوع‌های فضای خطی*** هم می‌گویند) را باید طبیعی دانست، زیرا چیزی جز بیان صوری ویژگی‌هایی از عمل جمع و عمل ضربِ در عدد نیستند، که به‌صورتی جدانشدنی با مفهوم این عمل‌ها (به هر صورت عامی که در نظر گرفته شوند) بستگی دارند. عمل‌هایی که مفهوم فیزیکی دارند، تنها وقتی به‌عنوان جمع و ضربِ در عدد

وارد میدان می‌شوند که، این عمل‌ها، با ویژگی‌های ۱ تا ۸ سازگار باشند.

برخی نتیجه‌های ناشی از این اصل موضوع‌ها را می‌آوریم:

a) عنصر خنثای ۰ ، منحصربه‌فرد است، یعنی تنها یک عنصر وجود دارد که در اصل موضوع ۲ صدق می‌کند؛

b) برای عنصر مفروض X، تنها یک عنصر متقابل X وجود دارد؛

c) «عمل تفریق» معنا دارد، یعنی با در دست داشتن مجموع و یکی از جمله‌های جمع، همیشه می‌توان جملهٔ دوم جمع را پیدا کرد؛ درضمن جملهٔ دوم منحصربه‌فرد است، اگر داشته باشیم: $X+Z=Y$، آن‌وقت $Z=Y+(-X)$؛

d) $0\cdot X=c\cdot 0=0$ ؛

e) اگر داشته باشیم $cX=0$، آن‌وقت یا $c=0$ یا $X=0$؛

f) $X=(-1)X$.

اثبات این نتیجه‌ها بسیار ساده است و ما به آن نمی‌پردازیم. از این به‌بعد، عنصرهای فضای خطی را بردار می‌نامیم.

وابستگی و استقلال خطی بردارها. اکنون به مفهوم وابستگی و استقلال خطی بردارها می‌پردازیم.

ترکیب خطی بردارهای X_1، X_2، ... ، X_m، به بردار

$$c_1X_1 + c_2X_2 + \cdots + c_mX_m$$

گفته می‌شود (به‌ازای مقدارهای عددیِ ضریب‌های c_1، c_2، ... ، c_m). اگر بین بردارهای X_1، X_2، ... ، X_m، دست‌کم یکی پیدا شود که ترکیب خطی بقیه باشد؛ در این صورت گویند، بردارهای X_1، X_2، ... ، X_m، وابسته خطی هستند. ولی اگر هیچ‌کدام از بردارهای X_1، X_2، ... ، X_m را نتوان به‌صورت ترکیبی خطی نسبت به دیگران نوشت، آن‌وقت، این بردارها را مستقلِ خطی گویند.

به‌سادگی دیده می‌شود، برای استقلال خطی بردارهای X_1، X_2، ... ، X_m، لازم و کافی است، رابطهٔ

$$c_1X_1 + c_2X_2 + \cdots + c_mX_m = 0$$

تنها برای $c_1=c_2=\cdots=c_m=0$ برقرار باشد.

برای بردارهایی که در فضای عادیِ سه‌بعدی باشند، مفهوم‌های وابستگی خطی و استقلال خطی، معنای هندسی ساده‌ای دارد.

دو بردار X_1 و X_2 را در نظر می‌گیریم. وابستگی خطی این دو بردار به این معنی است که یکی از این بردارها، «ترکیب خطی» دیگری است، یعنی خیلی ساده، تنها در یک ضریب عددی با آن فرق دارد. این وضع، به‌معنای آن است که دو بردار روی یک خط راست قرار دارند (در یک‌جهت یا در دو جهت مخالف).

برعکس، اگر دو بردار روی یک خط راست واقع باشند، آن‌وقت، یک وابستگی خطی بین آن‌ها برقرار است. بنابراین، استقلال خطی دو بردار، به‌معنای آن است که نمی‌توان آن‌ها را روی یک خط راست جا داد: دو بردار دارای دو امتداد مختلف‌اند.

اکنون ببینیم، وابستگی خطی و استقلال خطی سه بردار چه معنایی دارد! فرض کنیم، بردارهای X_1، X_2 و X_3، وابسته خطی باشند. برای مشخص بودن وضع، بردار X_3 را ترکیبی خطی از بردارهای X_1 و X_2 می‌گیریم. روشن است، در این حالت، بردار X_3، روی صفحه‌ای قرار دارد که از بردارهای X_1 و X_2 گذشته است؛ یعنی هر سه بردار روی یک صفحه‌اند. همچنین به‌سادگی دیده می‌شود، اگر سه بردار روی یک صفحه باشند، بین آن‌ها یک وابستگی خطی برقرار است. درواقع، اگر بردارهای X_1 و X_2 بر یک خط راست نباشند، آن‌وقت X_3 را می‌توان برحسب X_1 و X_2 تجزیه کرد، یعنی X_3 را به‌صورت ترکیبی خطی از X_1 و X_2 نوشت. اگر هم X_1 و X_2 روی یک خط راست باشند، آن‌وقت، X_1 و X_2 وابسته خطی هستند.

به این ترتیب، وابستگی خطی سه بردار، هم‌ارز با آن است که، این سه بردار، روی یک صفحه قرار گیرند. بنابراین، استقلال خطی سه بردار، وقتی و تنها وقتی تأمین می‌شود که روی یک صفحه نباشند.

چهار بردار در فضای سه‌بعدی، همیشه وابسته خطی‌اند. درواقع، اگر بردارهای X_1، X_2 و X_3 وابسته خطی باشند، آن‌وقت به‌ازای هر بردار X_4، بردارهای X_1، X_2، X_3 و X_4 هم، وابسته خطی هستند. و اگر X_1، X_2 و X_3مستقل خطی باشند، به‌معنای آن است که روی یک صفحه نیستند و هر بردار دل‌خواه X_4 را می‌توان برحسب X_1، X_2 و X_3 تجزیه و آن را به‌صورت ترکیبی از این سه بردار نشان داد.

پس در فضای سه‌بعدی، بردارهای X_1، X_2، ... ، X_k $(k \geq 3)$ ، وقتی و تنها وقتی وابسته‌خطی هستند که در فضایی با تعداد بُعدهای کمتر از k واقع باشند.

بعد از آن‌که تعریف دقیق «زیرفضا» و مفهوم «بُعد» را بیاوریم، خواهیم دید در حالت کلی هم، بستگی خطی بردارهای X_1، X_2، ... ، X_k، هم‌ارز با آن است که این بردارها در فضایی با تعداد بُعدهای کمتر از k واقع باشند، یعنی مفهوم «هندسی» بستگی خطی، به همان مفهومی که برای بردارهای فضای سه‌بعدی وجود داشت، باقی می‌ماند.

قضیه‌ای را که در این‌جا می‌آوریم، در نظریهٔ فضاهای خطی، نقش عمده‌ای دارد. اگر بردارهای X_1، X_2، ... ، X_m، به‌صورت ترکیب‌هایی خطی از بردارهای Y_1، Y_2، ... ، Y_k باشند و در ضمن $m > k$، آن‌وقت خودِ بردارهای X_1، X_2، ... ، X_m، وابسته خطی‌اند (قضیهٔ مربوط به بستگی خطیِ ترکیب‌های خطی).

قضیه برای $k = 1$ روشن است. برای $k > 1$ هم می‌توان قضیه را به‌سادگی و به‌یاری روش استقرای ریاضی ثابت کرد.

پایه‌ها و بُعد فضا. در فضای سه‌بعدی، هر سه بردار X_1، X_2 و X_3 را که بر یک صفحه واقع نباشند (یعنی، دارای استقلال خطی باشند)، می‌توان **پایهٔ** این فضا دانست؛ و این به‌معنای آن است که هر بردار فضای سه‌بعدی را می‌توان برحسب X_1، X_2 و X_3 تجزیه کرد، یعنی آن را به‌صورت ترکیبی خطی از این سه بردار نوشت.

فضاهای برداری خطی را می‌توان به دو گونه تقسیم کرد.

ممکن است، بردارهای مستقل خطی، به هر تعداد دل‌خواه (و به هر اندازه زیاد) در فضا وجود داشته باشد. چنین فضاهایی را بی‌نهایت‌بُعدی می‌نامند و بررسی آن‌ها، از چارچوب جبر خطی بیرون است. این فضاها، موضوع خاص یک نظام ریاضی را تشکیل می‌دهند که آنالیز تابعی (یا آنالیز فونکسیونل) نام دارد (فصل نوزدهم را ببینید).

فضای خطی را، «با بُعد محدود» گویند، وقتی که در آن، مرز محدودی برای تعداد بردارهای مستقل خطی وجود داشته باشد، یعنی بتوان عددی مثل n پیدا کرد، به‌نحوی که در فضا، n بردار مستقل خطی وجود داشته باشد؛ ولی اگر تعداد بردارها از n بیشتر باشد، با بستگی خطی به هم مربوط باشند. عدد n را «بُعد فضا» گویند.

برای نمونه، فضای بردارهای هندسی سه‌بُعدی معمولی، به این تعریف کلی، دارای بُعد سه است. درواقع، در فضای هندسی سه‌بُعدی، به تعداد دل‌خواه، بردارهای سه‌گانهٔ مستقل خطی وجود دارد، ولی هر چهار بردار دل خواه در آن، وابسته خطی هستند.

فضای ماتریس‌های ستونی با n درایه، به مفهوم تعریفی که در این‌جا آوردیم، یک فضای

n بُعدی است. درواقع، در این فضا، n بردار مستقل خطی وجود دارد، مثل بردارهای

$$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\quad,\qquad e_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}\quad,\ \ldots\ ,\quad e_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} \qquad (2)$$

ولی هر بردار $\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ از این فضا، به‌صورت ترکیبی خطی از آن‌هاست، از جمله به‌صورت

$$x_1e_1+x_2e_2+\cdots+x_ne_n$$

بنابراین، بنا بر قضیهٔ مربوط به رابطهٔ خطیِ ترکیب‌های خطی، اگر تعداد بردارها از n تجاوز کند، با بستگی خطی به هم مربوط‌اند.

چندجمله‌ای‌های با یک متغیر، یک فضای خطی را تشکیل می‌دهند. درواقع، برای چندجمله‌ای‌ها، تعریف طبیعی عمل جمع و عمل ضرب در عدد، با اصل موضوع‌های ۱ تا ۸ سازگار است. ولی این فضا بی‌نهایت‌بُعدی است، زیرا بردارهای ۱ ، x، ... ، x^N، برای هر N مستقل خطی هستند. البته، مجموعهٔ چندجمله‌ای‌هایی که درجهٔ آن‌ها از عدد مفروض N تجاوز نمی‌کند، فضایی با بُعد محدود را تشکیل می‌دهند که بُعد آن برابر است با $N+1$. درواقع، بردارهای ۱ ، x، ... ، x^N، مستقل خطی هستند و تعداد آن‌ها برابر است با $N+1$. ولی هر چندجمله‌ای دل‌خواه، که درجهٔ آن از N تجاوز نکند، ترکیبی خطی از ۱ ، x، ... ، x^N است. به این ترتیب، بنا بر همان قضیهٔ مربوط به رابطهٔ خطی، هر مجموعهٔ چندجمله‌ای‌های دل‌خواه، که درجهٔ آن‌ها کوچکتر یا برابر N، ولی تعداد آن‌ها بیشتر از $N+1$ باشد، وابسته خطی هستند.

اکنون، مفهوم **پایه** را برای فضای nبُعدی روشن می‌کنیم. پایه، به مجموعه‌ای از بردارهای مستقل خطی فضا گفته می‌شود که بتوان هر بردار دل‌خواه فضا را به‌صورت ترکیب خطی بردارهای این مجموعه درآورد. از جمله در فضای ماتریس‌های ستونی، برای نمونه، پایه عبارت است از مجموعهٔ بردارهای (۲). در فضای چندجمله‌ای‌های با درجهٔ کوچکتر یا برابر N، می‌توان «بردارهای» ۱ ، x، ... ، x^N را به‌عنوان پایه پذیرفت. در فضای هندسی سه‌بعدی، نقش پایه، به‌عهدهٔ هر سه بردار مستقل خطی است.

در فضای خطی nبُعدی، هر مجموعه‌ای از n بردار مستقل خطی (وجود دست‌کم یکی

از این مجموعه‌ها، اساس تعریف فضای nبُعدی است)، پایهٔ این فضا را تشکیل می‌دهد. درواقع، فرض کنید e_1، e_2، ... ، e_n، بردارهای مستقل خطی فضای خطی nبُعدی، و X بردار دل‌خواهی از این فضا باشد. در این صورت، بردارهای X، e_1، ... ، e_n وابستگی خطی دارند (زیرا تعداد آن‌ها از n بیشتر است)، یعنی عددهای c، c_1، c_2، ... ، c_n پیدا می‌شود، که همه با هم و به‌طور هم‌زمان برابر صفر نیستند، و داشته باشیم:

$$cX + c_1 e_1 + c_2 e_2 + \cdots + c_n e_n = 0$$

درضمن $c \neq 0$، زیرا برای $c = 0$، بردارهای e_1، e_2، ... ، e_n بستگی خطی پیدا می‌کنند. بنابراین

$$X = -\frac{c_1}{c} e_1 - \frac{c_2}{c} e_2 - \cdots - \frac{c_n}{c} e_n$$

یعنی هر بردار این فضا، ترکیبی خطی از بردارهای e_1، e_2، ... ، e_n است.

هر پایهٔ فضای خطی nبُعدی، درست شامل n بردار است. درواقع بردارهای پایه، استقلال خطی دارند و، به همین دلیل، تعداد آن‌ها نمی‌تواند از n بیشتر شود. از طرف دیگر فرض کنید e_1، e_2، ... ، e_k، پایهٔ یک فضای nبُعدی را تشکیل دهند. دیدیم $k \leq n$. بنابر تعریف پایه، هر برداری از فضا، ترکیبی خطی از بردارهای e_1، e_2، ... ، e_k است و بنابر قضیهٔ وابستگی خطی ترکیب‌های خطی، اگر تعداد بردارها را بزرگتر از k بگیریم، باید در وابستگی خطی باشند. از این‌جا نتیجه می‌شود، بُعد فضا، یعنی n، بیشتر از تعداد بردارهای پایه، یعنی k، نیست. بنابراین $k = n$.

اکنون تلاش می‌کنیم، مختصات بردار مفروض را، برحسب پایهٔ e_1، e_2، ... ، e_n بنویسیم. همان‌طور که گفتیم، هر بردار X، ترکیبی خطی از بردارهای پایه است؛ درضمن، این روش نمایش بردار منحصربه‌فرد است. درواقع، اگر فرض کنیم بردار X با دو روش برحسب بردارهای پایه نوشته شده باشد:

$$X = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n ,$$
$$X = x'_1 e_1 + x'_2 e_2 + \cdots + x'_n e_n ,$$

در این صورت، باید داشته باشیم:

$$(x_1 - x'_1)\, e_1 + (x_2 - x'_2)\, e_2 + \cdots + (x_n - x'_n)\, e_n = 0$$

و از این‌جا، بنا بر استقلالِ خطی بردارهای e_1، e_2، ... ، e_n نتیجه می‌شود:

$$x_1 = x'_1 \ \ , \ \ x_2 = x'_2 \ \ , \ \ x_n = x'_n$$

ضریب‌های x_1، x_2، ... ، x_n را که در تجزیهٔ بردار X، برحسب بردارهای پایه به‌دست می‌آید، **مختصات** (یا **مؤلفه‌های**) بردار X در این پایه گویند. به این ترتیب، ***تنها اگر پایهٔ فضا انتخاب شده باشد،*** هر بردار به‌طور طبیعی با یک سطر (یا یک ستون)، یعنی یک ماتریس سطری (یا ستونی) از مختصات خود متناظر می‌شود؛ برعکس هر سطر (یا ستون) از n عدد را می‌توان به‌عنوان مختصات یک بردار در نظر گرفت.

عمل جمع بردارها و عمل ضرب بردار در یک عدد، متناظر است با همین عمل‌ها روی سطرهای مختصات (یا ستون‌های آن).

به این ترتیب، هر فضای خطی nبُعدی، بدون توجه به طبیعت عنصرهای آن (تابع‌ها، ماتریس‌ها، نوعی از کمیت‌های فیزیکی وغیره)، در رابطه با این عمل‌ها، هیچ تفاوتی با فضای سطرها (یا ستون‌ها) ندارد. بنابراین، همان‌طور که پیش از این هم گفته‌ایم، تعمیم مفهوم فضای خطی و اصل موضوعی کردن آن، هیچ‌گونه پیچیدگی، در مقایسه با تفسیر فضا به‌عنوان فضای سطرها، پدید نمی‌آورد، ولی دایرهٔ کاربرد این مفهوم را، بی‌اندازه گسترش می‌دهد.

در ریاضیات، یکی بودن ویژگی‌های دومجموعه‌ای را، که هرکدام شامل عضوهای خاص خود است، نسبت به دستگاه مفروضی از عمل‌ها (یا رابطهٔ دیگری بین عضوها)، **یکریختی** (ایزومُورفیسم) گویند. تعریف دقیق یکریختی دستگاه‌های جبری را ما در بخش بیستم خواهیم دید. با استفاده از این اصطلاح می‌توانیم بگوییم، همهٔ فضاهای خطی n بُعدی، بدون ارتباط با طبیعت عنصرهای تشکیل‌دهندهٔ آن‌ها، نسبت به هم یکریخت (ایزومُورف) هستند. درضمن، همهٔ آن‌ها، با فضای سطرها، یکریخت‌اند.

زیرفضا. مجموعه‌ای از بردارهای فضایِ خطی nبُعدیِ R_n، به‌شرطی که هر ترکیب خطی از بردارهای دل‌خواه این مجموعه، به خودمجموعه تعلق داشته باشد، ***زیرفضای*** R_n نامیده می‌شود. روشن است زیرفضای R_n، خود یک فضای خطی است و بنابراین دارای پایه و بُعد است. همچنین روشن است، تعداد بُعدهای زیرفضا، از تعداد بُعدهای تمامی فضا تجاوز

نمی‌کند و تنها در حالتی می‌تواند برابر تعداد بُعدهای فضای اصلی باشد که زیرفضا بر تمامی فضا منطبق شود.

برای فضای برداری سه‌بُعدی، می‌توان صفحه و خط راست را، با تقریب انتقال، زیرفضا به‌حساب آورد؛ به‌زبان دقیق‌تر، مجموعهٔ همهٔ بردارهای واقع بر یک صفحه یا بر یک خط راست، با این شرط که صفحه و یا خط راست از مبداء مختصات بگذرند، زیرفضاهایی از فضای برداری سه‌بُعدی هستند.

اغلب بررسی زیرفضاهایی پیش می‌آید که از دستگاهی از بردارها «پدید آمده‌اند». این‌گونه زیرفضاها به این‌گونه تعریف می‌شوند. فرض کنید، دستگاهی از بردارهای X_1، X_2، ... ، X_m (وابسته خطی یا مستقل خطی) در فضای R_n داده شده باشد. در این صورت، مجموعهٔ همهٔ ترکیب‌های خطی این بردارها، یعنی

$$\{c_1X_1 + c_2X_2 + \cdots + c_mX_m\}$$

زیرفضایی از فضای R_n را می‌سازد و به آن، زیرفضایی گویند که از بردارهای X_1، X_2، ... ، X_m پدید آمده است.

بُعد این فضا را، **رتبهٔ** دستگاه بردارهای X_1، X_2، ... ، X_m می‌نامند. به‌سادگی دیده می‌شود، رتبهٔ دستگاه بردارها، برابر است با حداکثر تعداد بردارهای مستقل خطی، که در این دستگاه وجود دارد.

«مجموعه‌ای» که تنها شامل بردار خنثا (بردار صفر) باشد، با همهٔ شرط‌های زیرفضا سازگار است. تعداد بُعدهای این زیرفضا را، برابر صفر می‌گیریم.

اگر دو زیرفضا از فضای R_n در دست باشد، می‌توان از آن‌ها، دو زیرفضای دیگر هم ساخت: مجموع برداری و اشتراک آن‌ها.

مجموع برداریِ دو زیرفضای P و Q به مجموعهٔ مجموع‌های بردارهایی گفته می‌شود که به زیرفضاهای P و Q تعلق دارند. مجموع برداری را هم می‌توان به‌عنوان زیرفضایی در نظر گرفت که از اجتماع پایه‌های دو زیرفضای P و Q پدید آمده است.

اشتراک دو زیرفضا، به مجموعهٔ همهٔ بردارهایی گفته می‌شود که به هر دو زیرفضا تعلق دارند. مجموع برداری دو صفحه (یعنی زیرفضاهای برداری دوبُعدی) در فضای سه‌بُعدی معمولی، عبارت است از همهٔ فضا (تنها به‌شرطی که دو صفحه منطبق بر هم نباشند)؛ و اشتراک آن‌ها، عبارت است از یک خط راست (با همان شرط).

اگر بُعد دو زیرفضای مفروض را p و q، و بُعد مجموع برداری و اشتراک آن‌ها را به‌ترتیب t و s بگیریم، رابطهٔ جالبی به‌دست می‌آید:

$$p + q = t + s$$

که در این‌جا، از اثبات آن می‌گذریم.

از این رابطه نتیجه‌هایی به‌دست می‌آید که به اشتراک زیرفضاها، در حالت‌های خاص، مربوط می‌شود. از جمله، دو صفحه‌ای که منطبق بر هم نباشند (یعنی، دو زیرفضای دوبعدی) در فضای چهاربعدی، در حالت کلی، تنها در یک نقطه مشترک‌اند (بُعد اشتراک آن‌ها برابر صفر است)؛ دو صفحه تنها وقتی یکدیگر را در یک خط راست قطع می‌کنند که مجموع برداری آن‌ها سه‌بُعدی باشد؛ یعنی، به‌شرطی که هر دو صفحه در یک فضای سه‌بُعدی واقع باشند. درواقع، در این حالت داریم: $t + s = ۲ + ۲ = ۴$، و از این‌جا، تنها به‌ازای $t = ۳$ به‌دست می‌آید: $s = ۱$.

فضای خطی مختلط. ضمن شرح فضای ماتریس‌های سطری و فضای خطی کلی، در این باره دقت نکردیم که، وقتی از عمل ضرب بردار در یک عدد صحبت می‌شود، با چگونه عددی سر و کار داریم؟ از آن‌جا که کار خود را با تعمیم بردارهای معمولی آغاز کردیم، یعنی از پاره‌خط‌های راست جهت‌دار در فضای هندسی سه‌بعدی، به‌طور طبیعی با عددهای حقیقی سر و کار داشتیم. فضای خطی که به این ترتیب ساخته می‌شود، فضای خطی حقیقی نام دارد که طبیعی‌ترین نوع تعمیم فضای سه‌بعدیِ بردارهای معمولی است. ولی برای بسیاری از مسأله‌های ریاضیات امروزی، بررسی ***فضاهای خطی مختلط*** می‌تواند سودمند باشد. این اصطلاح به‌معنای مجموعه‌ای از موضوع‌هاست که دربارهٔ آن‌ها، عمل جمع و عمل ضرب در یک عدد مختلط، تعریف شده است و درضمن این عمل‌ها، با اصل موضوع‌های ۱ تا ۸ سازگارند. نمونهٔ فضای مختلط را، می‌توان فضای سطرهایی دانست که درایه‌های آن را عددهای مختلط تشکیل داده باشند.

نظریهٔ فضای مختلط، از نظر صوری، هیچ تفاوت اساسی با نظریهٔ فضای حقیقی ندارد.

با این‌همه، ولو برای فضای مختلط دوبُعدی هم، نمی‌توان تعبیر هندسی عینی پیدا کرد. درواقع، فضای مختلط nبُعدی را می‌توان به‌عنوان فضای حقیقی هم به‌حساب آورد، زیرا وقتی برای عنصرهای آن، عمل ضرب در هر عدد مختلط تعریف می‌شود، به‌خودی‌خود،

ضرب در عددهای حقیقی را هم در بر می‌گیرد. ولی تعداد بُعدهای فضای nبعدی مختلط، وقتی به‌عنوان فضای حقیقی در نظر گرفته شود، برابر با $۲n$ یعنی دوبرابر می‌شود. اگر e_1، e_2، ... ، e_n، پایهٔ فضای مختلط باشد، وقتی به‌عنوان فضای حقیقی در نظر گرفته شود، می‌توان پایهٔ آن را، در مثل این بردارها دانست:

$$e_1 \ , ie_1 \ , e_2 \ , ie_2 \ , \dots \ , e_n \ , ie_n \ (i=\sqrt{-1}\ )$$

بنابراین فضای دوبعدی مختلط را، می‌توان همچون یک فضای حقیقی تفسیر کرد، ولی تعداد بُعدهای این فضای حقیقی، برابر ۴ می‌شود.

سپس، اگر به‌جای عددهایی که «بردارهای» فضا در آن‌ها ضرب می‌شوند، مجموعهٔ دیگری از عددها، غیر از مجموعهٔ همهٔ عددهای حقیقی یا همهٔ عددهای مختلط، انتخاب کنیم، باز هم نظریهٔ فضای خطی از لحاظ صوری بی‌تغییر می‌ماند، تنها شرط این است که: مجموعهٔ انتخابی عددها باید چنان باشد که نتیجهٔ عمل‌های اصلی حساب ـ جمع، تفریق، ضرب و تقسیم ـ به همان مجموعه تعلق داشته باشد. مجموعهٔ عددهایی که با این شرط بسازند، ***هیأت عددی*** نامیده می‌شود (بحث مفصل‌تر دربارهٔ این مفهوم را در بخش بیستم ببینید). از جمله، هیأت عددهای گویا، نمونهٔ یک هیأت عددی است.

نظریهٔ فضاهای خطی روی هیأت دل‌خواه، در برخی از شاخه‌های جبر، که به نظریهٔ عددها نزدیک‌اند، کاربردهای زیادی دارد.

فضای اقلیدسی nبُعدی. پیش از این، صحبتی از تعمیم بعضی مفهوم‌های مهم فضای برداری عادی، مثل مفهوم طول بردار و یا زاویهٔ بین بردارها نداشتیم. می‌دانیم، این مفهوم‌ها، در هندسهٔ تحلیلی هم، در حالت‌هایی که به برخورد خط‌های راست و صفحه و یا توازی آن‌ها مربوط می‌شود، کاربرد دارد. ویژگی‌هایی از فضا را که شرح آن‌ها، نیازی به مفهوم‌های طول و زاویه نداشته باشند، می‌توان به‌عنوان ویژگی‌هایی مشخص کرد که با هر تبدیل آفین از بین نمی‌روند (جلد اول، بخش سوم، بند ۱۱ را ببینید). به همین دلیل، فضاهای خطی را، که در آن‌ها مفهوم طول بردار تعریف نشده باشد، ***فضای آفین*** گویند.

ولی در بسیاری از مسأله‌های ریاضیات، به تعمیم مفهوم طول بردار و مفهوم زاویه در فضای n بعدی نیاز داریم. این تعمیم، در قیاس با نظریهٔ بردارها در صفحه و در فضا، پدید آمده است.

در آغاز فضای حقیقی سطرها را در نظر می‌گیریم. طول بردار

$$X=(x_1, x_2, \dots, x_n)$$

بنابر تعریف، برابر است با عدد $|X|=\sqrt{x_1^2+x_2^2+\dots+x_n^2}$ ، و این طبیعی است، زیرا برای $n=2$ و $n=3$ ، به‌یاری همین برابری، طول بردار را برحسب مختصات آن، در دستگاه محورهای مختصات دکارتی، محاسبه می‌کنند.

مفهوم زاویهٔ بین بردارها هم، به‌طور طبیعی، با ملاحظهٔ زیر در نظر گرفته می‌شود: زاویهٔ بین بردارهای X و Y در صفحه و در فضا، عبارت است از زاویهٔ رأس A در مثلثی با ضلع‌های $AB=|X|$ ، $AC=|Y|$ و $BC=|X-Y|$.

برای فضای nبعدی هم، همین ملاحظه را به‌عنوان تعریف زاویهٔ بین بردارها، می‌پذیرند؛ یعنی فرض می‌کنند، بتوان دو بردار را از فضای nبعدی «جدا کرد» و روی صفحه‌ای «قرار داد»، به‌نحوی که طول و زاویهٔ بین آن‌ها بی‌تغییر بماند. با وجود این، در این تعریف، نوعی بی‌دقتی وجود دارد: وجود مثلث ABC با بردارهای به طول‌های $|X|$ ، $|Y|$ و $|X-Y|$ ، نیاز به اثبات دارد.

با صرف‌نظر از این بی‌دقتی، رابطهٔ مربوط به محاسبهٔ زاویه را می‌دهیم. بنابر رابطهٔ معلومِ مثلثات داریم:

$$BC^2=AB^2+AC^2-2AB\cdot AC\cdot\cos\varphi$$

از آن‌جا

$$\begin{aligned}\cos\varphi &= \frac{|X|^2+|Y|^2-|X-Y|^2}{2|X|\cdot|Y|}\\ &= \frac{x_1^2+\dots+x_n^2+y_1^2+\dots+y_n^2-(x_1-y_1)^2-\dots-(x_n-y_n)^2}{2|X|\cdot|Y|}\\ &= \frac{x_1y_1+\dots+x_ny_n}{|X|\cdot|Y|}\end{aligned}$$

اگر شبیه حالت فضای سه‌بعدی اصطلاح «ضرب اسکالر» را به معنای حاصل‌ضرب طول دو بردار در کسینوس زاویهٔ بین آن‌ها بگیریم، آن‌وقت حاصل‌ضرب اسکالر بردارها را، می‌توان با این رابطه محاسبه کرد:

$$X\cdot Y=x_1y_1+\dots+x_ny_n$$

که برای $n = ۲$ و $n = ۳$، با رابطهٔ مربوط به ضرب اسکالر بردارهای معمولی تطبیق می‌کند.

به‌زبان دقیق‌تر، عبارت $x_۱ y_۱ + \cdots + x_n y_n$ را باید به‌عنوان تعریف ***حاصل‌ضرب اسکالر*** پذیرفت و سپس، زاویهٔ بین دو بردار را، با این را طه تعریف کرد:

$$\cos\varphi = \frac{X \cdot Y}{|X| \cdot |Y|} \qquad (۳)$$

ما هم، به همین ترتیب عمل می‌کنیم.

به‌منظور قانونی بودن این تعریف برای زاویه، باید ثابت کنیم، مقدار سمت راست برابری (۳)، از لحاظ قدرمطلق، از واحد تجاوز نمی‌کند، یعنی $(X \cdot Y)^۲ \leq |X|^۲ + |Y|^۲$.

این نابرابری را می‌توان چنین نوشت:

$$(x_۱ y_۱ + \cdots + x_n y_n)^۲ \leq (x_۱^۲ + \cdots + x_n^۲)\ (y_۱^۲ + \cdots + y_n^۲)$$

این، نابرابری کوشی ـ بونیاکُوْسکی است و می‌توان آن را به‌طور مستقیم، و البته با محاسبه‌ای طولانی، ثابت کرد. در این‌جا آن را، با روش دیگری ثابت می‌کنیم.

یادآوری می‌کنیم، حاصل‌ضرب اسکالر بردارها، دارای این ویژگی‌ها هستند:

'۱. $X \cdot X = |X|^۲ > ۰$ برای $X \neq ۰$؛

'۲. $X \cdot Y = Y \cdot X$؛

'۳. $(cX) \cdot Y = c\,(X \cdot Y)$؛

'۴. $(X_۱ + X_۲) \cdot Y = X_۱ \cdot Y + X_۲ \cdot Y$.

اگر ضرب اسکالر را برحسب مختصات بنویسیم، درستی این ویژگی‌ها به‌طور مستقیم روشن می‌شود.

اکنون بردار $Y + tX$ را وارد می‌کنیم که در آن، t یک عدد حقیقی است. داریم $|Y + tX|^۲ \geq ۰$، زیرا مربع طول یک بردار نمی‌تواند منفی باشد. ولی بنابر ویژگی ضرب اسکالر

$$|Y + tX|^۲ = |Y|^۲ + ۲t X \cdot Y + t^۲ |X|^۲$$

از طرف دیگر، می‌دانیم سه‌جمله‌ای درجهٔ‌دوم به‌ازای همهٔ مقدارهای متغیر حقیقی t، تنها وقتی غیرمنفی می‌ماند که ریشه‌های آن موهومی یا برابر باشند، یعنی وقتی که مبین سه‌جمله‌ای منفی یا صفر باشد. ولی مبین سه‌جمله‌ای $|Y|^۲ + ۲tX \cdot Y + t^۲ |X|^۲$ برابر است با

$4(X\cdot Y)^2-4|X|^2|Y|^2$ و بنابراین

$$(X\cdot Y)^2 - |X|^2|Y|^2 \le 0$$

که با نابرابری کوشی ـ بونیاکُوْسکی هم‌ارز است.

از این نابرابری نتیجه می‌شود: $\frac{|X\cdot Y|}{|X|\,|Y|} \le 1$ و بنابراین، رابطهٔ (۳) برای تعیین زاویه قانونی است.

سپس، این نابرابری هم به‌سادگی ثابت می‌شود:

$$|X|-|Y| \le |X+Y| \le |X|+|Y|$$

که از آن می‌توان وجود مثلث با ضلع‌های $|X|$، $|Y|$ و $|X-Y|$ را نتیجه گرفت. به این ترتیب، عدم دقتی که از لحاظ هندسی، در تعریف عینی زاویه وجود داشت، برطرف می‌شود و، این تعریف، نیروی قانونی پیدا می‌کند.

تعریف اصل موضوعیِ فضای nبعدی اقلیدسی. در بند قبل با مفهوم‌های طول بردار، زاویه و ضرب اسکالر در فضای سطرها، آشنا شدیم. در تعریف کلی اصل موضوعی فضای خطی حقیقی n بعدی، خود این مفهوم‌ها هم به‌صورت اصل موضوعی بیان می‌شوند و درضمن، مفهوم ضرب اسکالر، مبنای کار قرار می‌گیرد.

ضرب اسکالر بردارهای فضای حقیقی خطی، به تناظر هر زوج بردار X و Y، با عددی حقیقی که حاصل‌ضرب اسکالر $X\cdot Y$ آن‌ها نام دارد، گفته می‌شود؛ درضمن، این تناظر باید با این شرط‌ها (اصل موضوع‌ها) سازگار باشد.

۱′. $X\cdot X>0$ برای $X\neq 0$؛ $0\cdot 0=0$؛

۲′. $X\cdot Y=Y\cdot X$؛

۳′. $(cX)\cdot Y=c(X\cdot Y)$؛

۴′. $(X_1+X_2)\cdot Y=X_1\cdot Y+X_2\cdot Y$.

سپس، به‌عنوان طول بردار، عدد $\sqrt{X\cdot X}$ و به‌عنوان کسینوس زاویهٔ بین بردارهای X و Y، عددی برابر $\frac{X\cdot Y}{|X|\cdot|Y|}$ را می‌پذیرند. برای قانونی بودن تعریف اخیر باید درستی نابرابری کوشی ـ بونیاکُوْسکی را ثابت کرد:

$$(X\cdot Y)^2 \le |X|^2\,|Y|^2$$

که شبیه آن‌چه اندکی قبل آوردیم، به نتیجه می‌رسد. در این اثبات، تنها از ویژگی‌های ′۱ تا ′۴ ضرب اسکالر استفاده می‌شود و ویژگی‌های فضای سطرها در این اثبات نقشی ندارد. درواقع، در یک فضای خطی که در آن، ضرب اسکالر وارد شده باشد، که با اصل موضوع‌های ′۱ تا ′۴ سازگار باشد، ***فضای اقلیدسی*** نامیده می‌شود.

در فضاهای خطی متفاوتی که در ریاضیات بررسی می‌شوند، حاصل‌ضرب اسکالر با روش‌های مختلف به‌دست می‌آید و نوع انتخاب این روش، مسألهٔ اصلی هر فضای خطی مشخص را تشکیل می‌دهد. از جمله در فضاهایی که عنصرهای آن‌ها را، تابع‌های با یک متغیر $X(t)$ تشکیل می‌دهند (در بازهٔ مفروض $a \le t \le b$)، به‌عنوان حاصل‌ضرب اسکالر دو عنصر $X(t)$ و $Y(t)$، اغلب عدد

$$\int_a^b X(t)\,Y(t)\,dt \quad \text{یا} \quad \int_a^b X(t)\,Y(t)\,p(t)\,dt$$

پذیرفته می‌شود که در آن، $p(t)$ یک تابع مثبت دل‌خواه است. به‌سادگی دیده می‌شود، در هریک از این تعریف‌ها، اصل موضوع‌های ′۱ تا ′۴ صدق می‌کنند.

تعامد. پایه‌های یکا متعامد. دو بردار از فضای اقلیدسی را متعامد (یا اُرتُوگُونال) گویند، وقتی حاصل‌ضرب اسکالر آن‌ها برابر صفر باشد. به‌سادگی دیده می‌شود، اگر بردارهای غیرصفری، دوبه‌دو متعامد باشند، همواره مستقل خطی‌اند. درواقع، فرض کنیم بردارهای غیرصفر X_1، X_2، ... ، X_m دوبه‌دو متعامد باشند و فرض کنیم داشته باشیم:

$$c_1X_1 + c_2X_2 + \cdots + c_mX_m = 0$$

آن‌وقت، بنا به ویژگی حاصل‌ضرب اسکالر داریم:

$$X_1(c_1X_1 + c_2X_2 + \cdots + c_mX_m) = c_1|X_1|^2 = 0$$

و از آن‌جا $c_1 = 0$. به همین ترتیب ثابت می‌شود: $c_2 = \cdots = c_m = 0$. بنابراین X_1، ... ، X_m مستقل خطی هستند.

از آن‌چه ثابت کردیم، نتیجه می‌شود، در یک فضای nبُعدی، می‌تواند n بردار غیرصفر دوبه‌دو متعامد (و نه بیشتر) وجود داشته باشد؛ هر مجموعه‌ای از n بردارِ دوبه‌دو متعامد پایهٔ فضا را تشکیل می‌دهند. به‌جز این، اگر طول هریک از این بردارها برابر واحد باشد،

پایه‌ای را که از آن‌ها تشکیل می‌شود، ***پایهٔ یکا متعامد*** گویند.

به‌سادگی می‌توان ثابت کرد (ولی در این‌جا، به آن نمی‌پردازیم)، در فضای اقلیدسی، پایه‌های یکا متعامد وجود دارد؛ از این گذشته، تعداد آن‌ها بی‌نهایت است. به‌جز این، اگر در فضای R، یک زیرفضای P انتخاب کنیم، می‌توان پایهٔ یکا متعامد زیرفضا را، با افزودن چند بردار، تا پایهٔ یکا متعامد تمامی فضا، کامل کرد.

بهتر است بردارهای فضای اقلیدسی، به یاری مختصات آن‌ها در یک پایهٔ یکا متعامد داده شود، زیرا در این حالت، به‌ویژه برای حاصل‌ضرب اسکالر، بیان ساده‌ای به‌دست می‌آید. درواقع، اگر مختصات بردارهای X و Y در پایهٔ یکا متعامد $e_۱$، $e_۲$، ...، e_n، به‌ترتیب

$$(x_۱ , x_۲ , \dots , x_n) \quad \text{و} \quad (y_۱ , y_۲ , \dots , y_n)$$

باشد، یعنی داشته باشیم:

$$X = x_۱ e_۱ + x_۲ e_۲ + \cdots + x_n e_n$$

و

$$Y = y_۱ e_۱ + y_۲ e_۲ + \cdots + y_n e_n$$

بنابه‌ویژگی حاصل‌ضرب اسکالر، به‌دست می‌آید:

$$\begin{aligned} X \cdot Y = & \; x_۱ y_۱ e_۱ e_۱ + x_۱ y_۲ e_۱ e_۲ + \cdots + x_۱ y_n e_۱ e_n + \\ & + x_۲ y_۱ e_۲ e_۱ + x_۲ y_۲ e_۲ e_۲ + \cdots + x_۲ y_n e_۲ e_n + \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + x_n y_۱ e_n e_۱ + x_n y_۲ e_n e_۲ + \cdots + x_n y_n e_n e_n = \\ & = x_۱ y_۱ + x_۲ y_۲ + \cdots + x_n y_n \end{aligned}$$

زیرا $e_i e_k = ۰$ $(i \neq k)$ و $e_i e_i = |e_i|^۲ = ۱$ (برای هر عدد طبیعی i از ۱ تا n). در حالت خاص

$$X \cdot X = x_۱^{\,۲} + x_۲^{\,۲} + \cdots + x_n^{\,۲}$$

به‌این ترتیب، طول بردار و حاصل‌ضرب برداری برحسب مختصات پایهٔ یکا متعامد، با همان رابطهٔ فضای سطرها بیان می‌شوند.

عبور از یکی از نمونه‌های فضای اقلیدسی ـ فضای سطرها ـ به فضای اقلیدسی کلی، که

به‌صورت اصل موضوعی تعریف شده باشد، هیچ پیچیدگی تازه‌ای پدید نمی‌آورد، ولی حوزهٔ کاربرد نظریه را گسترش می‌دهد.

به مسألهٔ تصویر قائم بردارها بر یک زیرفضا هم، توجه می‌کنیم. R_n را یک فضای اقلیدسی nبعدی و P_m را یک زیرفضای mبعدی آن می‌گیریم. فرض می‌کنیم e_1، e_2، ... ، e_m، f_1، ... ، f_{n-m}، پایهٔ یکا متعامد در R_n، شامل پایهٔ یکا متعامد زیرفضای P_m باشد. زیرفضای Q_{n-m} را، که از بردارهای f_1، f_2، ... ، f_{n-m}پدید آمده است، **متمم قائم** زیرفضای P_m گویند. بُعد آن برابر با $n-m$ است. این زیرفضای متمم قائم Q_{n-m} را می‌توان همچون مجموعهٔ همهٔ بردارهایی دانست که بر بردارهای زیرفضای P_m عمودند.

هر بردار Z از فضای R_n را می‌توان، به‌صورت یک‌ارزشی، به‌عنوان مجموعی از دو بردار X و Y نوشت که یکی از آن‌ها به زیرفضای P_m و دیگری به زیرفضای Q_{n-m} تعلق داشته باشد. این مطلب نتیجهٔ آن است که می‌توان بردار Z را، به‌طور یک‌ارزشی، به این صورت نشان داد:

$$Z = x_1 e_1 + \cdots + x_m e_m + y_1 f_1 + \cdots + y_{n-m} f_{n-m}$$

به‌نحوی که

$$X = x_1 e_1 + \cdots + x_m e_m \quad ; \quad Y = y_1 f_1 + \cdots + y_{n-m} f_{n-m}$$

بردار X را تصویر قائم بردار Z روی P_m گویند.

فضای یکانی. مفهوم طول بردار و حاصل‌ضرب اسکالر بردارها را، در فضای مختلط هم، می‌توان تعریف کرد. همچون گذشته، مفهوم ضرب اسکالر مبنای کار قرار دارد، و آن را اول ترتیب تعریف می‌کنیم. هر دو بردار X و Y از فضای مختلط را با عددی مختلط متناظر می‌کنیم و آن را حاصل‌ضرب اسکالر آن‌ها $(X \cdot Y)$ می‌نامیم. عمل ضرب اسکالر باید با این اصل موضوع‌ها سازگار باشد:

″۱. $X \cdot X$ برای $X \neq 0$ حقیقی و مثبت است؛ $0 \cdot 0 = 0$؛

″۲. $Y \cdot X = (X \cdot Y)'$؛

در این‌جا، نماد پریم $(')$، به‌معنای گذر به عدد مزدوج مختلط است؛

″۳. $(cX) \cdot Y = c(X \cdot Y)$ برای هر عدد مختلط c؛

″۴. $(X_1 + X_2) \cdot Y = X_1 \cdot Y + X_2 \cdot Y$، قانون پخشی.

در فضای سطرها با عددهای مختلط، می‌توان حاصل‌ضرب اسکالر بردارهای

$$(x_1, x_2, \dots, x_n) \text{ و } (y_1, y_2, \dots, y_n)$$

را برابر عدد $x_1 y_1' + x_2 y_2' + \dots + x_n y_n'$ گرفت و به‌سادگی ثابت کرد، برای این تعریف، اصل موضوع‌های "۱ تا "۴ صادق‌اند.

به‌عنوان طول بردار، عدد $\sqrt{X \cdot X}$ را می‌پذیرند. مفهوم زاویه بین بردارها تعریف نمی‌شود.

فضای خطی مختلط، همراه با حاصل‌ضرب برداریِ صادق در اصل موضوع‌های "۱ تا "۴ را، ***فضای یکانی*** گویند.

۳. دستگاه معادله‌های خطی

دستگاه دومعادلهٔ دومجهولی و سه‌معادلهٔ سه‌مجهولی. دستگاه دومعادلهٔ دومجهولی در حالت کلی، به این صورت است:

$$\begin{cases} a_1 x + b_1 y = c_1 \\ a_2 x + b_2 y = c_2 \end{cases}$$

برای حل این دستگاه، معادلهٔ اول را در b_2 و معادلهٔ دوم را در $-b_1$ ضرب و سپس، با هم جمع می‌کنیم، به‌دست می‌آید:

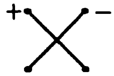
$$(a_1 b_2 - a_2 b_1)x = c_1 b_2 - c_2 b_1$$

به همین ترتیب، اگر معادلهٔ اول را در $-a_2$ و معادلهٔ دوم را در a_1 ضرب و سپس با هم جمع کنیم، به‌دست می‌آید:

$$(a_1 b_2 - a_2 b_1)y = a_1 c_2 - a_2 c_1$$

از این برابری‌ها، مقدارهای x و y به‌دست می‌آید، تنها با این شرط که عبارت $a_1 b_2 - a_2 b_1$، یعنی ضریب مجهول‌های x و y مخالف صفر باشد. این عبارت را ***دترمینانِ*** ماتریس $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ که از ضریب‌های مجهول‌ها در دستگاه درست شده است، می‌نامند.

دترمینان را به‌صورت $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ نشان می‌دهند. از این‌جا نتیجه می‌شود که دترمینان، با این طرح محاسبه می‌شود:



,

و نیازی به توضیح بیشتر ندارد.

به حل دستگاه برمی‌گردیم. عبارت‌های $c_1b_2-c_2b_1$ و $a_1c_2-a_2c_1$ هم، بنا بر همان تعریف عبارت‌اند از دترمینان‌های

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \quad \text{و} \quad \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

به این ترتیب، اگر مقدار دترمینان $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ برابر صفر نباشد، آن‌وقت به این دستورها، برای جواب دستگاه می‌رسیم:

$$x=\frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \quad \text{و} \quad y=\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \qquad (۴)$$

اگر بخواهیم دقیق‌تر باشیم، باید استدلال خود را کامل‌تر کنیم. عمل‌هایی که روی معادله‌ها انجام دادیم تا به دستورهای جواب دستگاه برسیم، به‌خودی‌خود، با این فرض همراه است که x و y عبارت‌اند از عددهایی که جواب دستگاه را تشکیل می‌دهند. ماهیت منطقی این استدلال، چنین است: اگر دترمینان ضریب‌های مجهول در دستگاه مخالف صفر باشد، و اگر جواب دستگاه وجود داشته باشد، آن‌گاه این جواب، از دستورهای (۴) به‌دست می‌آید. بنابراین، هنوز لازم است ثابت کنیم، مقدارهایی که برای مجهول‌ها پیدا کرده‌ایم، به‌واقع در هر دو معادلهٔ دستگاه صدق می‌کند. این آزمایش به‌سادگی انجام می‌شود.

بنابراین، اگر دترمینان ماتریس ضریب‌های مجهول‌ها در دستگاه مخالف صفر باشد، دستگاه جوابی منحصر دارد که با دستورهای (۴) داده می‌شود.

برای دستگاه سه‌معادلهٔ سه‌مجهولیِ

$$\begin{aligned} a_1x+b_1y+c_1z&=d_1 \\ a_2x+b_2y+c_2z&=d_2 \\ a_3x+b_3y+c_3z&=d_3 \end{aligned}$$

هم می‌توان شبیه حالت دومعادلهٔ دومجهولی استدلال کرد و به نتیجه رسید. برای این منظور، باید معادله‌ها را به‌ترتیب در ضریب‌هایی ضرب کنیم که از مجموع آن‌ها، یکباره دوتا از مجهول‌ها حذف شوند. در مثل، برای حذف مجهول‌های y و z، باید به‌ترتیب، ضریب‌های $b_۲c_۳-b_۳c_۲$، $b_۳c_۱-b_۱c_۳$ و $b_۱c_۲-b_۲c_۱$ را برگزید که می‌توان، به‌سادگی، محاسبه را آزمایش کرد.

درنتیجه معلوم می‌شود، اگر عبارت

$$\Delta=a_۱b_۲c_۳-a_۱b_۳c_۲+a_۲b_۳c_۱-a_۲b_۱c_۳+a_۳b_۱c_۲-a_۳b_۲c_۱$$

برابر صفر نباشد، دستگاه دارای جوابی منحصر است که از این دستورها به‌دست می‌آید:

$$x=\frac{\Delta_۱}{\Delta} \quad , \quad y=\frac{\Delta_۲}{\Delta} \quad , \quad z=\frac{\Delta_۳}{\Delta}$$

در این دستورها، $\Delta_۱$، $\Delta_۲$ و $\Delta_۳$، به این طریق به‌دست می‌آیند که، به‌ترتیب، در Δ، به‌جای ضریب‌های مجهولِ مربوط، مقدارهای ثابت را قرار دهیم.

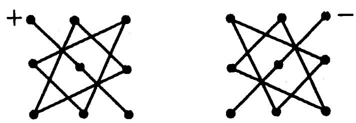
عبارت Δ را دترمینانِ ماتریسِ

$$\begin{pmatrix} a_۱ & b_۱ & c_۱ \\ a_۲ & b_۲ & c_۲ \\ a_۳ & b_۳ & c_۳ \end{pmatrix}$$

گویند و به این صورت نشان می‌دهند:

$$\begin{vmatrix} a_۱ & b_۱ & c_۱ \\ a_۲ & b_۲ & c_۲ \\ a_۳ & b_۳ & c_۳ \end{vmatrix}$$

برای محاسبهٔ دترمینان، بهتر است از این طرح استفاده کنیم:



روی طرح سمت چپ، به‌وسیلهٔ خط‌های راست (قطر دو مثلث)، نقطه‌هایی به‌هم وصل شده‌اند که درواقع، درایه‌هایی از دترمینان را نشان می‌دهند که حاصل‌ضرب آن‌ها در حاصل دترمینان با علامت مثبت آمده است؛ همچنین در طرح سمت راست، جمله‌هایی از دترمینان مشخص شده است که علامت منفی دارند.

برای دستگاه دومعادلهٔ دومجهولی و دستگاه سه‌معادلهٔ سه‌مجهولی، به نتیجه‌های مشابهی رسیدیم. در هر دو حالت، دستگاه دارای جواب منحصر است، به‌شرطی که دترمینان ماتریس ضریب‌ها (ضریب‌های مجهول‌ها) مخالف صفر باشد. دستورهای مربوط به جواب هم، شبیه یکدیگرند: در مخرج هریک از مجهول‌ها، دترمینان ماتریس ضریب‌های مجهول‌ها قرار دارد؛ در صورت‌ها هم، دترمینان ماتریسی قرار دارد که از دترمینان مخرج، با تبدیل ضریب‌های مجهول مربوط، به عددهای ثابت، به‌دست می‌آید.

تعمیم مستقیم این نتیجه‌گیری‌ها، به دستگاه nمعادلهٔ nمجهولی (برای هر عدد طبیعی n) دشواری‌هایی دارد. ولی با روشی غیرمستقیم، می‌توان کار را ساده‌تر کرد: در آغاز، مفهوم دترمینان را روی ماتریس مربعیِ با مرتبهٔ دل‌خواه، تعمیم می‌دهیم؛ سپس، با بررسی ویژگی‌های دترمینان، آن‌ها را در بررسی نظریهٔ دستگاه‌ها به‌کار می‌بریم.

دترمینان مرتبهٔ n. ضمن بررسی دترمینان‌های مرتبهٔ دوم و مرتبهٔ سوم، یعنی

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 \quad ,$$

و

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1$$

متوجه می‌شویم، هر جمله از حاصل دترمینان، حاصل‌ضرب یک درایه از هر سطر در یک درایه از هر ستونِ دترمینان است؛ درایه‌هایی که یک جمله را پدید می‌آورند، در یک سطر یا یک ستون نیستند؛ درضمن همهٔ حاصل‌ضرب ممکن از این‌گونه در حاصل دترمینان دیده می‌شود که با علامت‌های مثبت یا منفی آمده‌اند. همین ویژگی اساس تعمیم مفهوم دترمینانِ ماتریس مربعی، از مرتبهٔ دل‌خواه، قرار دارد: دترمینان مربعی ماتریس مرتبهٔ n، یا کوتاه‌تر، **دترمینان مرتبهٔ n**، به مجموع جبری همهٔ حاصل‌ضرب‌های ممکنِ درایه‌های ماتریس گفته می‌شود، به‌نحوی که در هر جملهٔ این مجموع جبری، از هر سطر و هر ستون، یکی و تنها یکی از درایه‌های آن سطر یا ستون انتخاب شده باشد؛ درضمن جمله‌های این مجموع، بنابر قانون معینی، با علامت مثبت یا منفی در نظر گرفته می‌شوند. این قانون به‌اندازهٔ کافی پیچیده است و در این‌جا به تنظیم آن نمی‌پردازیم. ولی یادآوری می‌کنیم، این قانون طوری تنظیم شده است که با این ویژگی‌های اصلی دترمینان سازگار باشد.

۱. با جابه‌جا کردن دو سطر از دترمینان، به‌نحوی که هرکدام از این دو سطر جای دیگری

را بگیرد، علامت حاصل دترمینان تغییر کند.

این ویژگی را دربارهٔ دترمینان‌های مرتبهٔ دوم و مرتبهٔ سوم، می‌توان به‌طور مستقیم و با محاسبه آزمایش کرد. در حالت کلی، این ویژگی، به‌یاری قانونی که برای علامت‌ها وجود دارد (و ما در این‌جا نیاوردیم) ثابت می‌شود.

دترمینان‌ها، ویژگی‌های مهم دیگری هم دارند، که به‌یاری آن‌ها می‌توان باموفقیت مسأله‌های نظری و محاسبه‌ای مربوط به دترمینان را (به هر اندازه که این دترمینان بزرگ باشد) حل کرد: به‌سادگی می‌توان محاسبه کرد که حاصل یک دترمینان مرتبهٔ n، دارای $n!$ جمله است که هر جملهٔ آن، شامل n عامل ضرب است که علامت آن بنابر قانونی معین می‌شود.

ویژگی‌های اصلی دترمینان را، بدون این که دربارهٔ اثبات تفصیلی آن‌ها بحث کنیم، دنبال می‌کنیم. نخستین ویژگی را دیدیم؛ اکنون به بقیهٔ ویژگی‌های دترمینان می‌پردازیم.

۲. دترمینان با تبدیل سطرها به ستون‌ها، و ستون‌ها به سطرها، به‌شرطی ردیف آن‌ها به‌هم نخورد، تغییر نمی‌کند.

اثبات این ویژگی، براساس بررسی مفصل قانون ترتیب علامت‌ها، در جمله‌های دترمینان، به‌دست می‌آید. این ویژگی، امکان می‌دهد، هر گزارهٔ مربوط به سطرهای دترمینان را، به‌ستونهای آن سرایت دهیم.

۳. دترمینان، نسبت به درایه‌های یک سطر (یا یک ستون خود)، تابعی خطی است؛ یعنی می‌توان دترمینان را به‌صورت

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \qquad (۵)$$

نوشت که در آن، A_{i1}، A_{i2}، ... ، A_{in}، عبارت‌هایی مستقل از سطر iام هستند.

این ویژگی نتیجه روشنی از این شرط است که هر جمله از حاصل دترمینان، شامل یک و تنها یک درایه از هر سطر و از جمله سطر iام است.

برابری (۵)، مُعرف باز کردن دترمینان برحسب درایه‌های سطر iام است؛ ضریب‌های A_{i1}، A_{i2}، ... ، A_{in} را مکمل‌های جبری درایه‌های a_{i1}، a_{i2}، ... ، a_{in} در دترمینان می‌نامند.

۴. مکمل جبری A_{ij} از درایهٔ a_{ij}، با تقریب علامت، برابر است با «زیردترمینان» Δ_{ij}،

یعنی دترمینانی با مرتبهٔ $(n-1)$، که از دترمینان داده‌شده، با حذف سطر iام و ستون jام به‌دست می‌آید. برای به‌دست آوردن این مکمل جبری، باید علامت آن را $(-1)^{i+j}$ گرفت. ویژگی‌های ۳ و ۴ این امکان را فراهم می‌آورند که بتوانیم دترمینانِ از مرتبهٔ n را، به محاسبهٔ n دترمینانِ از مرتبهٔ $(n-1)$ منجر کنیم.

با توجه به همین ویژگی‌ها، می‌توان ویژگی‌های جالب دیگری از دترمینان نتیجه گرفت.

۵. دترمینانی که درایه‌های دو سطر آن یکسان باشند، برابر صفر است.

درواقع، اگر دترمینانی، دو سطر یکسان داشته باشد، با توجه به ویژگی، باید با تبدیل این دو سطر یکسان به یکدیگر، به دترمینانی تبدیل شود که مقدار آن علامتی مخالف علامت دترمینان اصلی داشته باشد؛ درحالی که با تبدیل این دو سطر به یکدیگر، در دترمینان تغییری پیدا نمی‌شود. بنابراین، تنها می‌تواند برابر صفر باشد.

۶. مجموع حاصل‌ضرب‌های درایه‌های یک سطر در مکمل جبری سطر دیگر، برابر صفر است.

درواقع، این مجموع، عبارت است از نتیجهٔ باز کردن دترمینانی با دو سطر یکسان، برحسب یکی از آن‌ها.

۷. اگر درایه‌های یک سطر، بخشیاب مشترکی داشته باشند، می‌توان این بخشیاب مشترک را از آن سطر به جلو دترمینان منتقل کرد.

این ویژگی نتیجه‌ای است از ویژگی ۳.

۸. اگر درایه‌های دو سطر، متناسب باشند، مقدار دترمینان برابر صفر است.

کافی است ضریب تناسب را از دترمینان بیرون ببریم و دترمینانی با دو سطر یکسان به‌دست آوریم.

۹. اگر به درایه‌های یک سطر، عددهایی متناسب با سطر دیگر را اضافه کنیم، مقدار دترمینان تغییر نمی‌کند.

درواقع، با توجه به ویژگی ۳، دترمینان جدید، برابر است با مجموع دو دترمینان: دترمینان اصلی و دترمینانی که درایه‌های دو سطر آن متناسب و، درنتیجه، مقداری برابر صفر دارد.

ویژگی آخر، وسیلهٔ خوبی برای محاسبهٔ مقدار دترمینان‌هاست. به‌یاری این ویژگی می‌توان، بدون این که مقدار دترمینان تغییر کند، ماتریس آن را طوری تبدیل کنیم که در یک سطر (یا یک ستون) همهٔ درایه‌ها، به‌جز یکی، برابر صفر شوند. سپس، اگر دترمینان را

نسبت به درایه‌های همین سطر (یا ستون) باز کنیم، به‌جای محاسبهٔ دترمینانِ مرتبهٔ n، به دترمینانی از مرتبهٔ $(n-1)$ می‌رسیم: مکمل جبری درایه‌ای از سطر که مخالف صفر است.

مثال. می‌خواهیم مقدار این دترمینان را محاسبه کنیم:

$$\Delta = \begin{vmatrix} 1 & 1 & -1 & 2 \\ 2 & -1 & 1 & 1 \\ -1 & 2 & 0 & 1 \\ 1 & 1 & -2 & 1 \end{vmatrix}$$

به‌درایه‌های ستون دوم، قرینهٔ درایه‌های ستون اول؛ به درایه‌های ستون سوم، درایه‌های ستون اول؛ و به درایه‌های ستون چهارم، قرینهٔ دوبرابر درایه‌های ستون اول را می‌افزاییم؛ به این دترمینان می‌رسیم:

$$\Delta = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -3 & 3 & -3 \\ -1 & 3 & -1 & 3 \\ 1 & 0 & -1 & -1 \end{vmatrix}$$

اگر Δ را برحسب درایه‌های سطر اول باز کنیم، به‌دست می‌آید:

$$\Delta = 1 \times (-1)^{1+1} \begin{vmatrix} -3 & 3 & -3 \\ 3 & -1 & 3 \\ 0 & -1 & -1 \end{vmatrix}$$

اکنون، در دترمینان مرتبهٔ ۳ که به‌دست آمده است، به‌جای سطر اول، مجموع آن را با سطر دوم می‌نویسیم و دترمینان را برحسب ستون اول باز می‌کنیم، به‌دست می‌آید:

$$\Delta = \begin{vmatrix} 0 & 2 & 0 \\ 3 & -1 & 3 \\ 0 & -1 & -1 \end{vmatrix} = 3 \times (-1)^{1+2} \begin{vmatrix} 2 & 0 \\ -1 & -1 \end{vmatrix} = -3 \times (-2) = 6$$

دترمینان ماتریس A را با نماد $|A|$ نشان می‌دهند.

در پایان، ویژگی بسیار مهم دیگری از دترمینان می‌آوریم.

دترمینان حاصل‌ضرب دو ماتریس مربعی، برابر است با حاصل‌ضرب دترمینان‌های دو

ماتریس، یعنی با توجه به نمادی که برای دترمینانِ ماتریس در نظر گرفتیم، داریم:

$$|AB| = |A| \ |B|$$

این ویژگی، به‌خصوص، این امکان را به ما می‌دهد که حاصل‌ضربِ دترمینان‌های هم‌مرتبه را، بنا بر قانون ضرب ماتریس‌ها، به‌دست آوریم.

دستگاه n معادلهٔ خطی با n مجهول. اکنون، با تکیه بر دترمینان، می‌توانیم نتیجه‌هایی را که برای دستگاه دومعادلهٔ دومجهولی و سه‌معادلهٔ سه‌مجهولی به‌دست آوردیم، تعمیم دهیم و برای دستگاهی که دارای n معادله با n مجهول است به‌کار بریم، به‌شرطی که دترمینان ماتریس ضریب‌های مجهول‌ها، مخالف صفر باشد.

این دستگاه n معادلهٔ خطیِ شامل n مجهول را در نظر می‌گیریم:

$$\begin{array}{l} a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1 \quad , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nj}x_j + \cdots + a_{nn}x_n = b_n \end{array}$$

دترمینان ماتریس ضریب مجهول‌های این دستگاه را Δ می‌نامیم. بنابر فرض $\Delta \neq 0$؛ همچنین مکمل جبری درایهٔ a_{ij} را با نماد A_{ij} نشان می‌دهیم. معادلهٔ اول دستگاه را در A_{1j}، معادلهٔ دوم را در A_{2j}، ...، معادلهٔ nام را در A_{nj} ضرب و سپس، همهٔ معادله‌ها را با هم جمع می‌کنیم. به‌دست می‌آید:

$$\Delta x_j = b_1 A_{1j} + \cdots + b_n A_{nj}$$

درواقع، ضریب‌های همهٔ مجهول‌ها، به‌جز x_j، برابر صفر می‌شوند، زیرا این ضریب‌ها، برابر با مجموع حاصل‌ضرب‌های درایه‌های ستون jام در درایه‌های ستون‌های دیگرند (ویژگی ۶ که دربارهٔ ستون‌ها به‌کار رفته است)؛ ضریب مجهول x_j هم برابر است با مجموع حاصل‌ضرب‌های درایه‌های ستون jام در مکمل‌های جبریشان یعنی Δ.

به این ترتیب

$$x_j = \frac{b_1 A_{1j} + \cdots + b_n A_{nj}}{\Delta} \qquad (j = 1, 2, \ldots, n) \qquad (6)$$

همان‌طور که پیش از این هم گفته‌ایم، این نتیجه‌گیری در حالتی درست است که x_1، x_2، ... ، x_n را جواب دستگاه به‌حساب آوریم که وجود آن را باید پذیرفته باشیم.

نتیجهٔ استدلال را می‌توان به این ترتیب آورد: اگر برای دستگاه، جوابی وجود داشته باشد، آن‌وقت، این جواب منحصربه‌فرد است و با برابری‌های (۶) به‌دست می‌آید.

برای کامل بودن حل دستگاه، باید وجود جواب ثابت شود که از راه قرار دادن مقدارهای حاصل، به‌جای مجهول‌ها، در همهٔ معادله‌های دستگاه به‌دست می‌آید. با استفاده از ویژگی دترمینان هم (وقتی برای سطرها به‌کار می‌رود) می‌توان قانع شد که جواب درواقع در همهٔ معادله‌های دستگاه صدق می‌کند.

به این ترتیب، این قضیه درست است: اگر دترمینان ماتریس ضریب‌های مجهول‌ها، در n معادلهٔ خطیِ شامل n مجهول، برابر صفر نباشد، آن‌گاه، دستگاه یک جواب منحصر دارد که با دستورهای (۶) به‌دست می‌آید.

با توجه به این که مجموع $b_1 A_{1j} + \cdots + b_n A_{nj}$ را می‌توان به‌صورت این دترمینان نوشت:

$$\Delta_j = b_1 A_{1j} + \cdots + b_n A_{nj} = \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix}$$

(جمله‌های ثابت در ستون jام‌اند)، می‌توان دستورهای (۶) را طور دیگری هم در نظر گرفت.

به این ترتیب، نتیجه‌گیری‌های مربوط به دستگاه‌های دومعادلهٔ دومجهولی و سه‌معادلهٔ سه‌مجهولی، به‌طور کامل برای دستگاه n معادلهٔ n مجهولی تعمیم داده شد، به‌نحوی که دستورهای جواب هم، به همان صورت حالت‌های خاص‌اند.

یکی از نتیجه‌های این قضیه را می‌آوریم:

اگر از قبل بدانیم، دستگاه معادله‌های خطی جواب ندارد یا تعداد جواب‌ها بیش از یکی است، به‌معنای آن است که دترمینان ماتریس ضریب‌های مجهول‌ها، برابر صفر است.

از این نتیجه، به‌ویژه دربارهٔ دستگاه‌های همگِن، یعنی دستگاه‌هایی که در آنها، مقدارهای ثابت b_1، b_2، ... ، b_n برابر صفرند، اغلب کاربرد دارد. یک دستگاه همگن، همیشه جواب «بدیهی» زیر را می‌پذیرد:

$$x_1 = x_2 = \cdots = x_n = 0$$

در حالتی، که دستگاه همگن، به‌جز این جواب، جواب «نابدیهی» هم داشته باشد، دترمینان ضریب‌ها برابر صفر است.

این گزاره، راه را برای استفاده از نظریهٔ دترمینان‌ها در شاخه‌های دیگر ریاضیات و در کاربرد آن، می‌گشاید.

برای نمونه، مثالی از هندسهٔ تحلیلی می‌آوریم.

مطلوب است معادلهٔ صفحه‌ای که از سه نقطهٔ (x_1, y_1, z_1)، (x_2, y_2, z_2) و (x_3, y_3, z_3) می‌گذرد، به‌شرطی که این سه نقطه بر یک خط راست نباشند.

از هندسهٔ مقدماتی می‌دانیم، چنین صفحه‌ای وجود دارد. معادلهٔ آن را به‌صورت $Ax + By + Cz + D = 0$ در نظر می‌گیریم. چون صفحه از سه نقطهٔ مفروض می‌گذرد، باید داشته باشیم:

$$Ax_1 + By_1 + Cz_1 + D = 0$$
$$Ax_2 + By_2 + Cz_2 + D = 0$$
$$Ax_3 + By_3 + Cz_3 + D = 0$$

اگر (x, y, z) را مختصات نقطهٔ دل‌خواهی از صفحه بگیریم، درضمن داریم:

$$Ax + By + Cz + D = 0$$

این چهار معادله را همچون دستگاه معادله‌های خطی همگن، نسبت به ضریب‌های A، B، C و D، به‌عنوان مجهول‌های دستگاه، در نظر می‌گیریم. این دستگاه، جواب «نابدیهی» دارد، زیرا صفحه‌ای که به‌دنبال آن هستیم، وجود دارد. درنتیجه، دترمینان ضریب‌های دستگاه برابر صفر است، یعنی

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x & y & z & 1 \end{vmatrix} = 0 \qquad (7)$$

همین معادله، معرف صفحهٔ مجهول است. درواقع معادلهٔ (۷)، نسبت به x، y و z از درجهٔ اول است، زیرا دترمینان نسبت به درایه‌های سطر آخر خطی است.

با توجه به این که نقطه‌های مفروض بر یک خط راست نیستند، می‌توان به‌سادگی تحقیق

کرد که، همهٔ ضریب‌های آن هم، برابر صفر نیستند. بنابراین، معادلهٔ (۷) به‌واقع معادلهٔ یک صفحه است. این صفحه از نقطه‌های مفروض می‌گذرد، زیرا مختصات آنها، به‌روشنی، در معادلهٔ (۷) صدق می‌کنند.

معرفی n معادله با n مجهول با نماد ماتریسی. دستگاه n معادلهٔ خطی n مجهولیِ

$$a_{11}x_1+\cdots+a_{1n}x_n=b_1$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{n1}x_1+\cdots+a_{nn}x_n=b_n$$

را با نمادهای ماتریسی، می‌توان به‌صورت یک برابری نوشت:

$$AX=B$$

که در آن، A به‌معنای ماتریس ضریب‌های مجهول‌ها، X ستون متناظر مجهول‌ها و B ستون عددهای ثابت است.

جواب دستگاه (به‌شرطی که دترمینان ماتریس A برابر صفر نباشد)، چنین است (دستور (۶) را ببینید):

$$x_1=\frac{A_{11}}{\Delta}b_1+\frac{A_{21}}{\Delta}b_2+\cdots+\frac{A_{n1}}{\Delta}b_n \quad ,$$
$$x_2=\frac{A_{12}}{\Delta}b_1+\frac{A_{22}}{\Delta}b_2+\cdots+\frac{A_{n2}}{\Delta}b_n \quad ,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x_n=\frac{A_{1n}}{\Delta}b_1+\frac{A_{2n}}{\Delta}b_2+\cdots+\frac{A_{nn}}{\Delta}b_n$$

ویا به‌صورت ماتریسی آن:

$$X=\begin{pmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \cdots & \frac{A_{n1}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \cdots & \frac{A_{n2}}{\Delta} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{A_{1n}}{\Delta} & \frac{A_{2n}}{\Delta} & \cdots & \frac{A_{nn}}{\Delta} \end{pmatrix} B$$

ماتریسی که در سمت راست برابری و به‌عنوان ضریب B آمده است، ***ماتریس وارونِ*** ماتریس A نام دارد و به‌طور طبیعی با نماد A^{-1} نشان داده می‌شود. اگر این نماد را به‌کار ببریم، جواب دستگاه $AX=B$ را می‌توانیم به‌صورتی بنویسیم که دستور جواب معادلهٔ خطی یک‌مجهولی را به‌یاد می‌آورد، یعنی به‌صورت

$$X = A^{-1}B$$

با استفاده از اصطلاح‌های جبر ماتریس‌ها، نماد دیگری هم برای این نتیجه می‌توان پیدا کرد. پیش از هر چیز، باید به ماتریس

$$E = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

به نام ***ماتریس همانی (یا واحد)***، که نقشی اساسی دارد، توجه کنیم.

ماتریس همانی در بین ماتریس‌های مربعی، همان نقشی را به‌عهده دارد که عدد ۱ در بین همهٔ عددها به‌عهده گرفته است: برای هر ماتریس A، برابری‌های $AE=A$ و $EA=A$ برقرار است که به‌سادگی و براساس ویژگی‌های ضرب ماتریس‌ها قابل تحقیق است.

ماتریس A^{-1}، به‌عنوان وارونِ ماتریس A، نقشی شبیه نقش وارونِ معمولی عدد به‌عهده دارد:

$$AA^{-1} = A^{-1}A = E$$

درستی این برابری‌ها براساس قانون‌های ضرب ماتریس‌ها و ویژگی‌های ۳ و ۶ دترمینان‌ها، قابل تحقیق است.

با در نظر گرفتن ویژگی‌های ماتریس همانی و ماتریس وارون، می‌توان جواب دستگاه $AX=B$ را به این صورت توضیح داد.

فرض کنید $AX=B$. در این صورت $A^{-1}(AX)=A^{-1}B$. ولی

$$A^{-1}(AX) = (A^{-1}A)X = EA = X$$

و بنابراین $X=A^{-1}B$.

اکنون فرض کنیم $X=A^{-1}B$. در این صورت

$$AX = AA^{-1}B = EB = B$$

به این ترتیب، معادلهٔ $AX=B$ ، وقتی و تنها وقتی دارای جواب منحصربه‌فرد $X=A^{-1}B$ است که A^{-1} وجود داشته باشد.

دیدیم ماتریس وارون A^{-1} برای ماتریس A با این شرط وجود دارد که دترمینان ماتریس A مخالف صفر باشد. این شرط، برای وجود ماتریس وارون نه‌تنها کافی، بلکه لازم است. درواقع، اگر برای ماتریس A ، ماتریس وارون A^{-1} وجود داشته باشد، یعنی برابری $AA^{-1}=E$ برقرار باشد، آن‌وقت بنا بر ویژگی دترمینان حاصل‌ضرب دو ماتریس

$$|A|\ |A^{-1}| = |E| = 1$$

که از آن‌جا نتیجه می‌شود، دترمینان ماتریس A برابر صفر نیست.

ماتریسی که دترمینان آن مخالف صفر باشد، *ناتکین* نامیده می‌شود. به این ترتیب روشن کردیم، برای ماتریس‌های ناتکین و تنها برای آنها، همیشه ماتریس وارون وجود دارد.

وارد کردن مفهوم ماتریس وارون، نه‌تنها برای نظریهٔ دستگاه‌های معادله‌های خطی، بلکه برای بسیاری از مسأله‌های دیگر جبر خطی، سودمند است.

در پایان یادآوری می‌کنیم، دستورهای جواب در دستگاه معادله‌های خطی، وسیله‌ای ضروری برای بررسی‌های نظری است، ولی برای به‌دست آوردن جواب عددی دستگاه‌ها کاربرد زیادی ندارد.

همان‌طور که پیش از این هم گفتیم، برای جواب عددی دستگاه‌ها، روش‌های مختلف و طرح‌های محاسبه‌ای زیادی وجود دارد و به‌دلیل اهمیت زیادی که این مسأله برای بررسی‌های عملی دارد، ساده‌تر کردن روش‌های پیدا کردن جواب عددی در دستگاه‌های خطی (به‌ویژه، وقتی تعداد مجهول‌ها زیاد است)، همیشه جالب بوده و امروز هم جالب است.

حالت کلی دستگاه معادله‌های خطی. به بررسی دستگاه معادله‌های خطی در حالت کلی، ولو این که تعداد معادله‌ها برابر با تعداد مجهول‌ها نباشد، می‌پردازیم. در این طرح کلی، نمی‌توان

انتظار داشت، همیشه جوابی برای دستگاه وجود داشته باشد ویا در حالتی که دستگاه جواب دارد، این جواب منحصربه‌فرد باشد. طبیعی است پیش‌بینی کنیم، اگر تعداد معادله‌ها کمتر از تعداد مجهول‌ها باشد، آن‌وقت دستگاه بی‌نهایت جواب داشته باشد. برای نمونه، در دو معادلهٔ درجه‌اول با سه مجهول، مختصات هر نقطهٔ دل‌خواه از خط راستی که محل برخورد دو صفحه است (دو صفحه‌ای که معادلهٔ هرکدام با یکی از معادله‌های دستگاه داده شده است)، در دو معادلهٔ دستگاه صدق کند، ولی ممکن است در این حالت، جوابی برای دستگاه وجود نداشته باشد، وقتی که دو صفحهٔ مُعرف معادله‌ها با هم موازی باشند. اگر تعداد معادله‌ها بیش از تعداد مجهول‌ها باشد، آن‌وقت دستگاه، اغلب بدون جواب است. با وجود این، در این حالت هم ممکن است دستگاه جواب داشته باشد و حتی بی‌نهایت جواب.

برای بررسی مسألهٔ مربوط به وجود جواب یا جواب‌ها، در دستگاهی که به این صورت کلی در برابر ما قرار دارد، به بررسی «هندسی» دستگاه رو می‌آوریم.

دستگاه معادله‌های

$$\begin{array}{l} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \quad , \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \quad , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{array} \qquad (8)$$

را در فضای n بُعدی ستون‌ها و به این صورت تفسیر می‌کنیم:

$$x_1A_1 + x_2A_2 + \cdots + x_nA_n = B$$

که در آن، A_1، A_2، ... ، A_n به‌معنای ستون‌هایی از ضریب‌های متناظر مجهول‌ها، و B ستون جمله‌های ثابت است.

در این تفسیر، مسألهٔ مربوط به وجود جواب در دستگاه، به این مسأله منجر می‌شود که: آیا بردار B، ترکیبی خطی از بردارهای A_1، A_2، ... ، A_n است یا نه؟

پاسخ به این پرسش کم و بیش روشن است. برای این که بردار B ترکیب خطی از بردارهای A_1، A_2، ... ، A_n باشد، لازم و کافی است، بردار B در زیرفضایی باشد که از A_1، A_2، ... ، A_n پدید آمده است، به‌زبان دیگر، زیرفضای پدید آمده از A_1، A_2، ... ، A_n و B منطبق بر زیرفضایی باشد که از A_1، A_2، ... ، A_n پدید آمده است.

از آن‌جا که زیرفضای اول (زیرفضای پدید آمده از بردارهای A_1، A_2، ... ، A_n و B) شامل زیرفضای دوم است، برای این که برهم منطبق باشند، لازم و کافی است که، بُعد آنها، برابر باشد. به‌یاد می‌آوریم، بُعد زیرفضایی که از دستگاه مفروضی از بردارها پدید آمده باشد، رتبهٔ این دستگاه بردارها نامیده می‌شود. بنابراین شرط لازم و کافی برای وجود جواب در دستگاه

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = B$$

عبارت است از برابری رتبه‌ها در دستگاه بردارهای

$$A_1 , A_2 , \ldots , A_n \quad \text{و} \quad A_1 , A_2 , \ldots , A_n , B$$

می‌توان ثابت کرد (که در این‌جا به آن نمی‌پردازیم)، رتبهٔ هر دستگاهی از بردارها، برابر است با رتبهٔ ماتریسی که شامل مختصات این بردارهاست. در این‌جا، منظور از ***رتبهٔ ماتریس***، عبارت است از بالاترین مرتبهٔ دترمینان مخالف صفری است که می‌توان از ماتریس مفروض، به‌کمک بخشی از سطرها و ستون تشکیل داد.

از آن‌جا که مختصات بردارهای A_1، A_2، ... ، A_n (در پایهٔ طبیعی برای فضای ستون‌ها) ضریب‌های دستگاه، و مختصات بردار B، جمله‌های ثابت آن هستند، بنابراین می‌توانیم شکل نهایی شرط وجود جواب در دستگاه را تنظیم کنیم.

برای وجود دست‌کم یک جواب در دستگاه معادله‌های خطی

$$\begin{array}{l} a_{11} x_1 + \cdots + a_{1n} x_n = b_1 \quad , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1} x_1 + \cdots + a_{mn} x_n = b_m \end{array}$$

لازم و کافی است، رتبهٔ ماتریسی که از ضریب‌های دستگاه تشکیل شده است، با رتبهٔ ماتریس شامل ضریب‌ها و مقدارهای ثابت، برابر باشد.

اکنون به بررسی خصلت مجموعهٔ جواب‌ها، اگر وجود داشته باشند، می‌پردازیم. x_1°، x_2°، ... ، x_n° را جوابی از دستگاه (۸) می‌گیریم و فرض می‌کنیم:

$$x_1 = x_1^\circ + y_1 \quad , \quad x_2 = x_2^\circ + y_2 \quad , \quad \ldots \quad , \quad x_n = x_n^\circ + y_n$$

در این صورت، با توجه به این که x_1°، x_2°، ... ، x_n° جوابی از دستگاه (۸) هستند، مجهول‌های

تازهٔ y_1 ، y_2 ، ... ، y_n باید در این دستگاه همگن، با همان ماتریس ضریب‌ها، صدق کنند:

$$a_{11}y_1+\cdots+a_{1n}y_n=0 \quad , \qquad (9)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{m1}y_1+\cdots+a_{mn}y_n=0$$

برعکس، اگر به جواب نخستینِ x_1°، x_2°، ... ، x_n° از دستگاه (۸)، هر جوابی از دستگاه همگنِ (۹) را اضافه کنیم، آن دوباره جوابی از دستگاه (۸) به‌دست می‌آید.

بنابراین، برای به‌دست آوردن جواب کلی دستگاه (۸)، باید یکی از جواب‌های جداگانهٔ آن را در نظر گرفت و جواب کلی دستگاه همگن (۹) را به آن افزود.

بنابراین، مسألهٔ مربوط به خصلت مجموعهٔ جواب‌های دستگاه (۸)، منجر به همین پرسش برای دستگاه همگن (۹) می‌شود. این مسأله را در این‌جا بررسی خواهیم کرد.

دستگاه‌های همگن. دستگاه همگنِ معادله‌های خطیِ

$$a_{11}y_1+a_{12}y_2+\cdots+a_{1n}y_n=0 \quad ,$$
$$a_{21}y_1+a_{22}y_2+\cdots+a_{2n}y_n=0 \quad ,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{m1}y_1+a_{m2}y_2+\cdots+a_{mn}y_n=0$$

را در فضای nبُعدی اقلیدسی تفسیر می‌کنیم (ضریب‌های دستگاه را حقیقی در نظر می‌گیریم. برای دستگاه با ضریب مختلط، می‌توان شبیه همین تفسیر را در فضای یکانی داد و درضمن، به همین نتیجه رسید).

A'_1، A'_2، ... ، A'_m، Y را بردارهایی از فضای اقلیدسی می‌گیریم که مختصات آنها در یک پایهٔ یکا متعامد، به‌ترتیب،

$$(a_{11}, a_{12}, \ldots, a_{1n}), (a_{21}, a_{22}, \ldots, a_{2n}), \ldots, (a_{m1}, a_{m2}, \ldots, a_{mn}),$$
$$(y_1, y_2, \ldots, y_n)$$

باشد. در این صورت، دستگاه چنین می‌شود:

$$A'_1Y=0 \quad , \quad A'_2Y=0 \quad , \ldots , \quad A'_mY=0$$

یعنی، هر جواب دستگاه، برداری را معین می‌کند که بر همهٔ بردارهای مُعرف ضریب‌های

معادله‌های جداگانه عمود باشد.

بنابراین، مجموعهٔ جواب‌ها زیرفضایی را تشکیل می‌دهند، که متمم قائم زیرفضایی است که از بردارهای A'_1، A'_2، ...، A'_m پدید آمده‌اند. بعد زیرفضای اخیر برابر رتبهٔ r ماتریسی است که از ضریب‌های دستگاه تشکیل شده است. بُعد متمم قائم، یعنی «فضای جواب‌ها» هم برابر $(n-r)$ می‌شود.

هر زیرفضا پایه‌ای دارد، یعنی دستگاه بردارهای مستقل خطی که تعداد آنها برابر با بُعد زیرفضاست، و ترکیب خطی آنها زیرفضا را پر می‌کند. به‌این‌ترتیب، بین جواب‌های دستگاه همگن، $n-r$ جواب خطی مستقل وجود دارد، به‌نحوی که همهٔ جواب‌های دستگاه، ترکیبی خطی از آنها هستند. در این‌جا، n به‌معنای تعداد مجهول‌ها و r به‌معنای رتبهٔ ماتریس ضریب‌هاست.

به این ترتیب، روش ساختن جواب‌های دستگاه همگن و درنتیجه، دستگاه ناهمگن، به‌طور کامل روشن شد. به‌ویژه، وقتی و تنها وقتی جواب منحصر و «بدیهی» $x_1=x_2=\cdots=x_n=0$ را می‌پذیرد که رتبهٔ ماتریس ضریب‌ها برابر با تعداد مجهول‌ها باشد. درضمن، برای دستگاه ناهمگن هم، همین شرط (وقتی که دستگاه شرط‌های سازگاری را داشته باشد)، برای داشتن یک جواب منحصر وجود دارد.

آن‌چه دربارهٔ بررسی دستگاه‌ها آوردیم، به‌خوبی روشن می‌کند، چگونه استفاده از مفهوم‌های تعمیم‌یافتهٔ هندسی، کار را در مسأله‌های دشوار جبری، عینی و ساده می‌کند.

۴. تبدیل‌های خطی

تعریف و مثال‌ها. در بسیاری از بررسی‌های ریاضی، لازم می‌شود، متغیرها را عوض کنیم، یعنی از دستگاه متغیرهای x_1، x_2، ...، x_n، خود را به متغیرهای دیگری مثل y_1، y_2، ...، y_n برسانیم که به‌وسیلهٔ بستگی‌های تابعی، به متغیرهای اصلی مربوط‌اند:

$$\begin{gathered}y_1=\varphi_1(x_1,x_2,\dots,x_n)\\ y_2=\varphi_2(x_1,x_2,\dots,x_n)\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ y_n=\varphi_n(x_1,x_2,\dots,x_n)\end{gathered}$$

به‌عنوان مثال، اگر متغیرها به‌معنای مختصات نقطه‌ها در صفحه یا در فضا باشند، برای گذر

از یک دستگاه مختصات به دستگاه مختصات دیگر، باید از تبدیل مختصات استفاده کنیم که، به‌نوبهٔ خود، نیاز به وجود رابطه‌هایی دارد که مختصات اولیه را برحسب مختصات تازه و برعکس، بیان کند.

به‌جز این، در ریاضیات با موضوع‌هایی سر و کار داریم که موقعیت یا حالت چیزی را، که وابسته به چند متغیر است، شرح می‌دهند و، بنابراین با تغییر موقعیت یا حالت آن، به‌ناچار منجر به تبدیل متغیرها می‌شود. نمونهٔ مشخص این نوع تبدیل را، می‌توان در تغییر مختصات نقطه‌های یک جسم، ضمن تغییرشکل آن، دید.

یک تبدیل مفروض دستگاه شامل n متغیر را به صورت تبدیلی (یا تغییرشکلی) از فضای n بُعدی درنظر می‌گیریم، یعنی هر بردار فضا (یا بخشی از آن) را که دارای مختصات x_1، x_2، ... ، x_n است با برداری از فضا به مختصات y_1، y_2، ... ، y_n متناظر می‌کنیم.

پیش از این گفتیم، هر تابع چند متغیرهٔ «هموار» (که مشتق‌های جزئی پیوسته دارد)، وقتی این متغیرها تغییر کوچکی داشته باشند، به‌تابع خطی نزدیک است. بنابراین، هر تبدیل «هموار» (یعنی تبدیلی که در بیان تحلیلی آن، تابع‌های φ_1، φ_2، ... ، φ_n، مشتق‌های جزئی پیوسته‌ای داشته باشند)، در بخش کوچکی از فضا، به این تابع‌های خطی نزدیک است:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 \quad , \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 \quad , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n \end{aligned} \tag{10}$$

همین وضع، بررسی تبدیل‌های خطی را به یکی از مهم‌ترین مسأله‌های ریاضیات تبدیل می‌کند. به‌عنوان نمونه، با توجه به نظریهٔ n معادلهٔ خطی n مجهولی می‌دانیم، شرط لازم و کافی برای یکتا بودن جواب دستگاه معادله‌های (۱۰) نسبت به x_1، x_2، ... ، x_n، یعنی وارون‌پذیر بودن تبدیل خطی متناظر، این است که، دترمینان ضریب‌ها مخالف صفر باشد. این موقعیت، بنیان اصلی قضیهٔ بسیار مهمی از آنالیز را تشکیل می‌دهد: برای این که تبدیلِ

$$\begin{aligned} y_1 &= \varphi_1(x_1, x_2, \ldots, x_n) \\ y_2 &= \varphi_2(x_1, x_2, \ldots, x_n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= \varphi_n(x_1, x_2, \ldots, x_n) \end{aligned}$$

که در همسایگی نقطهٔ مفروض هموار است، تبدیل وارونِ هموار داشته باشد، لازم و کافی

است، در این نقطه، دترمینان

$$\begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_1} & \cdots & \frac{\partial \varphi_2}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial \varphi_n}{\partial x_1} & \cdots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix}$$

مخالف صفر باشد.

بررسی تبدیل خطی کلی (۱۰)، سر آخر منجر به بررسی تبدیل همگن، که همان ضریب‌ها را دارد، می‌شود:

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \quad , \\ y_2 &= a_{21}x_1 + \cdots + a_{2n}x_n \quad , \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned} \tag{۱۱}$$

از این به‌بعد، هرجا از تبدیل خطی نام می‌بریم، منظورمان تبدیل همگن است.

تبدیل خطی فضای nبُعدی را می‌توان، بدون در نظر گرفتن دستورهای (۱۱)، یعنی بدون استفاده از مختصات نقطه‌های متناظر، و با تعریف ویژگی‌های ذاتی تبدیل معین کرد. تعریف مفهوم تبدیل خطی، بدون استفاده از مختصات، از این جهت سودمند است که بستگی به‌انتخاب پایه ندارد. این تعریف را می‌توان این‌طور تنظیم کرد:

تبدیل خطی فضای خطی nبُعدی، با تابع $Y = A(X)$ تعریف می‌شود که در آن، متغیر X و مقدار Y، بردارها هستند. این تابع، با شرط خطی بودن، سازگار است:

$$A(c_1X_1 + c_2X_2) = c_1A(X_1) + c_2A(X_2) \tag{۱۲}$$

از این به‌بعد، وقتی از تبدیل خطی فضا صحبت می‌کنیم، به‌ویژه آن را به‌مفهوم همین تعریف می‌گیریم.

این تعریف، با تعریف مختصاتیِ قبل، هم‌ارز است. درواقع، اگر تابع $Y = A(X)$، بردار X با مختصات x_1، x_2، ... ، x_n را با بردار Y به مختصات y_1، y_2، ... ، y_n متناظر کند، به‌نحوی

که y_1، y_2، ...، y_n برحسب x_1، x_2، ...، x_n به‌صورت تابع‌های خطی همگن بیان شوند، آنگاه این تابع‌ها به‌روشنی با (۱۲) سازگارند. برعکس، اگر تابع $Y=A(X)$ با (۱۲) سازگار و e_1، e_2، ...، e_n پایه‌ای از فضا باشد، آنگاه

$$A(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1A(e_1) + x_2A(e_2) + \cdots + x_nA(e_n)$$

مختصات بردار $A(e_j)$ (در همین پایه) را با a_{1j}، ...، a_{nj} $(j = 1,\cdots,n)$ نشان می‌دهیم؛ در این صورت، مختصات بردار $Y=A(X)$ چنین خواهد بود:

$$\begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \quad , \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \quad , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{array}$$

بنابراین، هر تبدیل خطی از فضای خطی، نسبت به پایهٔ مفروض، متناظر با یک ماتریس مربعی است. خودِ تبدیل به زبان ماتریس‌ها و به‌صورت $Y=AX$ نوشته می‌شود. در این‌جا، X مُعرف ستون مختصاتِ بردار اولیه، Y مُعرف ستونِ مختصات بردار تبدیل‌شده و A مُعرف ماتریس ضریب‌های تبدیل است. ستون‌های ماتریس A، مُعرف مختصات بردارهایی هستند که بردارهای پایه به آنها تبدیل می‌شوند. در رابطه با بیان ماتریسی، از این به بعد، خود تبدیل خطی را هم، اغلب به‌صورت $Y=AX$ و با حذف پرانتز نشان می‌دهیم.

از دستور

$$A(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1A(e_1) + x_2A(e_2) + \cdots + x_nA(e_n)$$

نتیجه می‌شود، تمامی فضا، بعد از تبدیل خطی، به‌صورت زیرفضایی درمی‌آید که از بردارهای $A(e_1)$، ...، $A(e_n)$ پدیدآمده است. بُعد این زیرفضا، برابر است با رتبهٔ دستگاه بردارهای $A(e_1)$، $A(e_2)$، ...، $A(e_n)$ و یا هم‌ارز آن، رتبهٔ ماتریسی که شامل مختصات آن است، یعنی رتبهٔ ماتریس A. این زیرفضا، وقتی و تنها وقتی بر تمامی فضا منطبق می‌شود که رتبهٔ ماتریس برابر n، یعنی وقتی که دترمینان ماتریس A، مخالف صفر باشد. در این حالت، تبدیل را ***ناتکین*** می‌نامند.

با توجه به نظریهٔ دستگاه معادله‌های خطی می‌دانیم، تبدیل ناتکین، به‌صورتی یکتا وارون‌پذیر است و مختصات بردار اولیه را می‌توان برحسب مختصات بردار تبدیل‌شده،

به‌وسیلهٔ دستور $X=A^{-1}Y$ بیان کرد.

تبدیلی که ماتریس آن، دترمینانی برابر صفر دارد، ***تبدیل تکین*** نامیده می‌شود. تبدیل تکین، وارون‌پذیر نیست. این مطلب، نتیجه‌ای است از نظریهٔ معادله‌های خطی، یا بهتر و عینی‌تر از آن، از این که، این‌گونه تبدیل، تمامی فضا را به بخشی از آن تبدیل می‌کند.

به‌عنوان نمونهٔ تبدیل ناتکین، می‌توان پیش از همه، از ***تبدیل همانی*** نام برد که هر برداری را به خودش تبدیل می‌کند. ماتریس تبدیل همانی، در هر پایه‌ای، ماتریس همانی E است.

تبدیل تشابهی هم، یک تبدیل ناتکین است که هر برداری از فضا را در یک عدد (که برای همهٔ بردارها یکی است) ضرب می‌کند. ماتریس تبدیل تشابهی، به انتخاب پایه بستگی ندارد و به‌صورت aE است که در آن، a ضریب تشابه است.

حالت خاص و مهمی از تبدیل ناتکین، تبدیل‌های متعامد است. مفهوم ***تبدیل متعامد*** در فضای اقلیدسی معنا دارد و به‌عنوان یک تبدیل خطی در نظر گرفته می‌شود که طول بردارها را ثابت نگه دارد. تبدیل متعامد (اُرتُوگُونال) تعمیمی از دوران فضا، ضمن ثابت بودن مبداء مختصات، در فضای n بُعدی، یا ترکیبی از یک دوران و یک بازتاب (یا انعکاس) در صفحه‌ای که از مبداء می‌گذرد، است.

به‌سادگی دیده می‌شود که در تبدیل متعامد، نه‌تنها طول بردارها، بلکه حاصل‌ضرب‌های اسکالر هم حفظ می‌شود و بنابراین، تبدیل متعامد، یک پایهٔ یکا متعامد فضا را به دستگاهی از بردارهای واحد دوبه‌دو عمود بر هم تبدیل می‌کند که به‌نوبهٔ خود و به‌ناگزیر، پایه‌ای یکا متعامد است.

ماتریس مربوط به تبدیل متعامد، نسبت به یک پایهٔ یکا متعامد، دارای ویژگی‌های خاصی است که در این‌جا می‌آوریم.

اول؛ مجموع مربع‌های عضوهای هر ستون برابر واحد است، زیرا این مجموع، عبارت است از مربع طول بردارهایی که بردارهای پایهٔ انتخابی به آنها شده‌اند. دوم؛ مجموع حاصل‌ضرب‌های عضوهای متناظری که از دو ستون مختلف گرفته شده‌اند، برابر صفر است، زیرا چنین مجموع‌هایی، حاصل‌ضرب‌های اسکالر بردارهایی هستند که بردارهای پایه به آنها تبدیل شده‌اند.

با استفاده از ماتریس‌ها، می‌توانیم، این دو ویژگی را با دستور زیر نشان دهیم:

$$\overline{P}P=E$$

که در آن، P ماتریس تبدیل متعامد (و البته نسبت به یک پایهٔ یکا متعامد) و $\overline{P}$ ماتریس

ترانهادهٔ آن، یعنی ماتریسی که سطرهای آن، ستون‌های ماتریس P، با حفظ ردیف آنهاست.

درواقع، درایه‌های قطری ماتریس $\overline{P}P$، بنا بر قانون ضرب ماتریس‌ها، برابر است با مجموع مربع‌های درایه‌های ستون متناظر P، و درایه‌های غیرقطری برابر است با مجموع حاصل‌ضرب‌های درایه‌های متناظری که از ستون‌های P انتخاب شده است.

به‌عنوان نمونه‌ای از تبدیل تکین، می‌توان تصویر قائم همهٔ بردارهای فضای اقلیدسی را بر یک زیرفضا در نظر گرفت (بند ۲ را ببینید). درواقع، به‌ازای این تبدیل، نگاشتِ تمامی فضا بر بخشی از خودش به‌دست می‌آید.

تبدیل مختصات. اکنون به‌مسألهٔ مربوط به‌تبدیل مختصات در فضای nبُعدی می‌پردازیم؛ یعنی ضمن عبور از یک پایه به پایهٔ دیگر، مختصات بردارها چگونه تغییر می‌کند.

پایهٔ نخستین را e_1، e_2، ...، e_n می‌گیریم و فرض می‌کنیم f_1، f_2، ...، f_n، پایهٔ دیگری از فضا باشد. سپس، فرض می‌کنیم

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}$$

ماتریسی است که ستون‌های آن مختصات بردارهای پایهٔ جدید f_1، f_2، ...، f_n نسبت به پایهٔ اولیه باشد. روشن است ماتریس C ناتکین است، زیرا که بردارهای f_1، f_2، ...، f_n، مستقل خطی هستند. ماتریس C ***ماتریس تبدیل مختصات*** نامیده می‌شود.

مختصات برداری مثل X را در پایهٔ e_1، e_2، ...، e_n با x_1، x_2، ...، x_n؛ و مختصات همین بردار را نسبت به پایهٔ f_1، f_2، ...، f_n با x'_1، x'_2، ...، x'_n نشان می‌دهیم. در این صورت

$$X = x'_1 f_1 + x'_2 f_2 + \cdots + x'_n f_n$$

و بنابراین، بردار X نسبت به پایهٔ اولیه، ستونی را تشکیل می‌دهند:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_{11}x'_1 + c_{12}x'_2 + \cdots + c_{1n}x'_n \\ c_{21}x'_1 + c_{22}x'_2 + \cdots + c_{2n}x'_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{n1}x'_1 + c_{n2}x'_2 + \cdots + c_{nn}x'_n \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ c_{21} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

به این ترتیب، مختصات اولیه برحسب مختصات تبدیل‌شده، به‌صورت خطی و همگن با استفاده از ماتریس C بیان می‌شود.

دستورهایی که بستگی بین مختصات را نسبت به پایهٔ اولیه و پایهٔ تبدیل‌شده بیان می‌کنند، از نظر صوری، با دستورهایی که مختصات بردارهای متناظر را ضمن تبدیل خطی ناتکین فضا به هم مربوط می‌کنند، تطبیق می‌کند. این وضع به ما امکان می‌دهد، تبدیل خطیِ همگن متغیرهای مفروض از ماتریس ناتکین را یا به‌عنوان تبدیل مختصات و یا به‌عنوان تبدیل خطی فضا، تفسیر کنیم. در هر حالت مشخص، انتخاب یکی از این دو تفسیر، به‌یاری مضمون مسأله‌ای که بررسی می‌کنیم، روشن می‌شود.

اکنون به این موضوع می‌پردازیم که، ضمن تبدیل مختصات، ماتریس تبدیل خطی فضا، چگونه تغییر می‌کند.

فرض کنیم نسبت به پایهٔ e_1، e_2، ...، e_n، تبدیل خطی مفروض، ماتریس A را داشته باشد، به‌نحوی که ستون Y از مختصات بردار تبدیل‌شده، با ستون X بردار اولیه، با این دستور مربوط باشد:

$$Y = AX$$

اکنون فرض کنید، تبدیل مختصات با ماتریس C انجام شده باشد. X' و Y' را به‌معنای ستون‌های مختصات اولیه و بردارهای تبدیل، نسبت به پایهٔ جدید می‌گیریم. در این صورت

$$X = CX' \quad , \quad Y = CY'$$

از آن‌جا

$$Y' = C^{-1}Y = C^{-1}AX = C^{-1}ACX'$$

به این ترتیب، ماتریس تبدیل مفروض، نسبت به پایهٔ جدید عبارت است از $C^{-1}AC$.

ماتریس‌های A و B که با رابطهٔ $B = C^{-1}AC$ به هم مربوط باشند (که در آن، C ماتریسی ناتکین است)، **متشابه** نامیده می‌شوند. اگر دسته‌ای از ماتریس‌های دوبه‌دو متشابه داشته باشیم، تنها یک تبدیل خطی نسبت به پایه‌های مختلف با آنها متناظر است.

ویژه‌بردارها و ویژه‌مقدارهای تبدیل خطی. مهم‌ترین گروه از تبدیل‌های خطی، تبدیل‌هایی هستند که به‌صورتی که در زیر می‌آوریم، تحقق پیدا کنند.

e_1، e_2، ...، e_n را بردارهایی از فضا می‌گیریم که نسبت به هم استقلال خطی داشته باشند. فرض کنید، این بردارها ضمن تبدیل، در عددهایی مثل λ_1، λ_2، ...، λ_n ضرب شوند. اگر بردارهای e_1، e_2، ...، e_n را به‌عنوان پایهٔ فضا در نظر بگیریم، آن‌وقت تبدیل را به‌وسیلهٔ

این ماتریس قطری نشان می‌دهند:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

تبدیل این دسته از بردارها، مفهوم هندسی عینی و ساده‌ای دارند (البته در فضای حقیقی و برای $n = 2$ یا $n = 3$). به‌ویژه، اگر همهٔ عددهای λ_i مثبت باشند، آن‌وقت این تبدیل شاهد انبساط (یا انقباض) فضا در جهت بردارهای e_1، e_2، ... ، e_n با ضریب‌های λ_1، λ_2، ... ، λ_n است. اگر برخی از λ_iها منفی باشند، آن‌وقت تغییرشکل فضا، با تغییرجهت برخی از بردارهای e_1، e_2، ... ، e_n در جهت مخالف خود، همراه است. سرانجام، اگر در مثل $\lambda_1 = 0$، آن‌وقت فضا در جهت موازی با e_1 و از زیرفضایی که از e_2، ... ، e_n پدید آمده است تصویر می‌شود.

این‌گونه تبدیل‌ها، از این جهت مهم‌اند که با وجود سادگی خود، بی‌اندازه کلی‌اند. ثابت می‌شود، هر تبدیل خطی که با شرط‌هایی نه‌چندان سخت محدود شده باشد، به این دسته از تبدیل‌ها تعلق دارد، یعنی برای آن می‌توان پایه‌ای پیدا کرد که در آن، با ماتریس قطری شرح داده شود.

اگر به تبدیل‌های خطی فضای مختلط توجه داشته باشیم، آن‌وقت محدودیت‌هایی که باید برای فضا در نظر بگیریم، آزادتر و اختیاری‌تر می‌شود. و ما، از این‌جا به بعد، همین فرض را انتخاب می‌کنیم.

این تعریف را می‌آوریم.

بردار غیرصفر X، که ضمن تبدیل خطی A از فضا، به بردار هم‌راستای λX منجر شود، ***ویژه‌بردار تبدیل*** نامیده می‌شود. به زبان دیگر، بردار غیرصفر X، وقتی و تنها وقتی ویژه‌بردار تبدیل A است که داشته باشیم:

$$AX = \lambda X$$

عدد λ را هم، ***ویژه‌مقدار*** تبدیل A گویند.

روشن است، اگر تبدیل در پایه‌ای، به‌صورت ماتریس قطری بیان شود، آن‌وقت این پایه، از ویژه‌بردارها تشکیل شده است، و درایه‌های قطر، ویژه‌مقدارهای تبدیل‌اند. برعکس، اگر

در فضا پایه‌ای وجود داشته باشد که از ویژه‌بردارهای تبدیل A تشکیل شده باشد، آن‌وقت در این پایه، ماتریس نمایندهٔ A یک ماتریس قطری است که شامل ویژه‌مقدارهای متناظر با بردارهای پایه است.

به بررسی ویژگی‌های ویژه‌بردارها و ویژه‌مقدارها می‌پردازیم. برای این منظور، بیان مختصاتیِ تعریف ویژه‌بردار را می‌آوریم. فرض می‌کنیم A، ماتریس مربوط به تبدیل A نسبت به یک پایه و X ستون مختصات بردار X در همان پایه باشد. از برابری $AX=\lambda X$ نتیجه می‌گیریم

$$(A-\lambda E)X=0$$

که با بازکردن آن، به این دستگاه می‌رسیم:

$$\begin{aligned}
&(a_{11}-\lambda)x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\ ,\\
&a_{21}x_1+(a_{22}-\lambda)x_2+\cdots+a_{2n}x_n=0\ ,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{n1}x_1+a_{n2}x_2+\cdots+(a_{nn}-\lambda)x_n=0
\end{aligned}$$

این دستگاه برابری‌ها را، می‌توان همچون دستگاه معادله‌های خطی همگن نسبت به x_1، x_2، ... ، x_n در نظر گرفت. ما به حالتی نیاز داریم که دستگاه جوابی غیرصفر داشته باشد، زیرا همهٔ مختصات ویژه‌بردار نباید برابر صفر باشند. می‌دانیم شرط لازم و کافی، برای وجود جواب «نابدیهی» در دستگاه معادله‌های خطی همگن این است که رتبهٔ ماتریس ضریب‌ها، کمتر از تعداد مجهول‌ها باشد؛ و این، هم‌ارز با آن است که دترمینان دستگاه برابر صفر شود:

$$\begin{vmatrix}
a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\
a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\
\cdots & \cdots & \cdots & \cdots\\
a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda
\end{vmatrix}=0$$

به این ترتیب، همهٔ ویژه‌مقدارهای تبدیل A، ریشه‌های چندجمله‌ای $|A-\lambda E|$ هستند و برعکس، هر ریشهٔ این چندجمله‌ای، ویژه‌مقداری از تبدیل است، زیرا هر ریشه دست‌کم با یک ویژه‌بردار متناظر است. چندجمله‌ای $|A-\lambda E|$ را، ***چندجمله‌ای مشخصه*** برای

ماتریس A گویند. معادلهٔ $|A-\lambda E| = 0$ را هم، **معادلهٔ مشخصه** ماتریس A و ریشه‌های آن را، **عددهای مشخصه** ماتریس گویند.

با توجه به قضیهٔ اصلی جبر (بخش چهارم، جلد اول)، هر چندجمله‌ای دست‌کم یک ریشه دارد؛ بنابراین هر تبدیل خطی، دست‌کم دارای یک ویژه‌مقدار یعنی دست‌کم دارای یک ویژه‌بردار است. ولی البته، ممکن است حالتی پیش آید که تبدیل با ماتریسی حقیقی بیان شود و همه یا بخشی از ویژه مقدارهای آن، مختلط باشد. در واقع در فضای حقیقی، قضیهٔ مربوط به وجود ویژه مقدارها (حقیقی) و ویژه بردارها برای هر تبدیل خطی درست نیست. برای نمونه، اگر تبدیل را، دوران دور مبداء مختصات به‌اندازهٔ زاویه‌ای غیر از ۱۸۰ درجه در نظر بگیریم، جهت همهٔ بردارهای صفحه را تغییر می‌دهد، به‌نحوی که ویژه‌بردارها برای این تبدیل وجود ندارد.

ریشه‌های چندجمله‌ای مشخصه ماتریس A، ویژه‌مقدارهای تبدیل A هستند و بنابراین، ماتریس‌هایی که متناظر با یک تبدیل در پایه‌های مختلف‌اند مجموعه‌های یکسانی از ریشه‌ها دارند. و این، شبیه این پیش‌بینی است که چندجمله‌ای مشخصه تبدیل خطی، به خود تبدیل خطی مربوط می‌شود نه به انتخاب پایه. این حکم را می‌توان با برآورد ظریفی براساس ویژگی‌های عمل در ماتریس‌ها و دترمینان‌ها تحقیق کرد.

می‌دانیم، اگر ماتریس A متناظر با تبدیل A در یک پایه باشد، آن‌وقت در پایهٔ دیگر، تبدیل A، دارای ماتریس $C^{-1}AC$ است که در آن، C یک ماتریس ناتکین است. ولی

$$|C^{-1}AC-\lambda E| = |C^{-1}AC - C^{-1}\lambda EC| = |C^{-1}(A-\lambda E)C| =$$
$$= |C^{-1}|\,|C|\,|A-\lambda E| = |C^{-1}C|\,|A-\lambda E| = |A-\lambda E|$$

به این ترتیب، ماتریس متناظر با یک تبدیل A در پایه‌های مختلف، در واقع دارای یک چندجمله‌ای مشخصه هستند که، به همین دلیل، می‌توان آن را چندجمله‌ای مشخصه تبدیل نامید.

اکنون فرض می‌کنیم، همهٔ ویژه‌مقدارهای تبدیل A مختلف باشند. ثابت می‌کنیم، ویژه‌بردارهایی که، هرکدام برای یک ویژه‌مقدار در نظر گرفته شده‌اند، نسبت به هم استقلال خطی دارند. در واقع، اگر فرض کنیم برخی از این بردارها، مثل e_1، e_2، ...، e_k استقلال خطی داشته باشند، ولی بقیه، و از جمله e_{k+1}، ترکیبی خطی از آنها باشد، آن‌وقت

$$e_{k+1} = c_1 e_1 + c_2 e_2 + \cdots + c_k e_k \tag{13}$$

اگر در هر دو طرف برابری، تبدیل خطی را به‌کار بریم، به‌دست می‌آید:

$$Ae_{k+1} = c_1 Ae_1 + c_2 Ae_2 + \cdots + c_k Ae_k$$

از این‌جا و با توجه به تعریف ویژه‌بردار نتیجه می‌شود:

$$\lambda_{k+1} e_{k+1} = c_1 \lambda_1 e_1 + c_2 \lambda_2 e_2 + \cdots + c_k \lambda_k e_k$$

اگر برابری (۱۳) را در λ_{k+1} ضرب کنیم و سپس از برابری بالا کم کنیم، به‌دست می‌آید:

$$c_1 (\lambda_{k+1} - \lambda_1)\, e_1 + c_2 (\lambda_{k+1} - \lambda_2) e_2 + \cdots + c_k (\lambda_{k+1} - \lambda_k) e_k = 0$$

از این‌جا و با توجه به استقلال خطی e_1، e_2، ... ، e_k نتیجه می‌شود:

$$c_1 (\lambda_{k+1} - \lambda_1) = c_2 (\lambda_{k+1} - \lambda_2) = \cdots = c_k (\lambda_{k+1} - \lambda_k) = 0$$

ولی فرض بر این بود که همهٔ ویژه‌مقدارها باهم فرق دارند و هر بردار برای یک ویژه‌مقدار مختلف در نظر گرفته شده است. بنابراین

$$\lambda_{k+1} - \lambda_1 \neq 0 \,,\ \lambda_{k+1} - \lambda_2 \neq 0 \,,\ \ldots ,\ \lambda_{k+1} - \lambda_k \neq 0$$

و برابری (۱۳) برقرار نیست، زیرا ضریب‌های c_1، c_2، ... ، c_k نمی‌توانند باهم و به‌طور همزمان برابر صفر باشند.

حال دیگر روشن است، اگر همهٔ ویژه‌مقدارهای تبدیل خطی مختلف باشند، آن‌وقت پایه‌ای وجود دارد که در آن، ماتریس تبدیل، شکلی قطری داشته باشد. درواقع، به‌عنوان چنین پایه‌ای می‌توان دستگاه ویژه‌بردارها را طوری انتخاب کرد که هرکدام از آنها، متناظر با یک ویژه‌مقدار باشد. همان‌طور که ثابت کردیم، این بردارها استقلال خطی دارند و تعداد آنها برابر است با بُعد فضا، که به معنی این است که این بردارها یک پایه را تشکیل می‌دهند.

قضیه‌ای را که ثابت کردیم، می‌توان با استفاده از اصطلاح‌های نظریه ماتریس‌ها این‌طور تنظیم کرد: اگر همهٔ ویژه‌مقدارهای ماتریس متفاوت باشند، آن‌وقت ماتریس

با یک ماتریس قطری، که درایه‌های قطری آن همین ویژه‌مقدارها است، متشابه است.

مسألهٔ مربوط به تبدیل ماتریس تبدیل خطی به ساده‌ترین صورت خود، در حالتی که برای چندجمله‌ای مشخصه، ریشه‌های برابر وجود دارد، مسأله‌ای دشوار است. در این‌جا تنها به شرح کوتاهی دربارهٔ نتیجهٔ پایانی می‌پردازیم.

به ماتریس مرتبهٔ m

$$I_{m,\lambda_i} = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

را «بلوک متعارف» می‌نامند که در آن، درایه‌هایی که نوشته نشده است، برابر صفرند.

«ماتریس متعارف ژُردان، به ماتریسی گفته می‌شود که درایه‌های قطر اصلی آن را «بلوک‌های متعارف» تشکیل می‌دهند و بقیهٔ درایه‌های آن برابر صفرند، یعنی ماتریسِ

$$\begin{pmatrix} \boxed{I_{m_1,\lambda_1}} & & & \\ & \boxed{I_{m_2,\lambda_2}} & & \\ & & \ddots & \\ & & & \boxed{I_{m_k,\lambda_k}} \end{pmatrix}$$

عددهای λ_i در «بلوک‌های» مختلف، می‌توانند دوبه‌دو مختلف نباشند. هر ماتریس می‌تواند تحویل به ماتریس متعارف ژُردانِ متناظر خود، که با آن متشابه است، بشود. اثبات این قضیه، بسیار پیچیده است. یادآوری می‌کنیم، این قضیه، در بسیاری از حالت‌های کاربرد جبر در دیگر مسأله‌های ریاضیات، و به‌ویژه در نظریهٔ دستگاه‌های معادله‌های خطی دیفرانسیلی، نقشِ جدی دارد.

ماتریس، تنها وقتی به‌صورت قطری درمی‌آید که مرتبهٔ m_i همهٔ «بلوک‌ها» برابر واحد باشد.

۵. صورت‌های درجه دوم

تعریف و ویژگی‌های ساده. چندجمله‌ای همگن درجه‌دوم با چند متغیر را، صورت درجهٔ دوم گویند.

صورت درجه دوم با n متغیر x_1، x_2، ... ، x_n، شامل دو نوع جمله است: جمله‌هایی که مجذور متغیرها هستند، و جمله‌هایی که از حاصل‌ضرب دوبه‌دوی متغیرها به‌دست می‌آیند؛ البته، هر جمله دارای ضریبی است. صورت درجه دوم را به‌این صورت می‌نویسند:

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \\ & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \end{aligned}$$

هر دو جملهٔ متشابه $a_{12}x_1x_2$ و $a_{21}x_2x_1$ وغیره با ضریب‌های یکسان نوشته می‌شود، به‌نحوی که هریک از آنها، ضریبی برابر نصف ضریب حاصل‌ضرب دو متغیر داشته باشد. به این ترتیب، هر صورت درجه دوم، به نحوی یکتا با ماتریس متقارن ضریب‌های آن متناظر است.

صورت درجه دوم را، خیلی ساده، می‌توان به‌صورت ماتریس نشان داد. X را ماتریس ستونی از متغیرهای x_1، x_2، ... ، x_n، و $\overline{X}$ را ماتریس سطری از همین متغیرها (یعنی ماتریس ترانهادهٔ X) می‌گیریم. در این صورت

$$\begin{aligned} f(x_1, x_2, \dots, x_n) = {} & x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) + \\ & + x_2(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n) + \cdots + x_n(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n) = \end{aligned}$$

$$= (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix} =$$

$$= (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \overline{X}AX$$

در بسیاری از شاخه‌های ریاضیات و کاربردهای آنها، با صورت‌های درجه دوم برخورد می‌کنیم.

در نظریهٔ عددها و در بلورشناسی، چنان صورت‌های درجه دوم بررسی می‌شود که در آنها، متغیرهای $x_۱$، $x_۲$، ... ، x_n تنها مقدارهای درست عددی را می‌پذیرند. در هندسهٔ تحلیلی، صورت درجه دوم به صورت معادلهٔ منحنی (یا سطح) درجه دوم ظاهر می‌شود. صورت‌های درجه دوم، در مکانیک و هم، برای بیان انرژی جنبشی دستگاه برحسب مؤلفه‌های سرعت تعمیم‌یافته وغیره، کاربرد دارد. به‌جز این، بررسی صورت‌های درجه دوم در آنالیز هم، برای مطالعهٔ تابع‌های با چند متغیر در مسأله‌هایی لازم است، که برای حل آنها باید روشن کرد، چگونه تابع مفروض در همسایگی نقطهٔ مفروض، از مقدار تقریبی تابع خطی آن منحرف می‌شود. نمونهٔ این‌گونه مسأله‌ها را می‌توان در بررسی ماکزیمم و می‌نیمم تابع‌ها مشاهده کرد.

برای نمونه، مسألهٔ مربوط به بررسی ماکزیمم و می‌نیمم را برای تابع با دو متغیر $w=f(x,y)$، که تا مرتبهٔ سوم، مشتق‌های جزئی پیوسته‌ای دارد، در نظر می‌گیریم. شرط لازم، برای این که نقطهٔ $(x_۰, y_۰)$ ماکزیمم یا می‌نیمم تابع w باشند، این است که مشتق‌های جزئی مرتبهٔ اول آن در نقطهٔ $(x_۰, y_۰)$ برابر صفر باشند. فرض می‌کنیم، این شرط برقرار باشد. برای متغیرهای x و y، نموهای کوچک h و k را در نظر می‌گیریم و نمو متناظر تابع، یعنی $\Delta w=f(x_۰+h\ , y_۰+k)-f(x_۰, y_۰)$ را بررسی می‌کنیم. با توجه به دستور تیلور، این نمو با دقت تا بی‌نهایت کوچک‌های مرتبهٔ بالاتر، برابر است با صورت درجه دوم

$$\frac{۱}{۲}\left(rh^۲ + ۲shk + tk^۲\right)$$

که در آن r، s و t عبارت‌اند از مقدارهای مشتق دوم $\frac{\partial^۲ w}{\partial x^۲}$، $\frac{\partial^۲ w}{\partial x \partial y}$ و $\frac{\partial^۲ w}{\partial y^۲}$ در نقطهٔ $(x_۰, y_۰)$. اگر این صورت درجه دوم، به‌ازای همهٔ مقدارهای h و k (به‌جز $h=k=۰$) مثبت باشد، آن‌وقت تابع w در نقطهٔ $(x_۰, y_۰)$ می‌نیمم دارد؛ و اگر این مقدارها منفی باشند، آن‌وقت تابع در نقطهٔ $(x_۰, y_۰)$ ماکزیمم دارد. سرانجام، اگر صورت درجه دوم، هم مقدارهای مثبت و هم مقدارهای منفی را قبول کند، آن‌وقت، تابع در این نقطه ماکزیمم یا می‌نیمم ندارد. تابعی را هم که تعداد بیشتری متغیر داشته باشد، می‌توان به همین صورت بررسی کرد.

بررسی صورت‌های درجه دوم، در اساس، مربوط به این است که مسألهٔ هم‌ارزی صورت‌ها را نسبت به مجموعه‌ای از تبدیل‌های خطیِ متغیرها بررسی کنیم. دو صورت درجه دوم را هم‌ارز گویند، وقتی بتوان یکی از آنها را به‌یاری تبدیلی از مجموعهٔ مفروض به دیگری منجر کرد. مسألهٔ تحویل یک صورت به صورت دیگر، یعنی تحویل آن به ساده‌ترین صورت ممکن، با مسألهٔ هم‌ارزی بستگیِ تنگاتنگ دارد.

در مسأله‌های مختلف مربوط به صورت‌های درجه دوم، مجموعه‌های مختلفی هم که امکان تبدیل متغیرها را به‌وجود می‌آورند، بررسی می‌شود.

در مسأله‌های آنالیز، هر تبدیل ناتکین متغیرها کاربرد دارد؛ در هندسهٔ تحلیلی، بیش از همه، به تبدیل‌های متعامد (اُرتوگونال)، یعنی تبدیل‌هایی که امکان عبور از یک دستگاه مختصات دکارتی متغیر را به دیگری پدید می‌آورد، توجه می‌شود. سرانجام، در نظریهٔ عددها و در بلورشناسی، تبدیل‌های خطی با ضریب‌های درست و دترمینان برابر واحد، بررسی می‌شود.

از این مسأله‌ها، دو مسأله را انتخاب می‌کنیم؛ مسألهٔ مربوط به رساندن صورت درجه دوم به ساده‌ترین صورت خود به یاری مجموعه‌ای از تبدیل‌ها؛ و همین مسأله دربارهٔ تبدیل‌های متعامد. قبل از همه، روشن کنیم، ماتریس صورت درجه دوم، ضمن تبدیل خطی متغیرها، به چه وضعی درمی‌آید!

فرض کنیم: $f(x_1, x_2, \dots, x_n) = \overline{X}AX$ که در آن، A ماتریس متقارنی از ضریب‌های فرم و X ماتریس ستونی از متغیرهاست.

تبدیل خطی متغیرها را که به‌صورت کوتاه $X = CX'$ نشان داده می‌شود، در نظر می‌گیریم. در این‌جا C به‌معنای ماتریس ضریب‌های این تبدیل، و X' به‌معنای ستونی از متغیرهای تازه است. در این صورت $\overline{X} = \overline{X}'\overline{C}$ و بنابراین

$$\overline{X}AX = \overline{X}'(\overline{C}AC)X'$$

به‌نحوی که ماتریس تبدیل صورت درجه دوم عبارت است از $\overline{C}AC$.

به‌سادگی می‌توان تحقیق کرد که ماتریس $\overline{C}AC$ به‌خودی خود متقارن از آب درمی‌آید. به این ترتیب، مسألهٔ مربوط به درآوردن صورت درجه دوم به ساده‌ترین صورت ممکن، هم‌ارز است با مسألهٔ مربوط به درآوردن ماتریس متقارن به ساده‌ترین صورت به‌کمک ضرب سمت راست و سمت چپ آن در ماتریس‌هایی که ترانهادهٔ هم هستند.

تبدیل صورت درجه دوم به‌صورت متعارف به‌وسیلهٔ جدا کردن متوالی مربع‌ها. ثابت می‌کنیم، هر صورت درجه دوم (حقیقی) را می‌توان به‌صورت مجموع مربع‌های متغیرهای تازه همراه با ضریب‌هایی به یاری تبدیل خطیِ ناتکین حقیقی نوشت.

برای اثبات، در آغاز ثابت می‌کنیم، اگر صورت متحد با صفر نباشد، آن‌وقت با تبدیل ناتکین متغیرها، می‌توان ضریب مربع متغیر اول را مخالف با صفر درآورد.

درواقع، فرض کنید:

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \\
& + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \\
& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2
\end{aligned}
$$

اگر $a_{11} \neq 0$، هیچ تبدیلی لازم نیست. اگر $a_{11} = 0$، ولی ضریبی از ضریب‌های قطری مخالف صفر باشد، برای نمونه $a_{kk} \neq 0$، آن‌وقت فرض می‌کنیم $x_1 = x'_k$، $x_k = x'_1$ و بقیهٔ متغیرهای اولیه را برابر با متغیرهای متناظر جدید می‌گیریم. این تبدیل ناتکین ما را به هدف می‌رساند. سرانجام، اگر همهٔ ضریب‌های قطری برابر صفر باشند، آن‌وقت دست‌کم یکی از ضریب‌های غیرقطری مخالف صفر است؛ این ضریب را a_{12} فرض می‌کنیم.

این تبدیل را انجام می‌دهیم:

$$
\begin{aligned}
x_1 &= x'_1 \\
x_2 &= x'_1 + x'_1
\end{aligned}
$$

و همهٔ بقیهٔ متغیرها را برابر متغیرهای تازه می‌گیریم؛ در این صورت به هدف می‌رسیم.

به این ترتیب، بدون این که به کلی بودن مطلب لطمه‌ای وارد بیاید، می‌توان فرض کرد: $a_{11} \neq 0$.

اکنون مربع تابع خطی را طوری جدا می‌کنیم که همهٔ جمله‌های شامل x_1 در این مربع وارد شده باشند.

این عمل را به‌سادگی می‌توان انجام داد. در واقع

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n + \\
& + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \\
& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 =
\end{aligned}
$$

$$= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - a_{11}\left(\frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 +$$

$$+ a_{22}x_2^2 + \cdots + a_{2n}x_2x_n +$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2$$

که اگر پرانتز دوم را باز و جمله‌های متشابه را جمع و جور کنیم، به‌دست می‌آید:

$$f(x_1, x_2, \ldots, x_n) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + f_1(x_2, \ldots, x_n)$$

که در آن f_1 صورتی است با $(n-1)$ متغیر.

تبدیلِ

$$x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n = x'_1 \quad ,$$

$$x_2 = x'_2 \quad ,$$

$$\cdots\cdots$$

$$x_n = x'_n$$

به‌روشنی ناتکین است. با انجام این تبدیل، به صورتی به این شکل می‌رسیم:

$$a_{11}x_1'^2 + f_1(x'_2, \ldots, x'_n)$$

اگر همین روش را ادامه دهیم، سرانجام به صورت لازم که صورتی «متعارف» است، می‌رسیم:

$$\alpha_1 z_1^2 + \alpha_2 z_2^2 + \cdots + \alpha_n z_n^2$$

که در آن، z_1، z_2، ...، z_n معرف دنبالهٔ متغیرهای انتخابیِ تازه است.

قانون ماند (اینرسی یا لَختی) در صورت درجه دوم. ضمن رساندن صورت درجه دوم به صورت متعارف، در انتخاب تبدیل متغیرها، آزادی بسیار وجود دارد. این آزادی دست‌کم در این است که پیش از به‌کار بردن روش جداکردن مجذورها، می‌توان هر تبدیل ناتکین متغیرها را انجام داد.

با این‌همه، با وجود این آزادی عمل، در نتیجهٔ نهایی به‌تقریب، صورت‌های درجه دوم متعارف یکسانی، بدون ارتباط با تبدیلی که داده شده است، به‌دست می‌آید. به‌ویژه، تعداد

مجذورهای متغیرهای جدید که ضریب‌هایی مثبت، منفی یا صفر دارند، با هر روش عمل ثابت می‌ماند. این قضیه را، ***قانون ماند*** در صورت‌های درجه دوم گویند. در این‌جا به اثبات این قضیه نمی‌پردازیم.

قانون ماند در صورت‌های درجه دوم، مسألهٔ مربوط به هم‌ارزی صورت‌های درجه دوم حقیقی را نسبت به همهٔ تبدیل‌های ناتکین، حل می‌کند. به‌ویژه، دو صورت وقتی و تنها وقتی هم‌ارزند که با رسیدن به صورت متعارف، صورت‌های متعارفی به‌دست آید که تعداد مجذورهای با ضریب‌های مثبت، منفی و صفر در آنها یکسان باشد.

صورت‌های درجه دوم که بعد از رسیدن به‌صورت متعارف به مجموعی از مجذورهای متغیرهای جدید تبدیل می‌شوند که در آنها، همهٔ ضریب‌ها مثبت‌اند، برای ما اهمیت دارند. این‌گونه صورت‌ها را، **معینِ مثبت** می‌نامند.

صورت‌های درجه دوم معین مثبت دارای این ویژگی هستند که هر مقداری از آنها به‌ازای مقدارهای حقیقی متغیرها، به‌شرطی که با هم و به‌طور هم‌زمان برابر صفر نباشند، مثبت است.

تبدیل متعامد صورت‌های درجه دوم به صورت متعارف. در میان روش‌های مختلفی که برای رساندن صورت درجه دوم به‌صورت متعارف وجود دارد، تبدیل متعامد جای خاصی دارد، یعنی به‌کمک تبدیل خطی متغیرهای با ماتریس متعامد. به‌ویژه، این‌گونه تبدیل‌ها، از جمله در هندسهٔ تحلیلی، در مسأله‌های مربوط به درآوردن معادلهٔ کلی منحنی یا سطح درجه دوم به‌صورت متعارف، بسیار جالب‌اند.

برای به‌انجام رساندن چنین تبدیلی، بهتر است صورت درجه دوم را به‌عنوان تابعی از بردار در فضای اقلیدسی در نظر بگیریم و متغیرهای x_1، x_2، ... ، x_n را، مختصات بردار متغیر نسبت به یک پایهٔ یکا متعامد به‌حساب آوریم. در این صورت، تبدیل متعامد متغیرها، همچون عبور از یک پایهٔ یکا متعامد به دیگری تفسیر می‌شود.

صورت درجه دوم

$$f(x_1, x_2, \dots, x_n) = a_{11}x_1^2 + \dots + a_{1n}x_1x_n + \\ \dots\dots\dots\dots\dots\dots \\ + a_{n1}x_nx_1 + \dots + a_{nn}x_n^2$$

در رابطه با تبدیل خطی A، که ماتریس

$$A=\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

را برای پایهٔ آن انتخاب کرده‌ایم، در نظر می‌گیریم. در این صورت، خودِ صورت درجه دوم می‌تواند همچون حاصل‌ضرب اسکالر $AX\cdot X$ در نظر گرفته شود (که در آن، X برداری است با مختصات x_1، x_2، ... ، x_n)؛ درضمن ضریب‌های a_{ij} به‌عنوان حاصل‌ضرب‌های اسکالر $Ae_i\cdot e_j$ باشد که در آن، e_1، e_2، ... ، e_n پایهٔ یکا متعامد انتخابی است.

به‌سادگی دیده می‌شود، با توجه به تقارن ماتریس A، برای هر دو بردار X و Y، این برابری برقرار است:

$$AX\cdot Y = X\cdot AY$$

در آغاز ثابت می‌کنیم، تبدیل A دست‌کم یک ویژه‌مقدار حقیقی و متناظر با آن یک ویژه‌بُردار دارد.

برای این منظور، مقدارهای صورت $AX\cdot X$ را با این فرض در نظر می‌گیریم که بردار X در کرهٔ واحد حرکت کند، یعنی مجموعهٔ همهٔ بردارهای واحد باشد. با این شرط صورت $AX\cdot X$ دارای ماکزیمم می‌شود. ثابت می‌کنیم، این ماکزیمم λ_1، ویژه‌مقداری از تبدیل A و بردار X_0، که به‌ازای آن، این ماکزیمم به‌دست می‌آید، ویژه‌بردار متناظر با آن است، یعنی $AX_0=\lambda_1 X_0$.

اثبات این گزاره را به‌طور غیرمستقیم می‌دهیم و ثابت می‌کنیم، بردار AX_0 بر همهٔ بردارهای عمود بر X_0 عمود است.

توجه کنیم، برای هر بردار Z، نابرابری

$$AZ\cdot Z \le \lambda_1\ |Z|^2$$

برقرار است. این حکم از این‌جا به‌دست می‌آید که $X=\dfrac{Z}{|Z|}$ بردار واحد است، و λ_1 ماکزیمم مقدارهای فرم $AX\cdot X$ در کرهٔ واحد است. $Z=X_0+\varepsilon Y$ را در نظر می‌گیریم که در آن، ε یک عدد حقیقی و Y بردار دل‌خواهی عمود بر بردار X_0 است. در این صورت

$$AZ\cdot Z=(AX_0+\varepsilon AY)\cdot(X_0+\varepsilon Y) = AX_0\cdot X_0+2\varepsilon AX_0\cdot Y + \varepsilon^2 AY\cdot Y=$$
$$=\lambda_1 + 2\varepsilon AX_0\cdot Y + \varepsilon^2 AY\cdot Y$$

به‌جز این

$$|Z|^2 = (X_0 + \varepsilon Y)\cdot(X_0 + \varepsilon Y) = |X_0|^2 + \varepsilon^2|Y|^2 = 1 + \varepsilon^2|Y|^2$$

زیرا

$$X_0 \cdot Y = 0 \quad , \quad |X_0|^2 = 1$$

بنابراین

$$\lambda_1 + 2\varepsilon AX_0 \cdot Y + \varepsilon^2 AY\cdot Y \le \lambda_1 + \varepsilon^2\lambda_1|Y|^2$$

از آن‌جا، با تقسیم بر ε^2، به‌دست می‌آید:

$$\frac{2}{\varepsilon}AX_0 \cdot Y \le \lambda_1|Y|^2 - AY\cdot Y \tag{14}$$

این نابرابری باید برای هر عدد حقیقی ε برقرار باشد، هرقدر که ε را از لحاظ قدرمطلق کوچک بگیریم.

ولی نابرابری، تنها با شرطِ $AX_0 \cdot Y = 0$ برقرار است، زیرا اگر $AX_0 \cdot Y > 0$، آن‌وقت نابرابری (۱۴) برای مقدارهای به‌اندازهٔ کافی کوچک و مثبت ε برقرار نیست؛ اگر هم $AX_0 \cdot Y < 0$، آن‌وقت نابرابری (۱۴)، برای مقدارهای به‌اندازهٔ کافی کوچک (از لحاظ قدرمطلق) و منفی ممکن نیست. به این ترتیب $AX_0 \cdot Y = 0$، یعنی AX_0 بر هر بردار عمود بر X_0، عمود است. بنابراین، بردارهای AX_0 و X_0 هم‌راستا هستند، یعنی $AX_0 = \lambda' X_0$ که در آن، λ' یک عدد حقیقی است. این که $\lambda' = \lambda_1$، به‌سادگی قابل تحقیق است:

$$\lambda_1 = AX_0 \cdot X_0 = \lambda' X_0 \cdot X_0 = \lambda'$$

اکنون می‌توان ثابت کرد، هر صورت درجه دوم به‌وسیلهٔ تبدیل متعامد، می‌تواند به‌صورت متعارف درآید.

e_1، e_2، ...، e_n را پایهٔ متعامد اولیه و f_1، f_2، ...، f_n را پایهٔ یکا متعامد جدید می‌گیریم که، درضمن، بردار f_1 ویژه‌بردار X_0 در تبدیل A است. فرض کنید x_1، x_2، ...، x_n مختصات بردار X در پایهٔ اولیه و x'_1، x'_2، ...، x'_n مختصات آن در پایهٔ جدید باشد. در این صورت

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

که در آن، P ماتریسی متعامد است.

در صورت درجه دوم $AX \cdot X$ ، به متغیرهای جدید می‌رویم. در متغیرهای جدیدِ صورت درجه دوم، برای ضریب‌ها داریم:

$$a'_{ij} = Af_i \cdot f_j$$

بنابراین

$$a'_{11} = Af_1 \cdot f_1 = \lambda_1 f_1 \cdot f_1 = \lambda_1 \quad ,$$
$$a'_{1j} = a'_{j1} = Af_1 \cdot f_j = \lambda_1 f_1 \cdot f_j = 0 \quad (j \neq 1)$$

یعنی صورت به این صورت درمی‌آید:

$$\lambda_1 x'^2_1 + \varphi(x'_2, \ldots, x'_n)$$

به این ترتیب، با تبدیل متعامد توانستیم، یک مجذور از متغیر جدید را جدا کنیم.

با استفاده از همین استدلال دربارهٔ صورت تازهٔ $\varphi(x'_2, \ldots, x'_n)$، سرانجام به‌جایی می‌رسیم که صورت، به‌وسیلهٔ زنجیری از تبدیل‌های متعامد، به‌صورت متعارف درمی‌آید. ولی روشن است که، زنجیری از تبدیل‌های متعامد، هم‌ارز با یک تبدیل متعامد است. و به این ترتیب، اثبات کامل قضیه به‌دست می‌آید.

۶. تابعِ ماتریس‌ها و برخی کاربردهای آن

تابع ماتریس‌ها. کاربردهای جبر خطی در شاخه‌های دیگر ریاضیات، بسیار زیاد و متنوع است. بدون اغراق، در بسیاری از بخش‌های ریاضیات و فیزیک نظریِ امروز، از اندیشه‌ها و نتیجه‌گیری‌های جبر خطی، به صورت‌های گوناگون و به‌طور عمده به‌صورت محاسبهٔ ماتریس‌ها استفاده می‌شود.

در این‌جا، به‌صورتی کوتاه، یکی از راه‌های کاربرد محاسبهٔ با ماتریس‌ها در نظریهٔ معادله‌های دیفرانسیلی عادی را می‌آوریم. در این‌جا، نقش اصلی، به‌عهدهٔ تابع‌هایی از ماتریس‌هاست.

در آغاز، تعریف توان‌های ماتریس مربعی A را می‌آوریم. فرض می‌کنیم:

$$A^0 = E \ , \ A^1 = A \ , \ A^2 = A \cdot A \ , \ A^3 = A^2 \cdot A \ , \ A^4 = A^3 \cdot A \ , \ \ldots$$

به‌سادگی ثابت می‌شود، برای هر دو عدد طبیعی m و n داریم:

$$A^m A^n = A^{m+n}$$

برای ماتریس، عمل جمع و عمل ضرب در یک عدد، تعریف شده است. درنتیجه، این امکان پیدا می‌شود که مقدار چندجمله‌ایِ ماتریس (با یک متغیر) را پیدا کنیم. اگر

$$\varphi(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n$$

آن‌وقت، تعریف می‌کنیم:

$$\varphi(A) = a_0 A^n + a_1 A^{n+1} + \cdots + a_n E$$

به این ترتیب، تعریف مفهوم ساده‌ترین تابع با آوند ماتریس، به‌صورت یک چندجمله‌ای، به‌دست می‌آید.

با عبور حدی، به‌سادگی می‌توان مفهوم تابع با آوند ماتریسی را برای گروه به‌مراتب وسیع‌تری از تابع‌ها، نسبت به چندجمله‌ای‌های با یک متغیر، تعمیم داد. بدون این که به این مسأله، با همهٔ گستردگی آن بپردازیم، خود را به بررسی تابع‌های تحلیلی محدود می‌کنیم.

در آغاز با مفهوم حد دنباله‌ای از ماتریس‌ها آشنا می‌شویم. دنبالهٔ ماتریس‌های

$$A_1 = \begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \ , \ A_2 = \begin{pmatrix} a_{11}^{(2)} & \cdots & a_{1n}^{(2)} \\ \cdots & \cdots & \cdots \\ a_{n1}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} \ , \ \ldots$$

به سمت ماتریس $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ هم‌گرا (یا A را حد دنبالهٔ ماتریس‌های A_i ،

$i = 1, 2, \ldots$) گویند، وقتی که برای هر i و j داشته باشیم:

$$\lim_{k \to \infty} a_{ij}^{(k)} = a_{ij}$$

سپس، مجموع رشتهٔ $A_1 + A_2 + \cdots + A_k + \cdots$ را به حد مجموع

$$\lim_{k \to \infty} (A_1 + A_2 + \cdots + A_k)$$

گویند، به‌شرطی که این حد وجود داشته باشد.

$f(z)$ را تابعی تحلیلی، که در همسایگی $z = 0$ منظم است، در نظر می‌گیریم. در این صورت، همان‌طور که می‌دانیم، می‌توان $f(z)$ را به‌صورت یک رشتهٔ توانی باز کرد:

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k + \cdots$$

طبیعی است، برای هر ماتریس مربعی A، فرض شود:

$$f(A) = a_0 E + a_1 A + a_2 A^2 + \cdots + a_k A^k + \cdots$$

معلوم می‌شود که این رشته، برای هر ماتریس A، هم‌گراست اگر ویژه‌مقدارهای A در درون دایرهٔ هم‌گرایی $f(z)$ قرار دارند:

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k + \cdots$$

در کاربردها، تابع‌های مقدماتیِ از ماتریس‌ها، می‌تواند برای ما جالب باشد.

برای نمونه، تصاعد هندسی $E + A + A^2 + \cdots + A^k + \cdots$، رشته‌ای هم‌گرا برای ماتریس‌هایی است که قدرمطلق ویژه‌مقدارهای آن‌ها، از ۱ کوچکتر است؛ برای مجموع این رشته، ماتریس $(E-A)^{-1}$ به‌دست می‌آید که، به‌طور کامل، متناظر با این دستور است:

$$1 + x + \cdots + x^k + \cdots = \frac{1}{1-x}$$

بیان $(E-A)^{-1}$ به‌صورت یک رشتهٔ بی‌پایان، وسیلهٔ مؤثری برای حل تقریبی دستگاه معادله‌های خطی است که ماتریس ضریب‌های آن نزدیک به ماتریس همانی باشد.

درواقع، اگر این دستگاه را به‌صورت

$$(E - A) X = B$$

بنویسیم، به‌دست می‌آید:

$$X = (E - A)^{-1} B = B + AB + A^2 B + \cdots \qquad (15)$$

که دستور راحتی برای حل دستگاه است، تنها به‌شرطی که رشتهٔ (۱۵)، خیلی سریع هم‌گرا باشد.

بررسی رشتهٔ دوجمله‌ای می‌تواند سودمند باشد:

$$(E+A)^m = E + \frac{m}{1}A + \frac{m(m-1)}{2!}A^2 + \cdots$$

که، اگر ویژه‌مقدارهای A از لحاظ قدرمطلق کمتر از واحد باشند، می‌تواند نه‌تنها برای توان‌های طبیعی m، بلکه برای توان‌های کسری و منفی هم به‌کار رود.

به‌ویژه، تابع نمایی از ماتریس‌ها، برای کاربرد، اهمیت زیادی دارد:

$$e^A = E + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

رشته‌ای که از تابع نمایی به‌دست می‌آید، برای هر ماتریس A هم‌گراست. تابع نمایی از ماتریس‌ها دارای ویژگی‌هایی است که ویژگی‌های تابع نمایی عادی را به‌یاد می‌آورد. از جمله، اگر ماتریس‌های A و B نسبت به ضرب از قانون جابه‌جایی پیروی کنند، یعنی داشته باشیم $AB=BA$، آن‌وقت $e^{A+B}=e^A\cdot e^B$. ولی در حالتی که A و B نسبت به ضرب از قانون جابه‌جایی پیروی نکنند، این دستور درستی خود را از دست می‌دهد.

کاربرد در نظریهٔ دستگاه‌های معادله‌های دیفرانسیلی خطی عادی. در نظریهٔ دستگاه‌های معادله‌های دیفرانسیلی خطیِ عادی، بهتر است ماتریس‌هایی را در نظر بگیریم که درایه‌های آن‌ها، تابع‌هایی از متغیر مستقل‌اند:

$$U(t)=\begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \cdots & \cdots & \cdots \\ a_{m1}(t) & \cdots & a_{mn}(t) \end{pmatrix}$$

برای این‌گونه ماتریس‌ها، مفهوم مشتق نسبت به آوند t، به‌صورتی طبیعی تعریف می‌شود:

$$\frac{dU(t)}{dt}=\begin{pmatrix} a'_{11}(t) & \cdots & a'_{1n}(t) \\ \cdots & \cdots & \cdots \\ a'_{m1}(t) & \cdots & a'_{mn}(t) \end{pmatrix}$$

به‌سادگی می‌توان تحقیق کرد، برخی دستورهای مقدماتی دیفرانسیل‌گیری، برای ماتریس‌ها هم درست است:

$$\frac{d(U+V)}{dt}=\frac{dU}{dt}+\frac{dV}{dt} \ ,$$

$$\frac{d(cU)}{dt}=c\frac{dU}{dt} \quad ,$$

$$\frac{d(UV)}{dt}=\frac{dU}{dt}V+U\frac{dV}{dt}$$

(ضرب باید درست به همان ردیفی باشد که در دستور داده شده است).

دستگاه معادله‌های دیفرانسیلی همگن خطیِ عادی

$$\frac{dy_1}{dt}=a_{11}(t)y_1+a_{12}(t)y_2+\cdots+a_{1n}(t)y_n \ ,$$

$$\frac{dy_2}{dt}=a_{21}(t)y_1+a_{22}(t)y_2+\cdots+a_{2n}(t)y_n \ ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{dy_n}{dt}=a_{n1}(t)y_1+a_{n2}(t)y_2+\cdots+a_{nn}(t)y_n$$

را با نمادهایی که به‌کار بردیم، می‌توان به این صورت، به‌طور کوتاه، نوشت:

$$\frac{dY}{dt}=A(t)\,Y$$

که در آن داریم:

$$Y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad , \quad A(t)=\begin{pmatrix} a_{11}(t) & \ldots & a_{1n}(t) \\ \cdots & \cdots & \cdots \\ a_{n1}(t) & \ldots & a_{nn}(t) \end{pmatrix}$$

یعنی به‌صورتی مشابه با یک معادلهٔ دیفرانسیلی همگن عادی.

اگر ضریب‌های دستگاه ثابت باشند، یعنی ماتریس A ثابت باشد، آن‌وقت جواب دستگاه، از نظر ظاهری، شبیه جواب معادلهٔ $y'=ay$ است. در این حالت $Y=e^{At}C$ ، که در آن C ، ماتریسی ستونی از ثابت‌های دل‌خواه است.

وقتی جواب به این صورت باشد، برای بررسی بسیار راحت است. موضوع این است که برای هر تابع تحلیلی $f(z)$، این برابری برقرار است:

$$f(B^{-1}LB) = B^{-1}f(L)\,B$$

چون هر ماتریس را می‌توان به فرم متعارف ژُردان تبدیل کرد (۴ را ببینید)، بنابراین محاسبهٔ تابعی از یک ماتریس، به محاسبهٔ تابعی از ماتریس‌های متعارف منجر می‌شود، که به‌سادگی می‌توان به‌انجام رسانید. به این ترتیب، اگر $A = B^{-1}LB$ که در آن، L ماتریس متعارف است، آن‌گاه

$$Y = e^{At}\,C = B^{-1}\,e^{Lt}\,BC = B^{-1}\,e^{Lt}\,C'$$

که در آن $C' = BC$، ستونی از ثابت‌های دل‌خواه است.

از این دستور به‌سادگی بیان صریح همهٔ مجهول‌های ستون Y به‌دست می‌آید.

ای. آ. لاپو-دانیلوسکی دانشمند روسی توانسته است نظریهٔ تابع‌های ماتریس‌ها را تکامل دهد و از آن برای بررسی دستگاه‌هایی که ضریب‌های متغیر دارند، استفاده کند. نتیجه‌گیری‌های این ریاضی‌دان، از جمله درخشان‌ترین موفقیت‌های ریاضیات در دهه‌های گذشته است.

بخش هفدهم

فضای انتزاعی

آ. د. آلکساندرف

از آن زمان که ن. ای. لباچوسکی، برای نخستین بار، امکان وجود هندسهٔ نااقلیدسی را ثابت کرد و تصور تازه‌ای دربارهٔ رابطهٔ هندسه با جهان خارج پدید آورد، موضوع هندسه، روش آن و کاربردهای آن، بی‌اندازه گسترش یافت. اکنون ریاضی‌دانان، «فضاهای» گوناگونی را مطالعه می‌کنند. همراه با فضای اقلیدسی، فضای لباچوسکی، فضای تصویری، فضاهای n بُعدی مختلف و از جمله فضای بی‌نهایت‌بُعدی، فضای ریمانی، فضای تُوپولوژیک و فضاهای دیگری هم در میدان بررسی ریاضی‌دانان است؛ تعداد این فضاها بی‌پایان است و هرکدام از آن‌ها ویژگی‌های خود و «هندسهٔ» خود را دارند. در فیزیک از مفهوم‌های «فضاهای فازی» و «فضاهای پیکربندی» استفاده می‌کنند؛ نظریهٔ نسبیت تصور دربارهٔ انحنای فضا و نتیجه‌گیری‌های دیگری از نظریهٔ هندسه‌های انتزاعی را به‌کار می‌گیرد.

این انتزاع‌های ریاضی، چگونه و از کجا پدید آمده‌اند؟ پایهٔ آن‌ها بر کدام واقعیت‌هاست و چه ارزش و کاربرد واقعی دارند؟ چه رابطه‌ای با دنیای واقع دارند؟ در ریاضیات، این‌ها را چگونه تعریف و بررسی می‌کنند؟ اندیشه‌های کلی هندسهٔ امروزی، چه ارزشی در ریاضیات دارند؟

در این بخش، باید به این پرسش‌ها پاسخ بدهیم. در این بخش، به خود نظریهٔ فضاهای انتزاعی ریاضی کمتر پرداخته شده است، زیرا به بحثی بسیار طولانی و طرح دستگاه‌هایی از ریاضیات که بی‌اندازه تخصصی‌اند، نیاز دارد. در این‌جا تنها می‌خواهیم، جوهر و درون‌مایهٔ اندیشه‌های تازهٔ هندسی را روشن کنیم و به پرسش‌هایی که مطرح کردیم، پاسخ دهیم؛ کاری که بدون استفاده از اثبات‌های پیچیده و آوردن دستورهای بغرنج، ممکن است.

تاریخ این موضوع از «مقدمات» اقلیدس و از اصل موضوع، یا آن‌طور که معمول شده است، پُوستولای دربارهٔ خط‌های موازی، سرچشمه می‌گیرد.

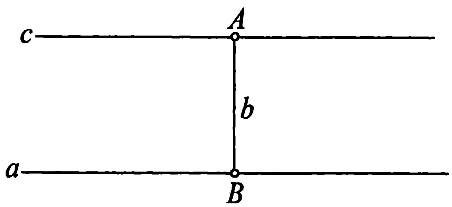
۱. تاریخ پوستولای اقلیدس

اقلیدس در کتاب «مقدمات» خود، پیش‌فرض‌های اساسی هندسه را، به‌صورت پُوستولاها و اصل موضوع‌ها تنظیم کرد. بین این پیش‌فرض‌ها، پُوستولای پنجم (در برخی نسخه‌های مقدمات، اصل موضوع یازدهم) وجود دارد که می‌توان آن را به این صورت بیان کرد: «از یک نقطهٔ واقع در بیرون خط راست، نمی‌توان بیش از یک خط راست، موازی با خط راست مفروض رسم کرد». می‌دانیم، خط راستی را موازی با خط راست مفروض گویند، وقتی که هر دو خط بر یک صفحه واقع باشند و یکدیگر را قطع نکنند؛ درضمن خط راستی را در نظر داریم که بی‌آغاز و بی‌انجام باشد، نه پاره‌خط راستی از آن.

به‌سادگی می‌توان ثابت کرد، از نقطهٔ A که بر خط راست مفروض a واقع نیست، همیشه می‌توان دست‌کم یک خط راست موازی a رسم کرد.

کافی است از نقطهٔ A، عمود b را بر خط راست a و سپس از A خط راست c را بر b عمود کنیم (شکل ۱). شکلی که به‌دست می‌آید، نسبت به b متقارن است، زیرا زاویه‌هایی که خط راست b در دو طرف خود، با خط‌های راست a و c می‌سازد، برابرند. بنابراین، اگر صفحه را روی خط راست b تا کنیم، نیم‌خط‌های راست a و هم نیم‌خط‌های راست c بر هم منطبق می‌شوند. از این‌جا نتیجه می‌شود، اگر خط‌های راست a و c، در یک سمت AB به هم برسند، آن‌وقت باید در سمت دیگر AB هم با یکدیگر برخورد داشته باشند. به این ترتیب، خط‌های راست a و c دو نقطهٔ برخورد پیدا می‌کنند که ممکن نیست، زیرا براساس ویژگی خط راست، از دو نقطه تنها یک خط راست می‌توان رسم کرد (به‌نحوی که، اگر خط‌های راستی، دارای دو نقطهٔ مشترک باشند، باید روی هم قرار گیرند).

به این ترتیب از ویژگی‌های اصلی خط راست و حرکت (زیرا برای تا کردن صفحه روی خط راست AB، باید نیم‌صفحه را دور این خط دوران داد) نتیجه می‌شود، همیشه



شکل ۱

می‌توان یک خط راست از نقطهٔ مفروض گذراند که با خط راست مفروض موازی باشد. و پُوستولای اقلیدس، این نتیجه‌گیری را با این گزاره کامل می‌کند که این خط راست یگانه است و هیچ خط راست دیگری نمی‌تواند از نقطهٔ مفروض بگذرد و با خط راست مفروض موازی باشد.

این پوستولا، بین بقیهٔ پوستولاها (و اصل موضوع‌ها)، تا اندازه‌ای، وضع خاصی دارد. خود اقلیدس، این پوستولا را به‌صورت پیچیده‌تری آورده است، ولی به‌صورتی هم که در این‌جا آوردیم، دشواری‌هایی را به‌همراه دارد. این دشواری به خودِ مفهوم خط‌های راست موازی مربوط می‌شود؛ در این‌جا صحبت بر سر همهٔ خط‌های موازی است. ولی چگونه می‌توان مطمئن شد که خط‌های راست مفروض موازی‌اند؟ برای این منظور، باید آن‌ها را از دو سمت «تا بی‌نهایت» ادامه داد و قانع شد که هرگز و **در *تمامی امتداد بی‌پایان* خود**، به‌هم نمی‌رسند. روشن است که همین تصور، سرچشمهٔ دشواری است. و به‌احتمالی، همین دشواری موجب شد که خود اقلیدس هم، جای خاصی را برای پوستولای پنجم در نظر بگیرد: در مقدمات اقلیدس، این پوستولا، از حکم بیست و نهم آغاز می‌شود، به‌نحوی که اقلیدس، در ۲۸ حکم نخست خود، از آن استفاده نمی‌کند. بنابراین، میل به کنار گذاشتن این پوستولا، و آزاد شدن از این دشواری، خیلی طبیعی است: در همان دوران باستان هم، تلاش‌هایی شد تا تعریف خط‌های راست موازی را تغییر دهند، یا شکل بیان پوستولا را عوض کنند و یا، بهتر از همه، آن را همچون یک قضیه از دیگر اصل‌ها و مفهوم‌های اساسی هندسه، نتیجه بگیرند.

به این ترتیب، با آغاز از همان دوران باستان، نظریهٔ خط‌های راست موازی که براساس پوستولای پنجم تنظیم شده بود، موضوعی برای بحث و تفسیر در کارهای بسیاری از هندسه‌دانان قرار گرفت. مسألهٔ اساسی در این زنجیرهٔ بررسی‌ها این بود که به‌کلی از پوستولای پنجم نجات یابند و آن را به‌عنوان قضیه‌ای، از بقیه اصل‌ها و قضیه‌های هندسه نتیجه بگیرند.

این مسأله، بسیاری از هندسه‌دانان بزرگ را به خود مشغول کرد: پُروکلِس یونانی (سدهٔ پنجم میلادی) مفسر اقلیدس، نصیرالدین توسی ایرانی (سدهٔ سیزدهم)، والیس انگلیسی (۱۶۱۶-۱۷۰۳)، ساکِری ایتالیایی (۱۶۶۷-۱۷۳۳)، فیلسوف و ریاضی‌دانان آلمانی لامبِرت (۱۷۲۸-۱۷۷۷)، لژاندر فرانسوی (۱۷۵۲-۱۸۳۳) و بسیاری دیگر؛ همهٔ آن‌ها در طول بیش از دوهزار سال بعد از پیدایش «مقدمات» اقلیدس، با تیزهوشی و با ظرافت هندسی، در

تلاش اثبات پوستولای پنجم بودند.

ولی همهٔ این تلاش‌ها بی‌نتیجه ماند. هر بار روشن می‌شد، مبتکر این یا آن اثبات، درواقع بر پیش‌فرضی تکیه کرده است که، با همهٔ سادگی خود، نمی‌شد آن را به‌صورتی منطقی، از دیگر پیش‌فرض‌های هندسه به‌دست آورد. به‌زبان دیگر، هر بار پوستولای پنجم با گزارهٔ دیگری عوض شده بود که می‌شد پوستولای پنجم را از آن نتیجه گرفت، درحالی که خودِ این گزاره نیاز به اثبات داشت[1].

ساکری و لامبرت، بیش از دیگران در مسأله عمیق شدند. ساکری در آغاز کوشید پوستولای پنجم را با روش برهان خُلف ثابت کند، یعنی از گزاره‌ای آغاز کند که پوستولای پنجم را نقض می‌کند و برپایهٔ آن پیش برود تا جایی که به تناقض برسد. با ادامهٔ این روش، به نتیجه‌هایی رسید که به‌نظرش پذیرفتنی نمی‌آمد و به‌کلی از تصور در ذهن دور بود؛ و او گمان کرد مسأله را حل کرده است. ولی او اشتباه می‌کرد، زیرا تضاد با تصورهای عینی و دور شدن از آن‌چه می‌بینیم و حس می‌کنیم، نمی‌تواند به‌معنای تضاد منطقی باشد. مسألهٔ اثبات منطقی پوستولای اقلیدس را باید براساس موقعیت‌ها و حکم‌های هندسی ثابت کرد، نه این که دوباره بر آن‌چه به‌نظر درست می‌رسد، تکیه شود. درواقع خودِ پوستولای پنجم هم، به خودی خود و به‌اندازهٔ کافی قانع‌کننده است. ولی تکرار می‌کنیم، «دلیل» عینی و محسوس برای قانع شدن چیزی، و اثبات منطقی چیز دیگری است.

لامبرت نسبت به ساکری و پیشینیان او، اندیشمندی عمیق‌تر بود. او هم از همین راه رفت، ولی به تناقض منطقی نرسید و اشتباه دیگران را تکرار نکرد: او هرگز اعلام نکرد، پوستولای پنجم را ثابت کرده است. ولی بعد از او و در آغاز سدهٔ نوزدهم، لژاندر دوباره پوستولای پنجم را «ثابت کرد» و در همان چاه اشتباه کهنه افتاد: به‌جای پوستولای پنجم، گزارهٔ دیگری گذاشت که خود نیاز به اثبات داشت.

به این ترتیب، در آغاز سدهٔ نوزدهم، مسألهٔ مربوط به اثبات پوستولای پنجم، حل‌نشده

---

۱. این گزاره‌های هم‌ارز با پوستولای پنجم بسیار زیاد بودند. در این‌جا، نمونه‌هایی از آن‌ها را می‌آوریم: ۱) خط راستِ موازی با خط راست دیگر، همه‌جا از آن به یک فاصله است (پروکلِس)؛ ۲) مثلث متشابه (ولی نابرابر) وجود دارند، یعنی مثلث‌هایی که در زاویه‌ها برابر، ولی در ضلع‌ها نابرابرند (والیس)، ۳) دست‌کم یک مستطیل وجود دارد، یعنی چنان چهارضلعی که همهٔ زاویه‌های آن قائمه باشند (ساکِری)؛ ۴) اگر خط راستی بر یک ضلع زاویهٔ حاده عمود باشد، ضلع دوم زاویه را هم قطع می‌کند (لژاندر)؛ ۵) مجموع زاویه‌های یک مثلث برابر دو زاویهٔ قائمه است (لژاندر)؛ ۶) مثلث‌هایی وجود دارند که مساحت آن‌ها به‌دل‌خواه بزرگ باشد (گُوس). این سیاهه را می‌توانیم تا هرجا که بخواهیم ادامه دهیم.

باقی بود و گامی جلوتر از زمان اقلیدس برنداشته شده بود. تلاش‌ها بیهوده بود و مسأله پیشرفتی نکرده بود. درواقع این، معمای بزرگی برای هندسه بود: مسأله‌ای که راه‌حل آن به نظر بهترین ریاضی‌دانان، بی‌تردید وجود داشت، در جریان دوهزار سال، تن به حل نمی‌داد.

در سدهٔ نوزدهم، مسألهٔ خط‌های راست موازی، به‌عنوان یکی از مسأله‌های مرکزی هندسه مطرح بود و ریاضی‌دانان زیادی درگیر با آن شدند: گوس، لاگرانژ، دالامبر، لژاندر، واختر، شه‌وِی‌کار، تااورینوس، فارکاش بایای و دیگران.

با همهٔ این‌ها، در راه اثبات پوستولای پنجم توفیقی به‌دست نیامد. دشواری از کجاست: آیا مربوط به ناتوانی هندسه‌دانان در حل مسأله است، یا این که مسأله به‌صورتی نادرست طرح شده است؟ این پرسش، به‌تدریج برخی از هندسه‌دانان را به اندیشهٔ ژرف‌تری نسبت به دیگران رساند. گوس، ریاضی‌دان مشهور آلمانی، کار روی مسأله را از سال ۱۷۹۲ آغاز کرد و اندک اندک طرح درست مسأله آشکار شد. سرانجام تصمیم گرفت پوستولای پنجم را کنار بگذارد و از سال ۱۸۱۳ آغاز به تنظیم هندسه‌ای کرد که در آن، به‌جای پوستولای پنجم، گزارهٔ متناقض آن را پذیرفته بود. اندکی دیرتر شه‌وِی‌کار ریاضی‌دان آلمانی در دوران اقامت خود در خارکوف و سپس تااورینوس، همین راه را در پیش گرفتند. ولی هیچ‌کدام نتوانستند به نتیجه‌ای قطعی برسند. گوس بادقت بررسی‌های خود را پنهان می‌کرد. شه‌وِی‌کار خود را به نوشتن نامه‌ای برای گوس محدود کرد؛ تنها تااورینوس بود که تکه‌هایی از هندسهٔ جدید را که براساس نفی پوستولای پنجم بود، چاپ کرد. ولی خود تااورینوس، امکان وجود چنین هندسه‌ای را نپذیرفتنی می‌دانست. به این ترتیب، هنوز مسألهٔ خط‌های راست موازی حل نشده بود: هیچ‌کدام از این ریاضی‌دانان، به پاسخ قطعی برای این پرسش نرسید. این پاسخ را، برای نخستین بار، ن. ای. لباچوسکی استاد جوان دانشگاه قازان پیدا کرد: در ۲۳ فوریهٔ سال ۱۸۲۶ میلادی، در نشست دانشکدهٔ فیزیک و ریاضی، سخن‌رانی خود را دربارهٔ نظریهٔ خط‌های راست موازی ایراد و در سال ۱۸۲۹ آن را در نشریهٔ دانشگاه قازان چاپ کرد.

۲. راه‌حل لباچوسکی

۱. درون‌مایهٔ راه‌حل مسألهٔ پوستولای پنجم را، خود لباچوسکی در نوشته‌ای با عنوان «مقدمات جدید هندسه» (۱۸۳۵)، به این ترتیب شرح می‌دهد:

«همه می‌دانند که در هندسه، نظریهٔ خط‌های راست موازی، هنوز به جایی نرسیده است. تلاش‌های بیهوده‌ای که از زمان اقلیدس در طول دوهزار سال در این باره شده است، مرا به این گمان نزدیک کرد که در خودِ مفهوم‌ها، هنوز آن چیزهایی که باید ثابت کرد و بتوان آن‌ها را همچون دیگر قانون‌های فیزیکی، و از جمله به یاری مشاهده‌های اخترشناسی، به آزمایش‌گذاشت، روشن ومشخص نشده‌است. من به‌تدریج به‌درستی گمان خود معتقد شدم و راه‌حل مسأله را برمبنای همان روشی دانستم که در سال ۱۸۲۶ دربارهٔ آن نوشته بودم».

به‌نظر می‌رسد، لباچوسکی در این بیان خود، که اندیشهٔ تازه‌ای را در مرکز توجه خود قرار می‌دهد، تنها حل مسألهٔ مربوط به پوستولای پنجم را در نظر ندارد، بلکه بازبینیِ همهٔ مفهوم‌های هندسه را (و نه‌تنها یک نوع هندسه) طلب می‌کند.

نیکلای لباچوسکی، از سال ۱۸۱۵ کار خود را دربارهٔ نظریهٔ خط‌های راست موازی آغاز کرد و تلاش کرد، شبیه دیگر هندسه‌دانان، درستی پوستولای پنجم را ثابت کند. در سال ۱۸۲۳ کم و بیش برای او روشن شده بود که همهٔ این‌گونه اثبات‌ها، «که تاکنون داده شده است، تنها نوعی روشنگری است، ولی نمی‌توانند ارزش اثبات ریاضی را، به‌مفهوم کامل خود، داشته باشند[1]». او در این‌جا به این نکته توجه می‌کند که «خود مفهوم‌ها شامل حقیقتی که می‌خواهیم اثبات کنیم نیستند»، به‌زبان دیگر، از مقدمه‌ها و مفهوم‌های اساسی هندسه، نمی‌توان پوستولای پنجم را بیرون آورد. چگونه به این موضوع قانع شد که چنین نتیجه‌گیری ممکن نیست؟

در این باره از این راه متقاعد شد که مدت‌ها راهی را که ساکِری و لامبرت رفته بودند، ادامه داد و از آن فراتر رفت. به‌عنوان پیش‌فرض، گزاره‌ای را انتخاب کرد که متناقض با پوستولای اقلیدس بود. او فرض را بر این گذاشت که: «از هر نقطه‌ای که بر خط راست مفروض واقع نباشد، می‌توان، نه یک، بلکه دست‌کم دو خط راست، موازیِ با خط راست مفروض رسم کرد». با انتخاب این گزاره به‌عنوان اصل موضوع و با توجه به پیش‌فرض‌های دیگر هندسه، می‌توانیم جلو برویم و نتیجه‌هایی به‌دست آوریم. اگر این اصل با اصل‌های دیگر هندسه ناسازگار باشد، در جایی با تناقض برخورد می‌کنیم که درنتیجه، پوستولای پنجم با روش برهان خُلف ثابت می‌شود؛ فرض نادرست همیشه ما را به تناقض می‌رساند.

---

۱. این مطلب را لباچوسکی در سال ۱۸۲۳ در دورهٔ هندسهٔ خود نوشت. این نوشته در زمان زندگی او چاپ نشد. دورهٔ «هندسه» لباچوسکی، برای نخستین بار در سال ۱۹۱۰ چاپ شد.

ولی اگر به چنین تناقضی برنخوریم، به دو نتیجه می‌رسیم که لباچوسکی هم بـه هـمان‌جا رسید.

نخستین نتیجه این است که پوستولای پنجم را نمی‌توان ثابت کرد. دوم این که برپایهٔ اصل موضوع‌های متناقض، آن‌گونه کـه در ایـن‌جا تـنظیم شـده‌انـد، مـی‌توان بـه‌طور جـداگـانه، زنجیره‌ای از نتیجه‌گیری‌ها و قضیه‌ها تشکیل داد که به تناقض برنخورند. این نتیجه‌گیری‌ها، می‌تواند از نظر منطقی امکان تنظیم نظریهٔ بی‌تناقضی را پدید آورد که به‌عنوان هـندسه‌ای تازه، هندسهٔ نااقلیدسی، معرفی شود. لباچوسکی بااحتیاط، این هندسه را «تخیلی» نامید، زیرا در آن زمان نمی‌توانست درستی آن را در دنیای واقع روشن کند. ولی امکـان مـنطقی بودنِ این هندسه، برای او روشن بود. لباچوسکی با اعتقاد پایدار به این اندیشه، نبوغ واقعی خود را نشان داد. این پایداری چنان بود که نه در اعتقاد او تزلزلی پدید آمد و نه به‌خاطر عدم درک و انتقادهای دیگران دچار هراس شد.

به این ترتیب، وقتی لباچوسکی به این دو نتیجه رسید، هم به اثبات‌ناپذیری پوستولای پنجم ایمان پیدا کرد و هم به امکان تکامل هندسهٔ جدید براساس اصل موضوعی متناقض با پوستولای پنجم اقلیدس. او مطمئن بود، هندسهٔ جدید، با آن‌که نتیجه‌های غیرقابل تصوری دارد و با آن‌چه به‌طور عینی دربارهٔ فضا دیده می‌شود نمی‌سازد، چه از نظر مضمون و چه از نظر کمال، منطقی و بی‌نقص است. و درواقع، لباچوسکی این هندسهٔ جدید را تکامل داد، هندسه‌ای که امروز نام او را بر خود دارد. او درضمن، به نتیجه‌ای کلی و بسیار مهم رسید: *از نظر منطقی، تنها یک هندسه وجود ندارد*. دربارهٔ ارزش و اهمیت این نتیجه‌گیری باز هم صحبت خواهیم کرد؛ در این نـتیجه‌گیری، بـخش عـمده‌ای از راه‌حـل‌های مسأله‌هـای مـربوط بـه فضاهای ریاضی جبری نهفته بود که در آغاز این بخش مطرح کرده بودیم.

دوباره به سخن لبـاچوسکی، کـه در بـالا آوردیـم، بـرمی‌گردیم. او مـی‌گوید، حـقیقت هندسی شبیه دیگر قانون‌های فیزیکی است که تنها تجربه می‌تواند درستی آن‌ها را تحقیق کند. این، در درجهٔ اول به این معناست که برای درک حقیقت، باید بین مفهوم‌های انتزاعی و دنیای واقعی جهان خارج رابطه‌ای و تناظری برقرار کرد. این رابـطه و تـناظر، تـنها بـه‌یاری تجربه به‌دست می‌آید و بنابرایـن، بـرای تـحقیق درسـتی ایـن یـا آن نـتیجه‌گـیری بـاید بـه بررسی‌های تجربی دست زد و تنها نتیجه‌گیری ذهنی منطقی، برای این منظور، کافی نیست. گرچه هندسهٔ اقلیدسی ویژگی‌های واقعی فضا را خیلی دقیق مـنعکس مـی‌کند، نـمی‌توان اطمینان داشت که بررسی‌های بعدی موجب این کشف نشود که هندسهٔ اقلیدسی، درستی

ویژگی‌های فضای واقعی را، تنها به‌تقریب منعکس می‌کند. در این صورت هندسه، به‌عنوان بازتاب‌دهندهٔ فضای واقعی (و نه به‌عنوان یک دستگاه منطقی)، باید تغییر کند و به‌یاری تجربه‌ها و آزمایش‌های تازه دقیق‌تر شود.

این اندیشهٔ پرنبوغ لباچوسکی، به‌وسیلهٔ شاخه‌های مختلف فیزیک و به‌وسیلهٔ نظریهٔ نسبیت تأیید شد.

خود لباچوسکی، براساس مشاهده‌های اخترشناسی به محاسبه پرداخت تا دقت هندسهٔ اقلیدسی را آزمایش کند. این محاسبه‌ها، در محدودهٔ دقتی که در دسترس بود، درستی این هندسه را تأیید کرد. این وضع اکنون تغییر کرده است، اگرچه بلافاصله باید گفت که هندسهٔ لباچوسکی برای کاربرد در فضا خیلی دقیق‌تر نیست، چرا که فضا ویژگی‌های دیگر و پیچیده‌تری دارد، ولی هندسهٔ لباچوسکی، قبل از این هم، پایداری و کاربردهای خود را در زمینه‌های دیگری پیدا کرده بود که دربارهٔ آن‌ها با تفصیل بیشتری صحبت خواهیم کرد.

باید تأکید کرد لباچوسکی به‌هیچ‌وجه هندسهٔ خود را، خیلی ساده و تنها به‌عنوان یک طرح منطقی که براساس پیش‌فرض‌هایی تصادفی بنا شده باشد، در نظر نگرفت. او مسألهٔ اساسی را، نه در تجزیه و تحلیل منطقی بنیان‌های هندسه، بلکه در بررسی رابطه‌هایی که با واقعیت داشتند، می‌دانست. از آن‌جا که تجربه نمی‌توانست پاسخ دقیق و انتزاعی را دربارهٔ درستی پوستولای اقلیدس بدهد، اندیشهٔ چنان امکان‌های منطقی پیش آمد که به‌یاری آن‌ها بتوان پنهانی‌ترین پیش‌فرض‌های هندسه را بررسی کرد و همین بررسی‌های ریاضی است که می‌تواند مسیرهای بررسی فیزیکی فضای واقعی را مشخص کند. به‌ویژه، از آن‌جا که هندسهٔ اقلیدسی، حالت حدی هندسهٔ لباچوسکی است، بنابراین هندسهٔ لباچوسکی امکان‌های بیشتری را برای بررسی فراهم می‌کند. از این دیدگاه، پوستولای پنجم اقلیدس محدودکننده است و سدی در برابر پیشرفت نظریه است. نظریه باید از این سد عبور کند و در جست‌وجو و شناخت راهی برای کشف حقیقت‌ها و قانون‌های تازه باشد. درک عمیق وابستگی ریاضیات با واقعیت‌ها، این امکان را به ما می‌دهد، از میان امکان‌های منطقی گوناگون، به انتخاب آن‌هایی بپردازیم که بیشترین و دقیق‌ترین مسیر شناخت طبیعت را به ما نشان می‌دهند. اگر به‌دنبال لباچوسکی، دانش هندسه، ریاضیات را با امکان‌های تازه‌ای برای شناخت دقیق‌تر فضا مسلح نمی‌کرد، آن‌وقت فیزیک امروز، چنان ابزار نیرومندی از ریاضیات را در اختیار نداشت که بتواند موقعیت نظریه نسبیت را تنظیم کند و تکامل دهد.

به این ترتیب، می‌توان نتیجه‌هایی را که ناشی از راه‌حل لباچوسکی برای پوستولای پنجم اقلیدس است، برشمرد:

۱°. پوستولای اقلیدس را ثابت نمی‌کنیم؛

۲°. اگر اصل موضوعی متناقض پوستولای پنجم به بنیان‌های هندسه اضافه کنیم، می‌توان هندسه‌ای را طرح ریخت که منطقی و کامل است، ولی از نظر مضمون با هندسهٔ اقلیدسی فرق دارد؛

۳°. درستی نتیجه‌هایی که با استدلال منطقی در این هندسه به‌دست می‌آید و کاربرد آن‌ها در فضای واقعی، تنها به‌یاری تجربه تأیید می‌شود. به‌نظر لباچوسکی، نباید با هندسه همچون یک طرح منطقی دل‌خواه برخورد کرد، بلکه هندسه، نظریه‌ای است که باید بتواند مسیرها و روش‌های ممکن را برای تکامل نظریه‌های فیزیکی فراهم آورد.

این راه‌حل، با راه‌حل‌های دیگری که در تلاش پدید آوردن هندسه‌ای بودند تا پوستولای پنجم اقلیدس را تأیید و اثبات کند، به‌کلی متفاوت است. این راه‌حل، به‌کلی با آن‌چه در ذهن ریاضی‌دانان می‌گذشت، مغایر بود. این، راه‌حلی به‌کلی تازه و ریشه‌ای بود. لباچوسکی، گره کور خط‌های راست موازی را گشود، ولی به شیوه‌ای که دیگران گمان نمی‌کردند.

۲. به‌تقریب هم‌زمان با ***لباچوسکی***، ***یانوش بایای*** (۱۸۰۲-۱۸۶۰ میلادی) هندسه‌دان مجار هم، ناممکن بودن اثبات پوستولای پنجم اقلیدس و امکان وجود هندسهٔ نااقلیدسی را کشف کرد. نتیجه‌گیری‌های ***بایای*** در سال ۱۸۳۲ میلادی، به‌صورت ضمیمهٔ رسالهٔ هندسی پدرش ***فارکاش بایای***، چاپ شد. پدر، نوشتهٔ پسرش را، پیش از چاپ برای ***گوس*** فرستاد و پاسخ تحسین‌آمیزی دریافت کرد؛ درضمن ***گوس*** در نامهٔ خود اطلاع داده بود که خود او هم مدت‌ها پیش به همین نتیجه رسیده است. ولی ***گوس*** نتیجه‌گیری‌های خودش را منتشر نکرد. او در یکی از نامه‌های خود نوشته است که نگران است، دیگران نتوانند موضوع را درک کنند.

این وضع در تاریخ دانش، بسیار پیش آمده است که وقتیِ موضوعی به‌اندازهٔ کافی پخته می‌شود، دانشمندان زیادی به‌طور هم‌زمان و بدون آگاهی از کار یکدیگر، به نتیجه‌ای یکسان رسیده‌اند. ***نیوتُن*** و ***لایب‌نیتس***، به‌طور هم‌زمان، محاسبهٔ انتگرالی و محاسبهٔ دیفرانسیلی را پیش بردند؛ ***والاس***، هم‌زمان با ***داروین*** و بدون رابطه با او، به همان اندیشهٔ ***داروین*** رسید. ***پوانکاره*** هم نظریهٔ نسبیت را هم‌زمان با ***اینشتین*** طرح کرد، و این فهرست را می‌توان همچنان

ادامه داد. این هم‌زمانی‌ها، یکبار دیگر ثابت می‌کند که دانش، به‌صورتی لازم و ضروری و از راه حلِ مسأله‌های حل‌نشدهٔ آن، تکامل می‌یابد، نه به‌صورت تصادفی و معمایی. امکان وجود هندسهٔ نااقلیدسی هم، به‌طور هم‌زمان و به‌وسیلهٔ چند هندسه‌دان کشف شد: ***لباچوسکی***، ***بایای***، ***گوس***، ***شه‌وِی‌کار*** و ***تاآورینوس***.

ولی، به این موضوع هم که همه‌جا در تاریخ دانش دیده می‌شود اشاره کنیم، همهٔ دانشمندانی که به نتیجه‌گیری تازهٔ واحدی رسیده‌اند، در استحکام بخشیدن به آن نقشی یکسان نداشته‌اند و نمی‌توان همهٔ آن‌ها را در یک ردیف قرار داد. در این‌جا، زمان اهمیت دارد: چه کسی زودتر از دیگران نظریه را طرح کرده است؛ به‌جز آن چه کسی آن را عمیق‌تر آورده و چگونه با پی‌گیری دنبال بحث را گرفته و به آن استحکام بخشیده است! نه ***شه‌وِی‌کار*** و نه ***تاآورینوس*** نتوانستند به هندسهٔ جدید جنبهٔ قانونی بدهند و تنها نتیجه‌های جداگانه‌ای به‌دنبال کارهای ***ساکِری*** و ***لامبرت*** به‌دست آوردند. گوس به‌ظاهر به وجود هندسهٔ جدید پی برده بود، ولی چنان پایداری از خود نشان نداد که آن را تا مرز لازم پیش ببرد. **گوس** ریاضی‌دانی ژرف‌بین و همه‌جانبه بود، ولی شجاعت ابراز اندیشه‌های تازهٔ خود را نداشت و، بنابراین، نتوانست دگرگونی لازم را، آن‌گونه که به‌وسیلهٔ لباچوسکی پدید آمد، در ریاضیات ایجاد کند. در این باره می‌توان **گوس** را نمونهٔ مشخص روشنفکران آلمانی آن زمان دانست که ضمن داشتن تفکر نظری عمیق، با دشواری‌های سیاسی روبه‌رو بودند. این مطلب را می‌توان با نمونهٔ ***هِگل*** که با **گوس** هم‌زمان بود، روشن کرد که به‌قول ***لنین***، دیالکتیک طبیعت و معرفت را «با نبوغ خود فهمید، ولی تنها فهمید و نه بیشتر» و به‌قول ***هِرتسن*** «جبر دگرگونی» را زیر تأثیر فیلسوفان ذهن‌گرا و ارتجاعی زمان خود درک کرد.

بایای غیرقابل حل بودن مسأله را متوجه شد، ولی از آن جلوتر نرفت و همچون لباچوسکی اندیشهٔ ضروری و عمیق تکامل آن را مطرح نکرد. تنها لباچوسکی بود که کشف خود را، برای نخستین بار در سخن‌رانی ۱۸۲۶ خود اعلام و در ۱۸۲۹ چاپ کرد. او در نوشته‌های بعدی خود، که تا سال ۱۸۵۵ ادامه داشتند، اندیشهٔ جدید را تکامل داد و بر آن پای فشرد. «هندسه‌ای» که او آورد، گرچه در آغاز فریب می‌نمود و مدتی طول کشید تا عقلانی بودن آن بر همگان روشن شد، مستحکم و قانع‌کننده بود. به همین دلیل است که هندسهٔ جدید، به‌حق به نام او شهرت یافته است.

نیکلا ایوانوویچ لباچوسکی، نه‌تنها هندسهٔ جدید را تکامل داد، بلکه به‌درستی پرسش مربوط به بستگی هندسه با جهان واقع را هم مطرح کرد. ذهن‌گرایان تا امروز هم علاقهٔ

زیادی به زنده نگه‌داشتن فلسفهٔ ***کانت*** دارند که، بنا بر آن، فضا صورت واقعیِ وجودِ ماده نیست و تنها انعکاسی از تصور ذهنی ما و بنابراین، مستقل از تجربه است، یعنی صورتی تخیلی دارد و نمی‌توان آن را وارد آزمایش کرد. به این ترتیب، به‌نظر ***کانت***، هندسه هم دانشی حضوری است و به آزمایش ربطی ندارد. ***لباچوسکی*** علیه دیدگاه نظری ***کانت*** بود. او می‌گفت، تنها از راه آزمایش می‌توان به حقیقت رسید. برخلاف ***کانت***، ***لباچوسکی*** تأکید می‌کرد، اگر «مفهوم‌های اولیه»، یعنی مفهوم‌های بنیانی «با آن‌چه محسوس است در تضاد باشند، نادرست‌اند». ***لباچوسکی***، نه‌تنها اعتقاد داشت، سرچشمهٔ تصورها و مفهوم‌های هندسی تجربه است (که ماتریالیست‌های پیش از او هم بر این امر تأکید داشتند)، بلکه روشن کرد که بستگی هندسه با واقعیت هم، باید به‌یاری تجربه دقیق‌تر شود. در این‌جا، اندیشهٔ تکامل هندسه هم دیده می‌شود، اندیشهٔ نزدیک‌تر شدن پیوسته و دایمی به حقیقت مطلق براساس پیشرفت آگاهی‌های ما.

به این ترتیب، ***لباچوسکی*** نه‌تنها یک هندسه‌دان نابغه، بلکه درضمن فیلسوفی ماتریالیست بود. او بسیار فعال و پرانرژی در جنبه‌های گوناگون و از جمله در زمینهٔ آموزش بود: استادی ریاضیات و نزدیک به ۲۰ سال ریاست دانشگاه قازان را برعهده داشت. واقع‌گرایی ***لباچوسکی***، که براساس آن با شجاعت و پایداری از اندیشه‌های خود دفاع می‌کرد، با وجود عدم درک و گاه استهزای دیگران، در تمامی فعالیت‌های گسترده‌ای که داشت، او را به‌عنوان دانشمند، مربی و سازمان‌ده معرفی می‌کند و او را یکی از بهترین چهره‌های آموزشی زمان خود می‌شناساند. دوران ***لباچوسکی***، دوران شکوفایی نبوغ روسی و خودآگاهی اجتماعی، دوران **پوشکین**، دکابریست‌ها و ***بِلینسکی*** بود. روسیه وارد عرصهٔ جهانی می‌شد، آن هم نه به‌عنوان شاگرد حرف‌شنو فرهنگ اروپایی، بلکه به‌عنوان نیرویی که به‌نوبهٔ خود تازه‌هایی را مطرح می‌کرد که اروپا از آن بی‌اطلاع بود. از جمله در ریاضیات، ***لباچوسکی*** مسأله‌ای را حل کرد که دوهزار سال بدون حل مانده بود و آن را به مسیر تازه و منطقی خود انداخت. واقع‌گرایی و شجاعت علمی او، همانند واقع‌گرایی و شجاعت علمی ***رادیشچف***، ***پِسْتِل*** و ***بِلینسکی*** بود. ***لباچوسکی*** تنها پدیدآورندهٔ هندسهٔ جدید نبود، او دانشمند، متفکر و انسانی بزرگ، به مفهوم واقعیِ این کلمه، بود. بعد از ***لباچوسکی***، هندسه از مرزهایی که او می‌توانست بیندیشد، بسیار جلوتر رفت، ولی به‌حق باید او را آغازگر دوران جدیدی در هندسه و در تمامی ریاضیات دانست.

۳. هندسهٔ لباچوسکی

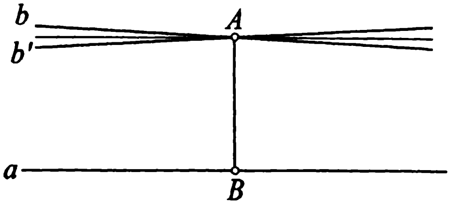
۱. به این ترتیب، ***لباچوسکی*** مبنای کار خود را بر گزاره‌ای متناقض پوستولای پنجم قرار داد: در صفحهٔ مفروض، از هر نقطه می‌توان دست‌کم دو خط راست رسم کرد که با خط راست مفروض، برخورد نداشته باشد. او برمبنای این گزاره، پیش رفت و هندسهٔ تازهٔ خود را ساخت. بنابراین، این هندسه، به‌عنوان نظریه‌ای قابل درک و به‌صورت مجموعه‌ای از قضیه‌هایی که از نظر منطقی قابل اثبات بودند، پدید آمد؛ او از گزارهٔ خود آغاز کرد و با استفاده از دیگر مفهوم‌های اصلی اقلیدسی[۱]، یا به‌گفتهٔ ***لباچوسکی***، هندسهٔ «معمولی» جلو رفت.

لباچوسکی، در نتیجه‌گیری‌های خود، همهٔ نتیجه‌ها را شبیه هندسهٔ مقدماتی «معمولی» به‌دست آورد، یعنی تا مثلثات نااقلیدسی و حل مثلث‌ها، تا محاسبهٔ مساحت‌ها و حجم‌ها پیش رفت. ما در این‌جا نمی‌توانیم زنجیرهٔ این نتیجه‌گیری‌ها را دنبال کنیم، نه به این دلیل که خیلی پیچیده‌اند، بلکه به‌دلیل محدود بودن جای مطلب. مگر نه این است که هندسهٔ «معمولی» دبیرستانی به‌اندازهٔ کافی پرصفحه است، و البته، نتیجه‌گیری‌های لباچوسکی نه ساده‌تر از نتیجه‌گیری‌های هندسهٔ «معمولی» است و نه کوتاه‌تر از آن‌ها. به همین مناسبت، در این‌جا تنها به برخی از نتیجه‌گیری‌های شگفتِ ***لباچوسکی*** می‌پردازیم و خوانندهٔ علاقه‌مند را، برای مطالعهٔ کامل مسیر هندسهٔ نااقلیدسی، به کتاب‌های تخصصی راهنمایی می‌کنیم. سپس برداشتی ساده از هندسهٔ نااقلیدسی در دنیای واقعی را هم می‌آوریم.

از نظریهٔ خط‌های راست موازی آغاز می‌کنیم. خط راست a و نقطهٔ A را واقع در بیرون a در نظر می‌گیریم. AB را عمود بر خط راست a می‌گیریم. بنا بر فرض، دست‌کم دو خط راست وجود دارد که از نقطهٔ A می‌گذرد و با خط راست a برخوردی ندارد. روی شکل ۲، خط‌های راست b و b'، برخلاف فرض ***لباچوسکی***، در امتداد خود، خط راست a را قطع می‌کنند. ولی از این بابت نباید تعجب کرد. ***لباچوسکی***، استدلال خود را نه روی شکل‌هایی که در صفحهٔ عادی رسم شوند، بلکه با نتیجه‌گیری‌های منطقی از فرض خود پیش بُرد، که البته با آن‌چه عادت کرده‌ایم روی شکل ببینیم مخالف است. در این‌جا، شکل نقشی کمکی دارد؛ روی این شکل‌ها نمی‌توان حقیقت‌های هندسهٔ نااقلیدسی را بادقت نشان داد، زیرا

---

۱. این مفهوم‌ها و به‌اصطلاح «بقیهٔ» اصل‌های هندسی، در بند ۵ همین بخش بادقت تنظیم شده است.



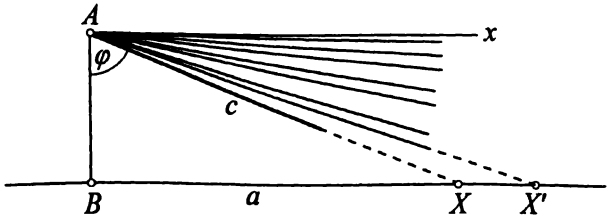
شکل ۲

شکل را، در مرز ممکنِ دقت، روی صفحهٔ عادی اقلیدسی رسم کرده‌ایم.

این تضاد بین امکان استدلال منطقی با تصور عینی، دشواری زیادی در درک هندسهٔ لباچوسکی پدید می‌آورد. ولی وقتی صحبت بر سر هندسه به‌عنوان یک نظریهٔ منطقی است، باید همهٔ توجه را به سمت استدلال دقیق منطقی برگرداند، نه به شکل‌های عادی که روی صفحه یا در فضای اقلیدسی رسم می‌شوند.

۲. دوباره به خط راست a و نقطهٔ A برمی‌گردیم. از نقطهٔ A، نیم‌خط راست x را طوری رسم می‌کنیم که با خط راست a برخورد نداشته باشد (برای نمونه، عمود بر AB)، آن‌وقت این نیم‌خط راست را دور نقطهٔ A طوری دوران می‌دهیم که زاویهٔ φ بین AB و x کوچک شود، ولی تا جایی که این نیم‌خط راست، خط a را قطع نکند. در این صورت، نیم‌خط راست x به سمت موقعیت حدی میل می‌کند که پاسخ‌گوی کمترین مقدار زاویهٔ φ است. این نیم‌خط راست حدی c هم، a را قطع نمی‌کند.

در واقع، اگر این خط راست حدی، خط راست a را در نقطه‌ای مثل X قطع کند (شکل ۳)، آن‌وقت می‌توانیم نقطهٔ X' را در سمت راست X و روی a در نظر بگیریم که درنتیجه نیم‌خط راست AX' به‌دست می‌آید که خط راست a را قطع می‌کند و درضمن با AB زاویهٔ بزرگتری می‌سازد. ولی این ممکن نیست، زیرا طبق تعریف نیم‌خط راست c، هر



شکل ۳

نیم خط راستی که با AB زاویه بزرگتری را بسازد خط راست a را قطع نمی‌کند.

بنابراین، نیم‌خط راست c، خط راست a را قطع نمی‌کند و درضمن، مرز خط‌های راستی است که از A می‌گذرند و a را قطع نمی‌کنند.

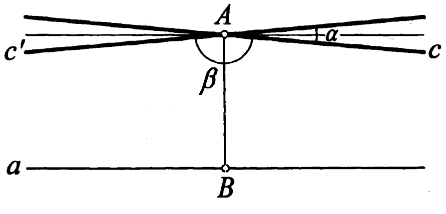
با توجه به تقارن روشن است، در طرف دیگر هم می‌توان نیم‌خط راست c' را پیدا کرد که a را قطع نکند و نیم‌خط راست مرزی در بین همهٔ این‌گونه نیم‌خط‌های راست باشد. اگر c و c' در امتداد هم باشند، آن‌وقت با هم خط راست $c+c'$ را تشکیل می‌دهند. در این صورت، این خط راست، یگانه خط راستی است که از نقطهٔ A می‌گذرد و با a موازی است، زیرا با اندک دوران c یا c'، خط راست a را قطع می‌کند. ولی فرض کرده بودیم که از نقطهٔ A، نه یک بلکه دو خط راست می‌توان موازی a رسم کرد، بنابراین c و c' در امتداد یکدیگر نیستند.

به این ترتیب، نخستین قضیهٔ هندسهٔ ***لباچوسکی*** را ثابت کردیم:

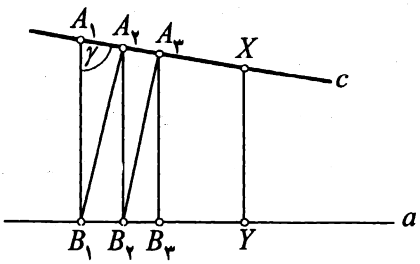
از نقطهٔ A که بر خط راست a واقع نیست، می‌توان دو نیم‌خط راست c و c' را طوری رسم کرد که خط راست a را قطع نکنند، ولی هر نیم‌خط راستی که از A بگذرد و در بیرون زاویهٔ بین c و c' باشد، خط راست a را قطع کند.

اگر نیم‌خط‌های راست c و c' را ادامه دهیم، دو خط راست به‌دست می‌آید (شکل ۴) که خط راست a را قطع نمی‌کنند و این ویژگی را دارند که هر خط راست که از نقطهٔ A بگذرد و در درون زاویهٔ α، بین این دو خط راست باشد، خط راست a را قطع نمی‌کند و هر خط راستی که از درون زاویهٔ β بگذرد، خط راست a را قطع می‌کند. ***لباچوسکی*** این‌گونه خط‌های راست c و c' را، ***موازی خط راست*** a می‌نامد: c از سمت راست و c' از سمت چپ، با خط راست a موازی است. ***لباچوسکی***، نصف زاویهٔ β را، ***زاویهٔ توازی*** می‌نامد؛ زاویهٔ توازی از زاویهٔ قائمه کمتر است، زیرا β از قائمه کوچکتر است.

۳. اکنون ببینیم اگر X در طول c حرکت کند (شکل ۵)، فاصلهٔ نقطهٔ X از خط راست c، تا



شکل ۴



شکل ۵

خط راست a، چگونه تغییر می‌کند. در هندسهٔ اقلیدسی، فاصلهٔ بین دو خط راست موازی، مقدار ثابتی است. ولی ثابت می‌کنیم، در این‌جا، با حرکت نقطهٔ X به سمت راست، فاصلهٔ آن از a (یعنی طول عمود XY) تنزل می‌کند.

از نقطهٔ A_1، عمود A_1B_1 را بر خط راست a در نظر می‌گیریم. از نقطهٔ B_1، عمود B_1A_2 را بر خط راست c رسم می‌کنیم (نقطهٔ A_2 در سمت راست A_1 واقع است، زیرا γ زاویه‌ای حاده است). سرانجام از نقطهٔ A_2، عمود A_2B_2 را دوباره بر خط راست a رسم می‌کنیم. ثابت می‌کنیم A_2B_2 از A_1B_1 کوچکتر است.

در هندسهٔ لباچوسکی هم، قضیه‌ای که می‌گوید، عمود از مایل کوتاه‌تر است، درست است، زیرا این قضیه را می‌توان بدون تکیه بر مفهوم خط‌های راست موازی ویا نتیجه‌های ناشی از آن ثابت کرد (این اثبات را در هر کتاب هندسهٔ مقدماتی می‌توان پیدا کرد). چون طول عمود از طول مایل کمتر است، بنابراین B_1A_2 به‌عنوان عمود بر خط راست c از A_1B_1 کوتاه‌تر و A_2B_2 هم به‌عنوان عمود بر خط راست a از B_1A_2 کوتاه‌تر است. درنتیجه A_2B_2 از A_1B_1 کوتاه‌تر می‌شود.

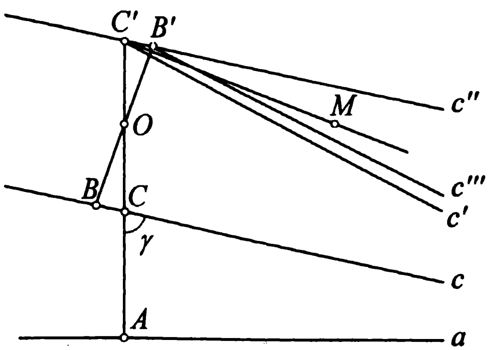
سپس اگر عمود B_2A_3 را از نقطهٔ B_2 بر خط راست c رسم و همان استدلال را تکرار کنیم، روشن می‌شود که A_3B_3 از A_2B_2 کوتاه‌تر است. با ادامهٔ این ساختمان، دنباله‌ای از عمودها به‌دست می‌آید که یکی از دیگری کوچکتر است. این عمودها، معرف فاصلهٔ نقطه‌های A_1 ، A_2 ، ... از خط راست a هستند که هرچه بیشتر به سمت راست برویم، طولی کوتاه‌تر دارند. اگر بخواهیم این استدلال ساده را کامل کنیم، می‌توان ثابت کرد که اگر در حالت کلی، نقطهٔ X'' در سمت راست X' باشد، آن‌وقت طول عمود $X''Y''$ از طول عمود $X'Y'$ کوچکتر است؛ در این‌جا به این اثبات نمی‌پردازیم. امیدواریم، این استدلال، به‌اندازهٔ کافی روشن‌کننده باشد، ولی باید توجه داشت که اثبات دقیق مسأله را در این‌جا نیاوردیم.

همچنین می‌توان ثابت کرد، فاصلهٔ XY، وقتی نقطهٔ X روی خط راست c به سمت راست حرکت کند، نه‌تنها تنزل می‌کند، بلکه وقتی X به سمت بی‌نهایت میل کند، این فاصله به سمت صفر میل می‌کند. یعنی **خط‌های راست موازی a و c ، به‌صورت مجانبی به‌هم نزدیک می‌شوند.** درضمن می‌توان ثابت کرد، در جهت عکس، فاصلهٔ بین خط‌های راست موازی a و c ، نه‌تنها افزایش می‌یابد، بلکه به سمت بی‌نهایت میل می‌کند.

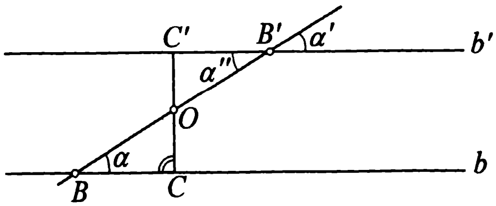
در هندسهٔ اقلیدسی، خط راست موازیِ با خط راست مفروض، از خط راست مفروض، فاصله‌ای ثابت دارد. ولی در هندسهٔ ***لباچوسکی***، چنین دو خط راستی وجود ندارد، چراکه در آن‌جا، دو خط راست همیشه، از یک طرف یا از دو طرف از هم بی‌نهایت دور شوند. خطی که به فاصلهٔ ثابت از خط راست مفروض باشد، هرگز خط راست نیست، بلکه یک منحنی است که به آن ***هم‌فاصله***، یعنی با فاصله‌های برابر گویند.

این نتیجه‌های حاصل از هندسهٔ لباچوسکی شگفتی‌آورند و با تصورهای عینی ناسازگارند. ولی پیش از این هم یادآور شدیم که این ناسازگاری با تصورهای عینی، نمی‌تواند دلیلی علیه هندسهٔ ***لباچوسکی*** باشد، زیرا هندسهٔ ***لباچوسکی*** نظریه‌ای انتزاعی است که به‌صورت منطقی از پیش‌فرض‌های پذیرفته‌شده، تکامل یافته است.

۴. اکنون زاویهٔ توازی را بررسی می‌کنیم، یعنی زاویهٔ γ که خط راست c موازی با خط راست مفروض a ، با عمود CA می‌سازد (شکل ۶). ثابت می‌کنیم، هرچه نقطهٔ C از خط راست a دورتر باشد، این زاویه کوچکتر است. برای این منظور، در آغاز این گزاره را ثابت می‌کنیم: اگر دو خط راست b و b' با قاطع BB' ، زاویه‌های برابر α و α' را بسازد، آن‌وقت این دو خط راست، دارای عمود مشترک هستند (شکل ۷).



شکل ۶



شکل ۷

برای اثبات از نقطهٔ O وسط پاره‌خط راست BB'، خط راست CC' را عمود بر خط راست b رسم می‌کنیم. دو مثلث OBC و $OB'C'$ به‌دست می‌آید. ضلع‌های OB و OB'، بنا بر ساختمان، طولی برابر دارند. زاویه‌های به رأس مشترک O برابرند. همچنین زاویهٔ α'' با زاویهٔ α' برابر است. زاویهٔ α هم طبق فرض با زاویهٔ α' برابر است. بنابراین زاویه‌های α و α'' هم، برابر می‌شوند. به این ترتیب، در مثلث‌های OBC و $OB'C'$، ضلع‌های OB و OB' وهم زاویه‌های کنار آن‌ها باهم برابرند. در این صورت، بنا بر قضیهٔ مشهور، مثلث‌ها و از جمله، زاویه‌های به رأس‌های C و C' برابرند. ولی زاویهٔ C قائمه است، زیرا خط راست CC' را عمود بر خط راست b رسم کرده بودیم. بنابراین زاویهٔ C' هم قائمه است، یعنی CC' بر خط راست b' هم عمود است. به این ترتیب، پاره‌خط راست CC' عمود مشترک خط‌های‌راست b و b' است. وجود عمودمشترک، ثابت شد.

اکنون ثابت می‌کنیم، با افزایش فاصله از خط راست، زاویهٔ توازی کاهش می‌یابد؛ یعنی اگر نقطهٔ C'، همان‌طور که در شکل ۶ دیده می‌شود، نسبت به خط راست a، دورتر از نقطهٔ C باشد، آنگاه خط راست c' که از نقطهٔ C' موازی a رسم شده است، با عمود $C'A$ زاویه‌ای کوچکتر از زاویهٔ بین خط راست موازی c که از نقطهٔ C رسم شده است با این عمود، می‌سازد.

برای اثبات، خط راست c'' را از نقطهٔ C' طوری می‌گذرانیم که همان زاویه‌ای را با $C'A$ بسازد که خط راست موازی c ساخته است. در این صورت، خط‌های راست c و c'' با قاطع CC' زاویه‌های برابر می‌سازند. بنابراین، همان‌طور که ثابت کردیم، این دو خط راست دارای عمود مشترک BB' هستند. در این صورت، از پای B' این عمود، می‌توان خط راست c''' را رسم کرد که با c موازی و با عمود زاویه‌ای کمتر از قائمه ساخته باشد، زیرا همان‌طور که می‌دانیم، خط راست موازی، با عمود زاویه‌ای کمتر از قائمه می‌سازد. اکنون نقطهٔ دل‌خواه M را در زاویهٔ بین c'' و c''' انتخاب و خط راست $C'M$ را از آن می‌گذرانیم. این خط راست در

درون زاویهٔ بین c'' و c''' است و امتداد آن نمی‌تواند c''' را قطع کند. از این گذشته، این خط راست، c را هم قطع نمی‌کند. ولی این خط راست با AC'، زاویه‌ای کوچکتر از زاویه‌ای که c'' با $A'C$ دارد، یعنی کوچکتر از زاویهٔ γ می‌سازد. از این گذشته، موازی c'، زاویهٔ بازهم کوچکتری می‌سازد، زیرا c'، در بین خط‌های راستی که از C' گذشته‌اند و a را قطع نمی‌کنند، مرزی‌ترینِ آن‌هاست. بنابراین، موازی c' با $C'A$، زاویه‌ای کمتر از c می‌سازد و این به‌معنای آن است که زاویهٔ توازی با انتقال به نقطهٔ دورتر C' کوچکتر شده است.

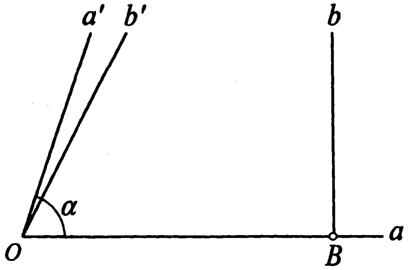
به این ترتیب ثابت کردیم، با دور شدن نقطهٔ C از خط راست a، زاویهٔ توازی کوچکتر می‌شود. ولی به‌نظر می‌رسد، می‌توان چیزی بیشتر از این را ثابت کرد: ***اگر نقطهٔ*** C ***به سمت بی‌نهایت برود، این زاویه به سمت صفر میل می‌کند***؛ یعنی هرچه فاصلهٔ از خط راست a بیشتر باشد، موازی این خط راست با عمود بر a، زاویه‌ای به‌دل‌خواه کوچک می‌سازد.[1] به زبان دیگر، اگر در نقطهٔ دوری از خط راست a، نسبت به عمود بر آن به‌اندازهٔ زاویهٔ کوچکی منحرف شویم، آن‌وقت هرگز خط راست a را، که نسبت به عمود به‌اندازهٔ زاویهٔ قائمه «انحراف» دارد، قطع نمی‌کند. این حقیقتِ هندسهٔ ***لباچوسکی*** هم، موجب تعجب می‌شود، ولی می‌بینید، از این به بعد به نتیجه‌هایی می‌رسیم که کمتر از آن‌چه تا این‌جا آمده است، تعجب‌آور نیست.

از جمله، زاویهٔ حادهٔ α را که نیم‌خط‌های راست a و a' با هم می‌سازند، در نظر می‌گیریم. اندکی دور از نقطهٔ O (رأس زاویهٔ α) عمود b را بر a رسم می‌کنیم، به‌نحوی که زاویهٔ توازی متناظر با فاصلهٔ OB، کوچکتر از α باشد (شکل ۸). چون زاویهٔ α بزرگ‌تر از زاویهٔ توازی است، بنابراین خط راست b' که از O موازی b رسم شود، با a زاویه‌ای کوچکتر می‌سازد. ولی این خط راست b'، خط راست b را قطع نمی‌کند. بنابراین به‌ویژه، a'هم آن را قطع نمی‌کند. و این به‌معنای آن است که، اگر بر ضلع‌های زاویهٔ حاده، در نقطه‌هایی دور از رأس، عمودهایی رسم کنیم، عمود بر یک ضلع، ضلع دیگر را قطع نمی‌کند.

۵. آن‌چه را نتیجه گرفتیم، دو هدف را دنبال می‌کرد. هدف اول، که مهم‌تر است، این بود که روی ساده‌ترین مثال‌ها نشان دهیم، چگونه می‌توان قضیه‌های هندسهٔ ***لباچوسکی*** را با

---

۱. اگر h طول عمود و γ زاویهٔ توازی باشد، بنابر اثبات لباچوسکی داریم: $\tan\frac{\gamma}{2} = e^{-h/k}$ که در آن k ثابتی است در بستگی با واحد طول و e عدد معروف مبنای لگاریتم طبیعی است. روشن است وقتی $h \to \infty$، مقدار $e^{-h/k}$ و همراه با آن مقدار γ به سمت صفر میل می‌کند.



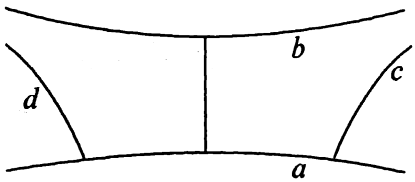
شکل ۸

آغاز از پیش‌فرض‌های پذیرفته‌شده، ثابت کرد. این، یکی از ساده‌ترین نمونه‌ها را در برابر ما می‌گذارد که چگونه ریاضی‌دانان در هندسهٔ انتزاعی، نتیجه‌های خود را به‌دست می‌آورند، و چگونه به‌طور کلی می‌توان به چنان نتیجه‌هایی رسید که ربطی به تصورهای بصریِ عادی ما نداشته باشد. دوم می‌خواستیم نشان دهیم، که در هندسهٔ ***لباچوسکی*** چه نتیجه‌های خاصی به‌دست می‌آیند. باز هم چند مثال می‌آوریم.

دو خط راست در صفحهٔ ***لباچوسکی*** یا متقاطع‌اند، یا به مفهوم ***لباچوسکی*** موازی‌اند که در این صورت، یا در یک سمت به‌صورت مجانبی به هم نزدیک و در سمت دیگر، تا بی‌نهایت از هم دور می‌شوند، یا دارای عمود مشترک‌اند و از هر دو طرف تا بی‌نهایت از هم دور می‌شود.

اگر خط‌های راست a و b عمود مشترک داشته باشند (شکل ۹)، آن‌وقت می‌توان دو عمود c و d را بر a، موازی با خط راست b رسم کرد (البته، موازی به مفهوم ***لباچوسکی***) که تمامی خط راست b در نوار بین خط‌های راست c و d واقع باشد.

حد دایره‌ای که شعاع آن تا بی‌نهایت بزرگ شده باشد، خط راست نیست، بلکه منحنی است که به آن **دایرهٔ حدی** گویند. از هر سه نقطه‌ای که بر یک خط راست نباشند، همیشه نمی‌توان دایره‌ای گذراند؛ از این سه نقطه ممکن است یک دایره بگذرد یا یک دایرهٔ حدی و



شکل ۹

یا **منحنی *هم‌فاصله*** (یعنی یک منحنی که از مجموعهٔ نقطه‌هایی تشکیل شده باشد که همهٔ آن‌ها از خط راستی به یک فاصله‌اند).

مجموع زاویه‌های هر مثلث، همیشه از دو قائمه کمتر است. اگر مثلث به‌گونه‌ای بزرگ شود که هر سه ارتفاع آن به‌طور نامحدود بزرگ شوند، آن‌وقت هر سه زاویهٔ آن به سمت صفر میل می‌کنند.

مثلث‌هایی که مساحتی به‌دل‌خواه زیاد داشته باشند، وجود ندارد.

اگر زاویه‌های دو مثلث، نظیر به نظیر برابر باشند، آن‌وقت دو مثلث برابرند.

طول محیط دایره (l) ، با شعاع آن (r) متناسب نیست و سریع‌تر رشد می‌کند (درواقع با قانونی نمایی) و این دستور برقرار است:

$$l=\pi k\left(e^{\frac{r}{k}}-e^{-\frac{r}{k}}\right) \qquad (۱)$$

که در آن k، عدد ثابتی وابسته به واحد طول است. از آن‌جا که

$$e^{\frac{r}{k}}=1+\frac{r}{k}+\frac{1}{2}\left(\frac{r}{k}\right)^2+\cdots \quad , \quad e^{-\frac{r}{k}}=1-\frac{r}{k}+\frac{1}{2}\left(\frac{r}{k}\right)^2-\cdots$$

با توجه به دستور (۱) به‌دست می‌آید:

$$l = 2\pi r\left(1+\frac{1}{6}\cdot\frac{r^2}{k^2}+\cdots\right) \qquad (۲)$$

که در این صورت، برای مقدارهای کوچک نسبت $\frac{r}{k}$ ، با تقریب کافی، به‌دست می‌آید: $l=2\pi r$.

همهٔ این گزاره‌ها، به‌عنوان نتیجه‌های منطقیِ پیش‌فرض «اصل *لباچوسکی*» و استفاده از اصل‌های هندسهٔ «معمولی» به‌دست می‌آیند.

۶. ویژگی بسیار مهم هندسهٔ *لباچوسکی* در این است که، اگر با حوزه‌های به‌اندازهٔ کافی کوچک سر و کار داشته باشد، اختلاف اندکی با هندسهٔ اقلیدسی دارد؛ هرچه با حوزهٔ کوچکتری سر و کار داشته باشیم، این اختلاف کمتر می‌شود. برای نمونه، وقتی با مثلث‌های به‌اندازهٔ کافی کوچک سَر و کار داشته باشیم، بستگی بین ضلع‌ها و زاویه‌های آن را می‌توان،

با دقت کافی، به‌وسیلهٔ دستورهای عادی مثلثاتی بیان کرد؛ درضمن، هرچه مثلث کوچکتر باشد، این دقت بیشتر است.

دستور (۲) نشان می‌دهد، برای شعاع‌های کوچک، طول محیط دایره با دقت خوبی، متناسب با شعاع آن است. به همین ترتیب، برای مثلث‌هایی که به‌اندازهٔ کافی کوچک باشند، مجموع زاویه‌ها، خیلی کم با دو قائمه اختلاف دارد.

در دستور محاسبهٔ محیط دایره، عدد ثابت k وجود دارد که با توجه به واحد طول معین می‌شود. اگر شعاع دایره نسبت به k کوچک باشد، یعنی اگر $\frac{r}{k}$ عدد کوچکی باشد، آن‌وقت همان‌طور که از دستور (۲) دیده می‌شود، طول l نزدیک به $2\pi r$ است. به‌طور کلی، هرچه اندازه‌های شکل نسبت به این مقدار ثابت کمتر باشد، ویژگی‌های شکل، به ویژگی‌های شکل متناظر خود در هندسهٔ اقلیدسی نزدیکتر می‌شود[1].

معیار انحراف ویژگی‌های شکل‌های هندسهٔ *لباچوسکی* نسبت به ویژگی‌های شکل‌های هندسهٔ اقلیدسی، عبارت است از نسبت $\frac{r}{k}$، به‌شرطی که r را معرف اندازهٔ شکل بدانیم (شعاع دایره، ضلع‌های مثلث وغیره).

از این‌جا به نتیجهٔ مهمی می‌رسیم.

فرض کنید با فضای واقعی سر و کار داشته باشیم و فاصله‌ها را با واحد کیلومتر اندازه بگیریم. ثابت k را خیلی بزرگ و برای مثال 10^{12} می‌گیریم.

---

۱. مثال. اگر a ، b و c طول ضلع‌های مجاور به زاویهٔ قائمه و وتر یک مثلث قائم‌الزاویه باشد، به‌جای رابطهٔ ***فیثاغورس*** داریم:

$$2\left(e^{\frac{c}{k}}+e^{-\frac{c}{k}}\right)=\left(e^{\frac{a}{k}}+e^{-\frac{a}{k}}\right)\left(e^{\frac{b}{k}}+e^{-\frac{b}{k}}\right)$$

که اگر به‌صورت رشته باز کنیم، چنین می‌شود:

$$c^2+\frac{c^4}{12k^2}+\cdots=a^2+b^2+\frac{a^4+6a^2b^2+b^4}{12k^2}+\cdots$$

به‌نحوی که به‌ازای عددهای بزرگ k به‌صورت $c^2=a^2+b^2$ یعنی قضیهٔ ***فیثاغورس*** درمی‌آید. همچنین، بنابر دستور لباچوسکی برای محاسبهٔ زاویهٔ توازی γ(پاورقی قبلی را ببینید)، داریم: $\tan\frac{\gamma}{2}=e^{-h/k}$. اگر $\frac{h}{k}$ کوچک، یعنی اگر خط‌های موازی نزدیک به هم باشند، آن‌وقت $\tan\frac{\gamma}{2}=e^{-h/k}\cong 1$ و $\gamma=90^\circ$. بنابراین، برای فاصله‌های کوچک، خط‌های موازی در صفحهٔ ***لباچوسکی***، خیلی کم با هندسهٔ اقلیدسی فرق دارند.

در این صورت، بنا بر دستور (۲)، برای محیط دایرهٔ با شعاع برابر ۱۰۰ کیلومتر، نسبت محیط به شعاع، کمتر از یک‌میلیاردم، با 2π اختلاف دارد. انحراف نسبت به اندازهٔ هندسهٔ اقلیدسی، در سایر زمینه‌ها هم، همین گونه خواهد بود. این انحراف در محدودهٔ یک کیلومتر، در ردیف $\frac{1}{k}$ یعنی 10^{-12} و در محدودهٔ یک متر، در ردیف 10^{-15}، یعنی به‌کلی ناچیز است. چنین انحراف‌هایی از هندسهٔ اقلیدسی، نمی‌تواند و نباید به‌حساب آید، زیرا برای نمونه، اندازهٔ یک اتم صد بار از این انحراف بیشتر است (اندازهٔ هر اتم به‌تقریب 10^{-13} کیلومتر است). از طرف دیگر، در محاسبه‌های مربوط به اخترشناسی، نسبت $\frac{r}{k}$ ممکن است چندان کوچک نباشد.

به همین دلیل، ***لباچوسکی*** می‌پذیرد که، در مقیاس‌های عادی، هندسهٔ اقلیدسی درست است و به نتیجه‌های دقیق می‌رسد؛ انحراف از هندسهٔ اقلیدسی وقتی پیش می‌آید که با مقیاس‌های نجومی سر و کار داشته باشیم. گرچه این فرض ***لباچوسکی*** درست است، ولی انحراف‌های جزئی از هندسهٔ اقلیدسی که امروزه در مقیاس‌های نجومی پیدا می‌شود، بغرنج‌تر از آن است که به این سادگی قابل توضیح باشد.

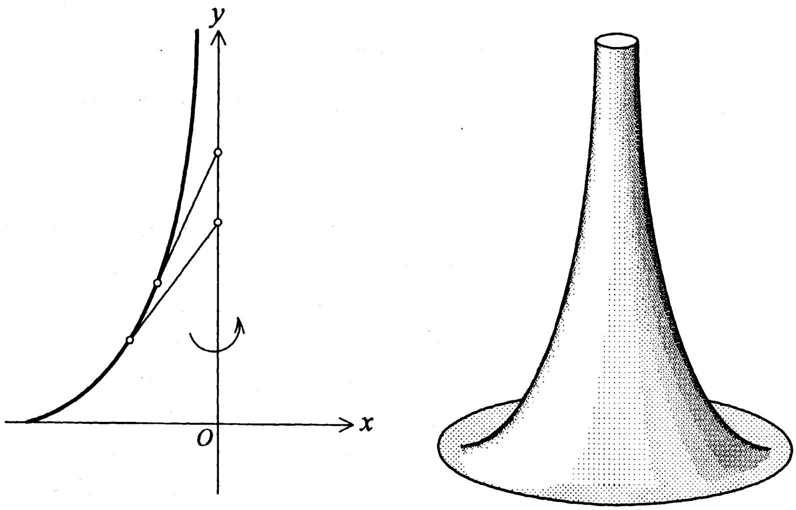
از آن‌چه گفتیم، نتیجهٔ مهم دیگری هم به‌دست می‌آید. چون انحراف از هندسهٔ اقلیدسی، با بزرگتر شدن ثابت k، کمتر می‌شود، با بزرگ کردن k به سمت بی‌نهایت، هندسهٔ ***لباچوسکی*** به سمت هندسهٔ اقلیدسی میل می‌کند. یعنی ***هندسهٔ اقلیدسی، چیزی جز حالت حدی هندسهٔ لباچوسکی*** نیست. بنابراین، اگر این حالت حدی را هم به هندسهٔ لباچوسکی بیفزاییم، آن‌وقت شامل هندسهٔ اقلیدسی هم می‌شود و به این مفهوم، به نظریهٔ عام‌تری تبدیل می‌شود. در این رابطه، ***لباچوسکی*** نظریهٔ خود را «هندسهٔ عام» می‌نامد، یعنی هندسه‌ای کلی. چنین وضعی همیشه، چه در ریاضیات و چه در دانش‌های طبیعی پیش می‌آید. نظریهٔ قدیمی به‌عنوان حالتی حدی از نظریهٔ تازه درمی‌آید و نتیجه‌گیری‌ها از حالت‌های جزئی به حالت‌های کلی می‌رسد.

ولی همهٔ این‌ها و همهٔ این‌گونه نتیجه‌گیری‌ها نوعی بازی فکری و غیرقابل فهم می‌نماید، مگر این که بتوانیم با روشی در حد امکان ساده، برای هندسهٔ ***لباچوسکی***، تعبیری در دستگاه عادی هندسهٔ اقلیدسی پیدا کنیم. این مسأله را خود ***لباچوسکی*** تا پایان حل نکرد؛ او تنها به بخشی از این موضوع پرداخت. این مسأله را چگونه می‌توان حل کرد، چیزی است که در این جا به آن می‌پردازیم.

۴. مفهوم واقعی هندسهٔ لباچوسکی

۱. نخستین تعبیر عینی هندسهٔ *لباچوسکی* را *بلترامی* هندسه‌دان ایتالیایی در سال ۱۸۶۸ میلادی ارائه داد. او متوجه شد، هندسهٔ درونی یک سطح خاص ـ شبه‌کره (یا کره نما) ـ با هندسهٔ تکه‌ای از صفحهٔ *لباچوسکی* تطبیق می‌کند. به‌یاد بیاوریم، منظور از هندسهٔ درونی یک سطح، عبارت است از مجموعهٔ ویژگی‌های شکل‌ها در روی آن، که تنها به یاری اندازهٔ طول‌ها در روی خود سطح، تعریف می‌شود. در شکل ۱۰ (سمت چپ) یک منحنی داده شده است که آن را *تراکتریس* گویند. این منحنی دارای این ویژگی است که طول پاره‌خط‌های راست مماس بر آن، در فاصلهٔ از نقطهٔ تماس تا محل برخورد آن با محور Oy، برای همهٔ نقطه‌های منحنی، مقداری ثابت است. محور Oy، مجانب آن است. اگر منحنی تراکتریس را دور مجانب‌خود دوران دهیم، شکلی به‌دست می‌آیدکه درشکل ۱۰ (سمت راست) نشان داده شده است و **شبه‌کره** نام دارد.

تفسیر هندسهٔ *لباچوسکی* به‌وسیلهٔ *بِلترامی* به این‌جا رسید که، همهٔ رابطه‌های هندسی در تکه‌ای از صفحهٔ *لباچوسکی* با رابطه‌های هندسی روی تکه مناسبی از شبه‌کره تطبیق می‌کند، اگر شرط زیر را در نظر داشته باشیم: نقش پاره‌خط‌های راست را، در این‌جا، خط‌های **ژئودزیک** ـ یعنی کوتاه‌ترین خط‌های روی سطح که دو نقطه از آن را به هم وصل می‌کند ـ



شکل ۱۰

به‌عهده داشته باشند. فاصلهٔ بین دو نقطه از یک سطح، با طول کوتاه‌ترین خط (خط ژئودزیک) که این دو نقطه را در روی سطح به هم وصل می‌کند، تعریف می‌شود. دو شکل را وقتی برابر به‌حساب می‌آورند که بتوان نقطه‌های آن‌ها را طوری روی هم قرار داد که فاصله‌های درونی نقطه‌های متناظر آن‌ها، با هم برابر باشند. جابه‌جایی شکل روی شبه‌کره، اندازه‌های آن را از دیدگاه هندسهٔ درونی حفظ می‌کند، ولو این‌که همراه با خم و پیچی باشد که معرف حرکت روی صفحهٔ *لباچوسکی* است. طول‌ها، زاویه‌ها و مساحت‌ها، به‌طور عادی روی سطح اندازه‌گیری می‌شوند و پاسخ‌گوی طول‌ها، زاویه‌ها و مساحت‌ها در هندسهٔ *لباچوسکی* است.

تفسیر *بِلترامی* نشان می‌دهد، با توجه به این شرط‌ها، هر گزاره‌ای از هندسهٔ *لباچوسکی* که مربوط به تکه‌ای از صفحه باشد، به حقیقت مستقیمی از هندسهٔ درونی شبه‌کره پاسخ می‌دهد. به این ترتیب، هندسهٔ *لباچوسکی* با مفهوم واقعی خود ظاهر می‌شود: این هندسه، چیزی جز نمایش انتزاعی هندسه روی شبه‌کره نیست.

باید یادآوری کرد، هندسهٔ درونی شبه‌کره، ۳۰ سال پیش از کشف آن به‌وسیلهٔ *بلترامی*، به‌وسیلهٔ **ف. مین‌دینگ** هم بررسی شده بود که درواقع، از ویژگی‌هایی نام می‌برد که با هندسهٔ *لباچوسکی* تطبیق می‌کند. ولی نه خودِ او و نه دیگران به این نکته اشاره‌ای نکردند؛ *بلترامی* بود که با مقایسهٔ نتیجه‌گیری‌های *لباچوسکی* و **مین‌دینگ**، بستگی بین آن‌ها را کشف کرد.

کشف *بلترامی* دیدگاه ریاضی‌دانان را نسبت به هندسهٔ *لباچوسکی* تغییر داد: در هندسهٔ «تخیلی» او، واقعیت را یافتند[1].

۲. با همهٔ این‌ها، همان‌طور که روی نمونهٔ شبه‌کره تأکید کردیم، هندسهٔ *لباچوسکی*، نه برای تمامی صفحهٔ خود، بلکه تنها برای تکه‌ای از آن قابل تفسیر است و با واقعیت تطبیق

---

۱. تاریخ قوام گرفتن مفهوم واقعی هندسهٔ *لباچوسکی*، بغرنج‌تر از این‌هاست. در درجهٔ اول، خود *لباچوسکی*، عدم تناقض کار خود را با اثبات برمبنای مُدل تحلیلی به‌طور مستقیم نشان داد، ولی نتوانست اثبات خود را تا پایان ادامه دهد. این اثبات، خیلی دیرتر به‌دست آمد. دوم، *ریمان* ریاضی‌دان آلمانی در سال ۱۸۵۴ ، نظریه‌ای آورد (بند ۱۰ را ببینید) که شامل همان نتیجه‌گیری‌های *بلترامی* بود. ولی *ریمان* نتیجه‌گیری‌های خود را خیلی روشن ننوشته بود و رساله‌ای کم و بیش غیرقابل فهم ارائه داد. این رساله بعد از مرگ ریمان در سال ۱۸۶۸ چاپ شد، وقتی که کار *بلترامی* منتشر شده بود. به‌طور کلی، سراسر تاریخ هندسهٔ *لباچوسکی*، از جست‌وجوی اثبات پوستولای *اقلیدس* تا روشنی کامل معنای هندسهٔ نااقلیدسی، از این جهت بی‌اندازه آموزنده است که نشان می‌دهد چه نیروهایی باید صرف شود و چه راه‌های غیرمستقیم و گاه دشواری باید پیموده شود تا حقیقتی، سرانجام به‌شکل ساده و قابل فهم خود عرضه شود.

می‌کند[۱]. بنابراین، مسألهٔ روشن کردن مفهوم واقعی هندسهٔ *لباچوسکی* در تمامی صفحه و به‌ویژه در فضا، همچنان حل‌نشده باقی می‌ماند. ولی این مسأله هم، خیلی زود و در سال ۱۸۷۰ ، به‌وسیلهٔ *کلاین* ریاضی‌دان آلمانی حل شد. ببینیم این راه‌حل چگونه بود؟

روی صفحهٔ عادی اقلیدسی، دایره‌ای انتخاب می‌کنیم و تنها درون آن را در نظر می‌گیریم، یعنی محیط و بخش بیرونی دایره را کنار می‌گذاریم. این بخش درونی دایره را «صفحه» می‌نامیم که نقش صفحه *لباچوسکی* را به‌عهده دارد. وترهای این دایره را «خط‌های راست» می‌نامیم؛ درضمن بنابر شرط، دو انتهای هر وتر را، که روی محیط دایره قرار دارند، از آن کنار می‌گذاریم. سرانجام، هرگونه تبدیلی از دایره که آن را به خودش و خط‌های راست را به خط‌های راست تبدیل کند، یعنی وترهای آن را تغییر ندهد، «حرکت» می‌نامیم. ساده‌ترینِ این‌گونه تبدیل‌ها، دوران دایره دور مرکز آن است، ولی درواقع، تعداد این‌گونه تبدیل‌ها بسیار زیاد است. دربارهٔ چگونگی این تبدیل‌ها صحبت می‌کنیم.

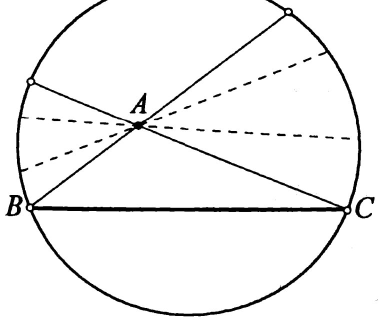
اگر این نام‌گذاری‌ها را بپذیریم، معلوم می‌شود، حقیقت‌های هندسهٔ معمولی درون دایرهٔ ما، منجر به قضیه‌های هندسهٔ *لباچوسکی* می‌شود. و برعکس: هر قضیه‌ای از هندسهٔ *لباچوسکی*، همچون حقیقتی از هندسهٔ عادیِ درون دایره تفسیر می‌شود.

از جمله، بنا بر اصل *لباچوسکی*، از نقطه‌ای که بر خط راست مفروض واقع نیست، می‌توان دست‌کم دو خط راست رسم کرد که خط راست مفروض را قطع نکنند. این اصل را به زبان هندسهٔ معمولی و سازگار با شرطی که انتخاب کرده‌ایم، می‌دهیم، یعنی به‌جای خط‌های راست، وترها را در نظر می‌گیریم. در این صورت به این گزاره می‌رسیم: از نقطهٔ واقع در درون دایره، که بر وتر مفروضی واقع نباشد، می‌توان دست‌کم دو وتر رسم کرد که با وتر مفروض، برخورد نداشته باشند. درستی این گزاره، با توجه به شکل ۱۱ روشن است.

---

۱. شبه‌کره در همه‌جا، انحنای گوسی منفی یکنواختی دارد. تمامی سطح‌های با انحنای منفی ثابت، دست‌کم در تکه‌های کوچکی از خود، دارای هندسهٔ درونی یکسانی هستند و بنابراین هر کدام می‌تواند برای بیان هندسهٔ *لباچوسکی* به‌کار آید. ولی همان‌طور که *هیلبرت* در سال ۱۹۰۱ ثابت کرد، هیچ‌یک از این سطح‌ها نمی‌تواند، بدون تکینی، از همه‌طرف، تا بی‌نهایت امتداد پیدا کند و درنتیجه نمی‌تواند مدلی برای *تمامی* صفحهٔ *لباچوسکی* باشد. از طرف دیگر، *کوییر*، ریاضی‌دانان جوان هلندی، در سال ۱۹۵۵ ثابت کرد، سطح‌های هموار وجود دارند که به مفهوم هندسهٔ درونی خود، تمامی صفحهٔ *لباچوسکی* را نمایش می‌دهند، ولی این‌گونه سطح‌ها، با وجود هموار، نمی‌تواند دارای خمیدگی پیوسته باشند؛ این سطح‌ها، انحنای معینی ندارند.

این را هم یادآور شویم که، ضمن نمایش هندسهٔ *لباچوسکی* روی سطح با انحنای منفی ثابت K ، ثابت k در دستورهای بند قبل، مفهوم ساده‌ای پیدا می‌کند: $k^2=-\frac{1}{K}$.



شکل ۱۱

بنابراین، اصل *لباچوسکی* در این‌جا محقق است.

سپس، به‌یاد می‌آوریم، در هندسهٔ *لباچوسکی*، بین خط‌های راستی که از نقطهٔ مفروض می‌گذرند و خط راست مفروض را قطع نمی‌کنند، دو خط راست مرزی وجود دارد که، به‌ویژه *لباچوسکی*، تنها آن‌ها، موازی با خط راست مفروض می‌نامد. و این، به‌معنای آن است که بین وترهایی که از نقطهٔ مفروض A می‌گذرند و وتر مفروض BC را قطع نمی‌کنند، دو وتر مرزی وجود دارد. و درواقع، این دو وتر مرزی، یکی به نقطهٔ B و دیگری به نقطهٔ C نزدیک می‌شود. این دو وتر با وتر BC نقطهٔ مشترکی ندارند، زیرا نقطه‌های واقع بر محیط دایره را کنار گذاشته بودیم. بنابراین، این قضیهٔ *لباچوسکی* هم در این‌جا برقرار است.

برای تفسیر بعدی قضیه‌های *لباچوسکی* به زبان هندسهٔ عادی درون دایره، باید روشن کنیم، در دایره، چگونه باید پاره‌خط‌های راست و زاویه‌ها را اندازه گرفت تا این اندازه‌گیری پاسخ‌گوی هندسهٔ *لباچوسکی* باشد. البته این اندازه‌گیری نمی‌تواند به مفهوم عادیِ آن باشد، زیرا به مفهوم عادی، وتر طولی محدود و خط راست، که در این‌جا با وتر مشخص می‌شود، طولی بی‌نهایت دارد. به همین جهت، ممکن است در این‌جا نوعی تضاد دیده شود، ولی خواهیم دید که هیچ‌گونه تضاد و تناقضی وجود ندارد.

به‌یاد بیاوریم، طول پاره‌خط‌های راست را چگونه اندازه می‌گیریم! طول پاره‌خط راستی مثل AB را به‌عنوان واحد انتخاب می‌کنیم و سپس، طول هر پاره‌خط راستی مثل XY را، ضمن مقایسه با پاره‌خط راست AB، اندازه می‌گیریم. درضمن، پاره‌خط راست AB را در طول پاره‌خط راست XY جابه‌جا می‌کنیم. اگر در این جابه‌جایی، بخشی از XY، که طولی کمتر از پاره‌خط راست AB دارد، باقی بماند، آن‌وقت پاره‌خط راست AB در مثل به ۱۰

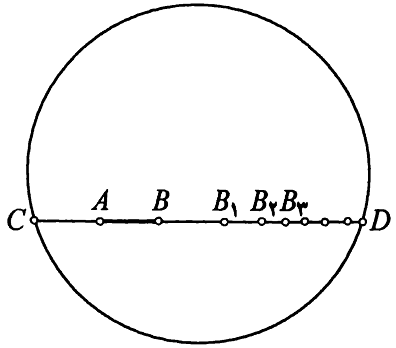
بخش برابر تقسیم می‌کنیم (برابر به این معنا که هرکدام از حرکت دیگری به‌دست آید)، این بخش‌ها را روی قسمت باقی‌ماندهٔ XY می‌گذاریم؛ سپس اگر لازم بود، پاره‌خط راست AB را به ۱۰۰ بخش برابر تقسیم می‌کنیم وغیره. درنتیجه، طول پاره‌خط راست XY، به‌صورت یک کسر دهدهی، که ممکن است بی‌پایان هم باشد، نشان داده می‌شود. به این ترتیب، اندازه‌گیری طول از راه جابه‌جایی پاره‌خط راست کامل، که به‌عنوان واحد انتخاب شده است ویا بخشی از آن به‌دست می‌آید، یعنی اندازه‌گیری براساس حرکت شکل می‌گیرد. ولی حرکت را تعریف کرده‌ایم (هر حرکت، به‌معنای تبدیلی از دایره است که خط‌های راست را به خط‌های راست منجر کند)، بنابراین معلوم می‌شود، کدام پاره‌خط‌های راستی برابرند و طول را چگونه باید اندازه گرفت. به این ترتیب، تعریف حرکت، گرچه به صورتی ناروشن، قانون اندازه‌گیری طول را در بر می‌گیرد. به همین ترتیب، زاویه‌ها را هم می‌توان با انتخاب زاویه‌ای به‌عنوان واحد، اندازه گرفت. بنابراین، قانون اندازه‌گیری زاویه هم، در تعریف حرکت وجود دارد.

قانون‌های اندازه‌گیری طول و زاویه، به‌نحوی که پاسخ‌گوی هندسهٔ ***لباچوسکی*** باشد، بسیار ساده است، اگرچه در ماهیت خود با تعریف‌های عادی تفاوت دارند. ما به نتیجه‌گیری از این قانون‌ها نمی‌پردازیم، زیرا این موضوع، برای بحث، اهمیت جدی ندارد[1].

قانون اندازه‌گیری طول به‌صورتی است که وتر، طولی برابر با بی‌نهایت دارد. و این به آن مناسبت است که، اگر به‌یاری تبدیلی که به‌عنوان حرکت پذیرفته‌ایم، پاره‌خط راست AB را در پاره‌خط راست BB_1، سپس در پاره‌خط راست B_1B_2 وغیره ببریم، آن‌وقت پاره‌خط راست B_kB_{k+1} را به‌دست می‌آوریم که، به‌مفهوم عادی، خیلی کوتاه‌تر است (گرچه به‌مفهوم مُدل هندسهٔ ***لباچوسکی*** ما، همهٔ این پاره‌خط‌های راست با هم برابرند؛ شکل ۱۲). نقطه‌های B_1، B_2، ...، B_k، ... با نزدیک شدن به انتهای وتر به هم نزدیک‌تر و متمرکزتر می‌شوند. ولی وتر ما، انتها ندارد: بنابر شرط انتهای آن را کنار گذاشته‌ایم و، بنابراین، به این مفهوم، بی‌پایان است. به‌مفهوم هندسهٔ ***لباچوسکی***، نقطه‌های B_1، B_2، ...، در جایی به هم نزدیک و متمرکز نمی‌شوند، آن‌ها به‌سمت بی‌نهایت می‌روند. از راه تبدیلی که به‌عنوان حرکت پذیرفته‌ایم، وقتی پاره‌خط‌های راست برابر را بر هم قرار می‌دهیم، نمی‌توان از درون

---

۱. قانون اندازه‌گیری طول، به این ترتیب است. پاره‌خط راست AB را روی وتر CD بگیرید (شکل ۱۲). پاره‌خط راست را با روش معمول اندازه می‌گیریم و به‌اصطلاح نسبت مرکب $\frac{CB}{CA}:\frac{DB}{DA}$ را تشکیل می‌دهیم؛ و لگاریتم آن را به‌عنوان طول پاره‌خط راست AB انتخاب می‌کنیم.



شکل ۱۲

دایره به محیط آن رسید.

برای این که چگونگی روی هم گذاشتن پاره‌خط‌های راست را در مُدل بهتر بفهمیم، تبدیلی را بررسی می‌کنیم که نقش جابه‌جایی در طول خط راست را به‌عهده دارد.

روی صفحه، دستگاه مختصات قائم را، که در آن مبداء مختصات در مرکز دایره است، در نظر می‌گیریم. برای مشخص بودن وضع، طول شعاع دایره را واحد می‌گیریم، به‌نحوی که معادلهٔ آن به‌صورت $x^2+y^2=1$ درمی‌آید و هر نقطهٔ درونی آن با نامعادلهٔ $x^2+y^2<1$ سازگار است.

تبدیلی را در نظر می‌گیریم که با این دستورها داده شده است:

$$x'=\frac{x+a}{1+ax} \quad , \quad y'=\frac{\sqrt{1-a^2}}{1+ax} \qquad (3)$$

که در آن x' و y'، مختصات نقطه بعد از تبدیل؛ x و y مختصات نخستینِ نقطه و a عدد مفروضی از لحاظ قدرمطلق کوچکتر از واحد است.

از دستورهای (۳) می‌توان به‌سادگی دستورهای عکس را، یعنی مقدارهای x و y را برحسب x' و y' پیدا کرد:

$$x=\frac{x'-a}{1-ax'} \quad , \quad y=\frac{y'\sqrt{1-a^2}}{1-ax'} \qquad (4)$$

تبدیل (۳) با دو شرطی که برای «حرکت» در مُدل خود داشتیم، سازگار است: ۱) دایره

را به خودش تبدیل می‌کند؛ ۲) خط‌های راست را به خط‌های راست منجر می‌کند.

برای اثبات ویژگی اول باید روشن کنیم که نابرابری یا برابری $x^2+y^2 \leq 1$، پس از تبدیل منجر به $x'^2+y'^2 \leq 1$ می‌شود و برعکس. برای نمونه ثابت می‌کنیم برای $x^2+y^2=1$، بی‌تردید به $x'^2+y'^2=1$ می‌رسیم، یعنی نقطه‌های واقع بر محیط دایره، باز هم روی محیط باقی می‌مانند.

با توجه به دستورهای (۳)، $x'^2+y'^2$ را محاسبه می‌کنیم و سپس قرار می‌دهیم $x^2+y^2=1$، یعنی $y^2=1-x^2$:

$$x'^2+y'^2=\frac{(x+a)^2+y^2(1-a^2)}{(1+ax)^2}=\frac{(x+a)^2+(1-x^2)(1-a^2)}{(1+ax)^2}=$$

$$=\frac{x^2+2ax+a^2+1-x^2-a^2+a^2x^2}{(1+ax)^2}=\frac{1+2ax+a^2x^2}{1+2ax+a^2x^2}=1$$

بنابراین برای $x^2+y^2=1$، باز هم به‌دست می‌آید: $x'^2+y'^2=1$. حالت‌های دیگر را هم، می‌توان به همین ترتیب تحقیق کرد.

ویژگی دوم دستورهای (۳) هم به همین سادگی ثابت می‌شود. می‌دانیم هر خط راست با یک معادلهٔ خطی معرفی می‌شود و، برعکس، هر معادلهٔ خطی نمایندهٔ یک خط راست است. این خط راست را در نظر می‌گیریم:

$$Ax+By+C=0 \qquad (۵)$$

بعد از تبدیل (۴) به‌دست می‌آید:

$$A\frac{x'-a}{1-ax'}+B\frac{y'\sqrt{1-a^2}}{1-ax'}+C=0$$

که اگر مخرج‌ها را از بین ببریم:

$$(A-aC)x'+B\sqrt{1-a^2}\,y'+(C-aA)=0$$

این، یک معادلهٔ خطی و مُعرف یک خط راست است؛ خط راست (۵)، بعد از تبدیل، به این خط راست منجر می‌شود.

به این نکته هم توجه کنیم که تبدیل (۳) محور Ox را به خودش منجر می‌کند، تنها جای نقطه‌ها در روی آن جابه‌جا می‌شود. این مطلب روشن است، زیرا روی این محور داریم $y = ۰$، که با توجه به دستور (۳) به‌دست می‌آید $y' = ۰$. برای محور Ox، از دستور تبدیل، به‌دست می‌آید:

$$x' = \frac{x + a}{۱ + ax} \quad (|a| < ۱) \qquad (۳')$$

روی این خط راست، پاره‌خط راست $x_۱x_۲$، بنا به دستور $(۳')$، به پاره‌خط راست $x'_۱x'_۲$ منجر می‌شود، و بنابر شرط، این دو پاره‌خط راست با هم برابرند. به این ترتیب، «روی هم گذاشتن پاره‌خط‌های راست» هم انجام می‌گیرد.

برای نقطهٔ O، مرکز دایره، داریم $x = ۰$ و متناظر با آن $x' = a$، یعنی ضمن تبدیل $(۳')$، مرکز به نقطهٔ A با مختص $x = a$ منتقل می‌شود.

از آن‌جا که a می‌تواند هر عدد دل‌خواه، با شرط $|a| < ۱$ باشد، مرکز دایره می‌تواند به هر نقطه‌ای روی قطر در طول محور Ox برود.

در همین تبدیل، نقطه‌ای که بر A قرار داشت، به نقطهٔ $A_۱$ با مختص

$$x_۱ = \frac{a + a}{۱ + a^۲} = \frac{۲a}{۱ + a^۲}$$

می‌رود. بنابراین، پاره‌خط راست OA، ضمن تبدیل (۳)، منجر به پاره‌خط راست $AA_۱$ می‌شود که ناشی از «قرار گرفتن» این پاره‌خط راست بر «خط راستی» است که مُعرف قطر دایره است.

با تکرار این تبدیل، می‌توانیم همین پاره‌خط راست را، هر چند بار که مایل باشیم، دورتر ببریم. در این صورت نقطهٔ A_n با مختص x_n به نقطهٔ $A_{n+۱}$ با مختص

$$x_{n+۱} = \frac{x_n + a}{۱ + ax_n}$$

می‌رسد. به این ترتیب، نقطه‌های A، $A_۱$، $A_۲$، ... با مختص‌های

$$x_۰ = a \quad , \quad x_۱ = \frac{۲a}{۱ + a^۲} \quad , \quad x_۲ = \frac{x_۱ + a}{۱ + ax_۱} = \frac{۳a + a^۳}{۱ + ۳a^۲} \quad , \quad \ldots$$

به‌دست می‌آید. چون همهٔ این پاره‌خط‌های راست $A_n A_{n+۱}$ از OA و با تبدیل معرف

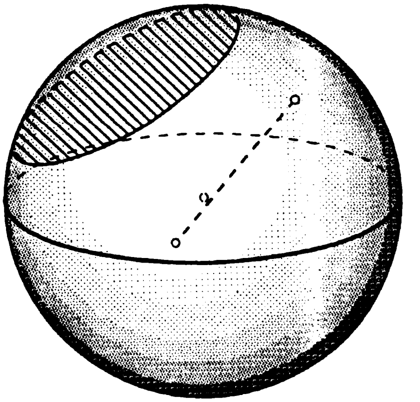
حرکت به‌دست می‌آیند، همهٔ آن‌ها با هم برابرند ـ البته، برابر به مفهوم هندسهٔ ***لباچوسکی*** و به همان صورتی که روی مُدل نشان داده شده است. به‌سادگی ثابت می‌شود که نقطه‌های A_n، در انتهای قطر غلیظ و غلیظ‌تر می‌شوند؛ و به مفهوم مُدل تا بی‌نهایت ادامه می‌یابند.

چون محور Ox را می‌توان در هر جهتی انتخاب کرد، بنابراین چنین تبدیل‌هایی را می‌توان در طول هر قطری انجام داد. اگر این تبدیل‌ها را، با دوران دور مرکز دایره و تصویرهای بر قطر ترکیب کنیم، همهٔ «حرکت‌ها»، آن‌طور که از مُدل فهمیده می‌شود، به‌دست می‌آید؛ این حرکت‌ها، شامل انتقال، دوران و تصویر انعکاسی‌اند. تفصیل این مطلب را در بند بعد می‌بینیم. در آن‌جا بادقت ثابت می‌کنیم که، درواقع، هندسهٔ ***لباچوسکی*** با مُدل ما سازگار است و به‌ویژه، تبدیل ناشی از «حرکت»، با همهٔ شرط‌ها (اصل موضوع‌ها) می‌سازد، اصل موضوع‌هایی که از مفهوم حرکت در هندسه پیروی می‌کنند.

دوباره به مدلی که ***کلاین*** برای هندسهٔ ***لباچوسکی*** پیشنهاد می‌کند، برمی‌گردیم. به‌عنوان صفحه، درون دایره در نظر گرفته می‌شود؛ نقطه عبارت است از هر نقطه‌ای در صفحه، خط راست یعنی وتری از دایره (که دو انتهای آن کنار گذاشته شده است)، حرکت به‌معنای تبدیلی است که دایره را به خودش و وترها را به وترها منجر می‌کند؛ ردیف نقطه‌ها (نقطه‌ای که بر خط راست واقع است؛ نقطه‌ای که بین دو نقطهٔ دیگر قرار دارد) با همان مفهوم عادی پذیرفته می‌شود. قانون اندازه‌گیری طول و زاویه‌ها (و همچنین مساحت‌ها) از تعریف حرکت ناشی می‌شود و به همین ترتیب برابری پاره‌خط‌های راست و زاویه‌ها (و هر شکلی) معین می‌شود، عمل قرار دادن یک پاره‌خط راست در طول دیگری هم، از همین‌جا تعریف می‌شود. ***با این شرط‌ها، تمامی نظریهٔ هندسهٔ لباچوسکی بر صفحه، متناظر با هندسهٔ اقلیدسی درونِ دایره می‌شود و، برعکس، هر حقیقتی که از هندسهٔ اقلیدسی در درون یک دایره به‌دست آید، پاسخ‌گوی قضیه‌ای از هندسهٔ لباچوسکی است.***

مُدل هندسهٔ ***لباچوسکی*** در فضا هم، به همین ترتیب به‌دست می‌آید. به‌عنوان فضا، درونِ یک کره (شکل ۱۳)، به‌عنوان خط راست، وتری از این کره، به‌عنوان صفحه، سطح دایره‌ای که محیط آن بر سطح کره واقع است، در نظر گرفته می‌شود. درضمن، خود سطح کره، یعنی دو انتهای هر وتر و محیط هر دایرهٔ واقع بر سطح کره، کنارگذاشته می‌شود؛ سرانجام حرکت به‌عنوان تبدیلی تعریف می‌شود که کره را به خودش و هر وتر را به وتر دیگری تبدیل کند.

وقتی این مُدل، برای هندسهٔ ***لباچوسکی*** داده شد، تعبیر واقعی ساده‌ای برای آن پیدا شد: هندسهٔ ***لباچوسکی*** از این جهت حقیقت دارد که می‌توان آن را همچون طرح هندسه در داخل



شکل ۱۳

دایره یا کره در نظر گرفت. همین وضع، بی‌تناقضی آن را هم ثابت می‌کند. نتیجه‌گیری‌های آن، نمی‌تواند منجر به تناقض شود، زیرا همهٔ این نتیجه‌ها را می‌توان با زبان هندسهٔ اقلیدسی معمولی در درون دایره به‌دست آورد (ویا در درون کره، اگر صحبت بر سر هندسهٔ *لباچوسکی* در فضا باشد[1]).

۳. بعد از *کلاین*، *پوانکاره* ریاضی‌دان فرانسوی، مُدل دیگری برای هندسهٔ *لباچوسکی* داد که از آن برای نتیجه‌گیری‌های مهمی در نظریهٔ تابع‌های با متغیر مختلط استفاده کرد[2]. به این ترتیب، *پوانکاره* از هندسهٔ *لباچوسکی* برای حل دشواری‌هایی استفاده می‌کند که در شاخهٔ دیگری از ریاضیات قرار دارند. هندسهٔ *لباچوسکی* کاربردهای دیگری، هم در ریاضیات و هم در فیزیک نظری پیدا کرد؛ از جمله *واری‌چک* فیزیک‌دان در سال ۱۹۱۳ کاربردی از این هندسه را در نظریهٔ نسبیت ارائه داد.

هندسهٔ *لباچوسکی* باموفقیت پیش می‌رود و در آن شاخه‌های تازه‌ای شکل می‌گیرد: نظریهٔ ساختمان‌های هندسی، نظریهٔ عمومی منحنی‌ها و سطح‌ها، نظریهٔ جسم‌های گوژ و غیره.

---

۱. معمول است که ریاضی‌دانان می‌گویند، هندسهٔ *لباچوسکی*، در هندسهٔ اقلیدسی قابل نمایش است و بنابراین، به‌همان اندازه که هندسهٔ اقلیدسی بی‌تناقض‌است، می‌توان هندسهٔ *لباچوسکی* را هم‌بی‌تناقض دانست.

۲. مدل *پوانکاره*، سرانجام به این‌جا می‌رسد که، به‌جای صفحهٔ *لباچوسکی*، همان درون دایره در نظر گرفته شود، ولی برای خط‌های راست، کمان‌هایی از دایره را در نظر می‌گیرد که بر محیط دایرهٔ مفروض عمودند؛ به‌عنوان حرکت هم، هر تبدیل همدیسی را در نظر می‌گیرد که دایره را به خودش منجر کند. این رابطهٔ با تبدیل‌های همدیس (کُنفُرم) منجر به رابطه با نظریهٔ تابع‌های با متغیر مختلط می‌شود.

۵. اصل موضوع هندسه و تحقیق آن‌ها در مُدل مفروض

۱. برای این که با دقت ریاضی ثابت کنیم که مُدل *کلاین* به‌واقع تفسیری از هندسهٔ *لباچوسکی* است، پیش از همه باید بادقت روشن کنیم، چه چیزی را باید ثابت کرد. تحقیق ردیفِ قضیه‌های *لباچوسکی* کاری بی‌معنی است؛ این قضیه‌ها زیادند و درضمن بسیار زیاد و نامحدود، زیرا همیشه می‌توان قضیه‌های تازه و تازه‌تری ثابت کرد. بنابراین، کافی است ثابت کنیم، در مُدل *کلاین*، پیش‌فرض‌های اصلی هندسهٔ *لباچوسکی*، پیش‌فرض‌هایی که همهٔ دیگر گزاره‌ها را می‌توان از آن‌ها نتیجه گرفت، صدق می‌کند. ولی در این صورت، باید در آغاز، خود این پیش‌فرض‌های اصلی را تنظیم کرد.

به این ترتیب، مسألهٔ مربوط به اثبات بی‌تناقضی هندسهٔ *لباچوسکی*، به مسألهٔ مربوط به تنظیم دقیق و کامل پیش‌فرض‌های اصلی آن، یعنی اصل موضوع‌ها (یا آکسیوم‌ها) منجر می‌شود. ولی از آن‌جا که پیش‌فرض‌های هندسهٔ *لباچوسکی* با پیش‌فرض‌های هندسهٔ اقلیدسی، تنها در اصل موضوع مربوط خط‌های راست موازی، با هم اختلاف دارند، مسأله به این‌جا می‌رسد که اصل موضوع‌های هندسهٔ اقلیدسی را بادقت و به‌طور کامل تنظیم کنیم. خود *اقلیدس* به چنین تنظیمی نرسیده است: او به‌ویژه تعریفی برای حرکت یا روی هم گذاشتن شکل‌ها ندارد، گرچه از این مفهوم‌ها استفاده کرده است. مسألهٔ مربوط به دقیق و کامل کردن اصل موضوع‌های *اقلیدس*، بارها به‌وسیلهٔ هندسه‌دانان و به‌ویژه با پیشرفت هندسهٔ *لباچوسکی* مطرح بود و تا پایان سدهٔ نوزدهم ادامه داشت.

سرانجام، با تلاش تعدادی از هندسه‌دانان، مسألهٔ مربوط به تنظیم اصل موضوع‌های هندسه حل شد.

اصل موضوع‌ها را، در حالت کلی، می‌توان متفاوت در نظر گرفت و مفهوم‌های اصلی را به گونه‌های مختلف انتخاب کرد. در این‌جا سیاهه‌ای از اصل موضوع‌های هندسهٔ روی صفحه را می‌آوریم که بر پایهٔ مفهوم‌های اساسی نقطه، خط راست، حرکت؛ همچنین این مفهوم که چگونه نقطهٔ X بر خط راست a واقع است؛ چگونه نقطهٔ B بین دو نقطهٔ A و C قرار دارد؛ چگونه حرکت، نقطهٔ X را به نقطهٔ Y می‌رساند، ساخته شده‌اند (در این صورت، بقیهٔ مفهوم‌ها به‌کمک این‌ها تعریف می‌شوند. از جمله، پاره‌خط راست همچون مجموعهٔ نقطه‌های واقع در بین دو نقطهٔ مفروض قابل تعریف است).

اصل موضوع‌ها از پنج گروه تشکیل شده‌اند.

I. اصل موضوع‌های وقوع (یا برخورد)

۱) از هر دو نقطه، یک خط راست و درضمن، تنها یک خط راست می‌گذرد.

۲) روی هر خط راست، دست‌کم دو نقطه وجود دارد.

۳) دست‌کم سه نقطه وجود دارد که بر یک خط راست واقع نباشند.

II. اصل موضوع‌های ترتیب

۱) از هر سه نقطهٔ واقع بر خط راست، تنها یکی بین دو نقطهٔ دیگر قرار دارد.

۲) اگر A و B، دو نقطه از خط راست باشند، آن‌وقت روی همین خط راست دست‌کم یک نقطهٔ C وجود دارد، به‌نحوی که نقطهٔ B، بین نقطه‌های A و C واقع باشد.

۳) خط راست، صفحه را به دو نیم‌صفحه تقسیم می‌کند (یعنی، همهٔ نقطه‌هایی از صفحه را که بر این خط راست واقع نیستند، به دو رده به‌نحوی تقسیم می‌کند که اگر دو نقطهٔ یک رده را با پاره‌خط راستی به هم وصل کنیم، خط راست را قطع نمی‌کند، ولی پاره‌خط راستی که دو نقطه از دو ردهٔ مختلف را به هم وصل کند، با خط راست برخورد دارد).

III. اصل موضوع‌های حرکت

(حرکت نه همچون تبدیل یک شکل جداگانه، بلکه همچون تبدیل تمامی صفحه فهمیده می‌شود.)

۱) حرکت، خط راست را به خط راست منجر می‌کند.

۲) دو حرکتی که یکی پس از دیگری انجام شود، هم‌ارز با یک حرکت است.

۳) A، A' و a، a' را دو نقطه و دو نیم‌خط راستی می‌گیریم که از آن‌ها آغاز شده‌اند، و α، α' را نیم‌صفحه‌هایی می‌گیریم که به‌وسیلهٔ ادامهٔ خط‌های راست a و a' محدود شده‌اند؛ در این صورت یک حرکت، و تنها یک حرکت، وجود دارد که در آن A به A'، a به a' و α به α' منجر می‌شود (به‌زبان دیگر، نقطهٔ A با انتقال به A' منتقل می‌شود، سپس با دوران نیم‌خط راست a، منجر به a' می‌شود؛ آن‌وقت یا نیم‌صفحهٔ α بر نیم‌صفحهٔ α' منطبق است و یا باید دورانی را دور خط راست a انجام داد تا بر هم منطبق شوند).

IV. اصل موضوع پیوستگی

نقطه‌های X_1، X_2، X_3، ... واقع بر خط راست طوری قرار دارند که هر نقطهٔ بعدی در

سمت راست نقطهٔ قبلی است؛ درضمن، نقطه‌ای مثل A وجود دارد که در سمت راست همهٔ آن‌هاست[1]. در این صورت، نقطه‌ای مثل B وجود دارد که در سمت راست همهٔ نقطه‌های X_1، X_2، ... واقع است، به‌نحوی که نقطهٔ X_n به‌دل‌خواه نزدیک به آن وجود دارد (یعنی نقطه‌ای در سمت چپ B پیدا می‌شود، مثل نقطهٔ C، که نقطهٔ X_n روی پاره‌خط راست CB واقع است).

V. اصل موضوع توازی (اقلیدس)

از نقطهٔ مفروض، تنها یک خط راست می‌توان رسم کرد که خط راست مفروض را قطع نکند.

این‌ها اصل موضوع‌هایی هستند که برای ساختمان هندسهٔ اقلیدسی روی صفحه کفایت می‌کنند. همهٔ قضیه‌های هندسهٔ مسطحه دبیرستانی را می‌توان از این اصل‌ها نتیجه گرفت، آن هم به‌صورتی دقیق.

اصل موضوع‌های هندسهٔ لباچوسکی، تنها در اصل موضوع توازی با این اصل موضوع‌ها اختلاف دارد.

V′. اصل موضوع توازی (لباچوسکی)

از نقطهٔ واقع در بیرون خط راست، دست‌کم دو خط راست می‌گذرد که خط راست مفروض را قطع نمی‌کند.

ممکن است اندکی عجیب به‌نظر آید که بین اصل موضوع‌ها، از جملهِ این هم وجود دارد که: «روی هر خط راست، دست‌کم یک نقطه وجود دارد». مگر نه این است که تصور عادی به ما می‌گوید، روی خط راست، مجموعه‌ای از بی‌نهایت نقطه وجود دارد. عجیب نیست که، نه *اقلیدس* و نه ریاضی‌دان دیگری تا پایان سدهٔ نوزدهم، به ذهنش نرسید چنین اصل موضوعی را تنظیم کند و همه آن را به خودی خود بدیهی دانسته‌اند. ولی، اکنون وضع تغییر کرده است. وقتی تفسیر تازه‌ای از هندسه می‌دهیم، زیر نام خط راست، مفهوم عادی آن را نمی‌فهمیم، بلکه به چیز دیگری توجه داریم: به خط‌های ژئودزیک روی سطح، به وتری از دایره که دو انتهای آن را کنار گذاشته‌ایم ویا چیزی دیگر. بنابراین مسأله تنظیم

۱. «سمت راست» را می‌توان همه‌جا به «سمت چپ» تغییر داد.

دقیقی از همه‌چیز پیش می‌آید، هر چیزی که ممکن است در «اشیا»یی که خط راست می‌نامیم، مورد نیاز باشد. دربارهٔ سایر مفهوم‌ها و اصل موضوع‌ها هم، همین وضع وجود دارد.

همان‌طور که پیش از این هم دیدیم، پیدایی تفسیرهای مختلف هندسه، یکی از مهم‌ترین انگیزه‌ها برای دقیق کردن پایه‌ها وبنیان‌های هندسه بود. ازنظر تاریخی هم وضع به‌همین‌گونه است: تنظیم دقیق اصل موضوع‌ها، بعد از مُدل‌های بلترامی، کلاین و پوانکاره انجام گرفت.

۲. اکنون ثابت می‌کنیم، همهٔ این اصل موضوع‌ها، به‌جز اصل موضوع توازی ***اقلیدس***، در مُدل ***کلاین*** صدق می‌کنند. همان‌طور که در بند قبل (شکل ۱۱) هم دیدیم، در این‌جا، نه اصل توازی ***اقلیدس***، بلکه اصل توازی ***لباچوسکی*** برقرار است. اکنون به تحقیق دربارهٔ اصل موضوع‌های I تا IV می‌پردازیم.

در این مُدل، درون دایره (که شعاع آن را واحد به‌حساب می‌آوریم) نقش صفحه را به‌عهده دارد. نقش نقطه، به‌عهدهٔ نقطه و نقش خط راست به‌عهدهٔ وتر است؛ مفهوم‌های «نقطه بر خط راست واقع است» و «نقطه بین دو نقطهٔ دیگر قرار دارد»، با همان مفهوم‌های عادی درک می‌شود. از این‌جا، نتیجه می‌گیریم که اصل موضوع‌های ترتیبی و اصل موضوع پیوستگی با مُدل ما سازگارند. از جمله، اصل موضوع سوم ترتیبی، به این معنای ساده است که وتر، دایره را به دو بخش تقسیم می‌کند.

این می‌ماند که دربارهٔ اصل موضوع‌های حرکت، آزمایش کنیم. حرکت را به‌معنای تبدیلی گرفتیم که دایره را به خودش و خط راست را به خط راست منجر کند. از این تعریف روشن است که، این‌گونه تبدیل‌ها، با دو اصل موضوع اول حرکت سازگارند: اصل موضوع اول، به این دلیل که خط‌های راست همان وترها هستند و بنابراین تبدیل وتر به وتر به‌معنای تبدیل خط راست به خط راست است؛ اصل موضوع دوم، به این دلیل که اگر دو حرکت را پشت سر هم انجام دهیم، دایره به خودش و وتر به وتر تبدیل می‌شود که خود به‌معنای یک تبدیل یا یک «حرکت» است.

اصل موضوع سوم حرکت باقی می‌ماند که اثبات سازگاری آن با مُدل، اندکی دشوارتر است.

قبل از همه، یادآوری می‌کنیم که، این اصل موضوع، شامل دو گزاره است.

A، A' را دو نقطه؛ a، a' را دو نیم‌خط راستی که از این دو نقطه آغاز شده‌اند و α، α' را دو نیم‌صفحه‌ای که به‌وسیلهٔ خط‌های راست a، a' محدود شده‌اند، در نظر می‌گیریم.

گزارهٔ اول می‌گوید، حرکتی وجود دارد که A را به A'، a را به a' و α را به α' تبدیل می‌کند.

گزارهٔ دوم حاکی است که، این حرکت، یکتا (منحصربه‌فرد) است.

می‌توان ادعا کرد که، این دو گزاره، در بخش سوم، بند ۱۴ (جلد اول) ثابت شده است، ولی ترجیح می‌دهیم، اثبات آن‌ها را در این‌جا، نه آن‌گونه که به بخش سوم ارتباط دارند، بلکه به‌صورت دیگری که کلی‌تر است، ثابت کنیم.

ثابت می‌کنیم، گزارهٔ اول، با مُدل (یعنی با مفهوم‌هایی که برای «نیم‌خط راست»، «نیم‌صفحه» و «حرکت» پذیرفته‌ایم) سازگار است.

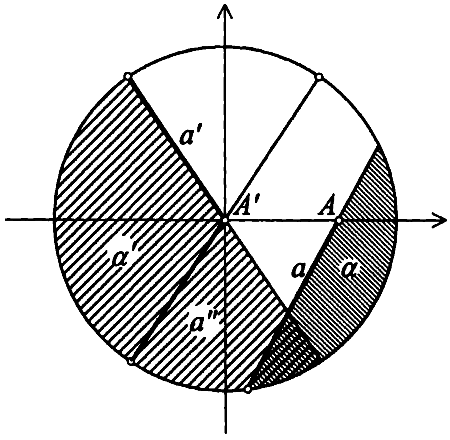
در آغاز فرض می‌کنیم A' بر مرکز دایره قرار گرفته باشد. محورهای مختصات را طوری انتخاب می‌کنیم که مبداء آن بر مرکز دایره منطبق و محور Ox از آن، از نقطهٔ A گذشته باشد (شکل ۱۴).

در بند قبل، این تبدیل را در نظر گرفته بودیم:

$$x'=\frac{x+a}{1+ax} \quad , \quad y'=\frac{y\sqrt{1-a^2}}{1+ax} \tag{۶}$$

در آن‌جا ثابت کردیم، این تبدیل، یک «حرکت» است (یعنی دایره را به خودش، و خط‌های راست را به خط‌های راست منجر می‌کند).

$x_\circ$ را طول نقطهٔ A می‌گیریم و برای عرض آن داریم: $y_\circ = ۰$. بنابراین، اگر فرض کنیم



شکل ۱۴

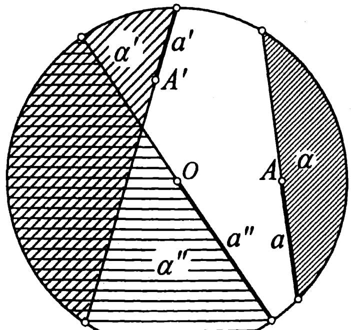
$a=-x$، آن‌وقت بنابر دستور (۶)، نقطهٔ A به نقطه‌ای با مختصات (۰ , ۰) می‌رود، یعنی بر A' منطبق می‌شود.

چون در این حرکت، خط‌های راست، به خط‌های راست منجر می‌شوند، بنابراین «نیم‌خط راست a (یعنی تکه‌ای از وتر) به وضعی مثل a'' درمی‌آید (شکل ۱۴). اکنون دوران دور مرکز، می‌تواند a'' را به a' برساند. «نیم‌صفحهٔ» α، قطعه‌ای از دایره محدود به «خط راست» (وتر) a است. اگر این «نیم‌صفحه»، بعد از حرکت بر α' منطبق باشد، آن‌وقت کار تبدیل تمام شده است؛ ولی اگر بر α' منطبق نشود، آن‌وقت با یک دوران (انعکاس نسبت به قطر a')، به نیم‌دایرهٔ α' می‌رساند.

به این ترتیب، ضمن ترکیب (۶) با دوران و اگر لازم باشد با انعکاس، A، a، α، به A'، a'، α' تبدیل، یعنی وجود حرکت لازم، ثابت می‌شود.

تا این‌جا، حالت خاصی را در نظر گرفتیم که نقطهٔ A'، بر مرکز دایره قرار گرفته باشد. اکنون فرض می‌کنیم، A' در نقطهٔ دل‌خواهی باشد. در این صورت، با توجه به آن‌چه هم‌اکنون ثابت کردیم، می‌توانیم آن را با یک حرکت به مرکز برسانیم. این حرکت را با D_1 نشان می‌دهیم. درضمن، «نیم‌خط راست» a' به «نیم‌خط راستی» مثل a'' تبدیل می‌شود که از مرکز گذشته است و «نیم‌صفحهٔ» α' به «نیم‌صفحهٔ» دیگری (نیم‌دایره‌ای) مثل α'' می‌رسد (شکل ۱۵).

همان‌طور که پیش از این دیدیم، می‌توان به‌یاری «حرکتی» مثل D_2، نقطهٔ A را به مرکز، «نیم‌خط راست» a را به a'' و «نیم‌صفحهٔ» α را به α'' رسانید. سپس، «حرکتی»، وارون D_1، A' را به جای قبلی و درضمن a'' و α'' را به وضع اولیهٔ خود a' و α'



شکل ۱۵

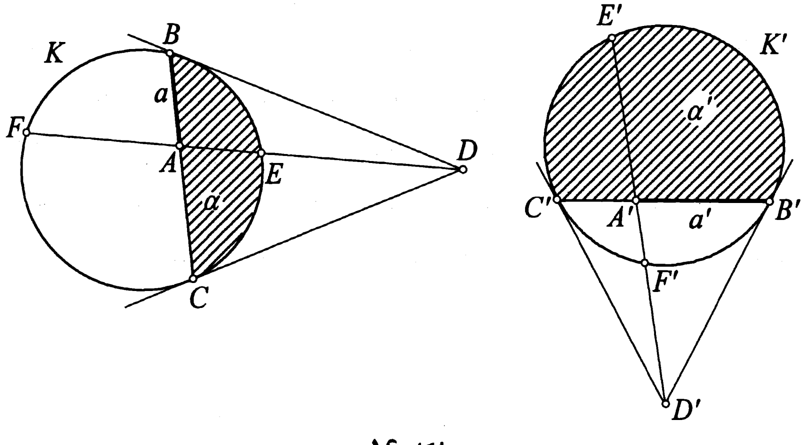
برمی‌گرداند[1].

به این ترتیب، درنتیجهٔ ترکیب «حرکت» D_2 با «حرکت» وارون D_1 می‌توانیم A ، a ، α را به A' ، a' ، α' منجر کنیم. ولی ترکیب دو «حرکت»، خود یک «حرکت» است؛ یعنی حرکتی وجود دارد که A ، a ، α را به A' ، a' ، α' تبدیل کند و این برای حالتی است که جای A و A' به‌دل‌خواه در درون دایره انتخاب شده‌اند. به این ترتیب، گزارهٔ اول اصل موضوع سوم حرکت، به‌طور کامل ثابت شد.

اکنون ثابت می‌کنیم، گزارهٔ دوم این اصل موضوع هم، با مُدل سازگار است. با توجه به مفهوم این گزاره، باید این گزاره را ثابت کنیم.

A ، A' را دو نقطهٔ درونی دایره؛ a ، a' را دو تکه وتری از این دایره که از A ، A' آغاز شده‌اند؛ α ، α' را بخش‌هایی از دایره که محدود به این وترها باشند، در نظر می‌گیریم. برای روشن‌تر بودن بحث، این نقطه‌ها، وترها و تکه‌دایره‌ها را در دو شکل جداگانه داده‌ایم (شکل ۱۶). البته، باید آن‌ها را متعلق به یک دایره دانست. گزارهٔ ما این است که تنها یک «حرکت» منحصر وجود دارد که به‌ترتیب، A را به A' ، a را به a' و α را به α' تبدیل می‌کند، یعنی این تبدیل با یک حرکت معین انجام می‌شود.

برای اثبات، تبدیلی را بررسی می‌کنیم که، نه‌تنها دایره، بلکه تمامی صفحه را در بر



شکل ۱۶

۱. اگر «حرکت» D_1 با دستورهای (۳) (بند ۴) بیان شود، «حرکت» عکس D_1 با دستورهای (۴) بیان می‌شود.

بگیرد[1]. «حرکت»، بنابر تعریف خود، خط‌های راست را به خط‌های راست تبدیل می‌کند. تبدیلی که دارای این ویژگی باشد، **تصویری** نامیده می‌شود. از این‌رو می‌توان گفت، برای ما «حرکت» یک تبدیل تصویری است که دایرهٔ مفروض را به خودش تبدیل می‌کند (این مطلب روی شکل ۱۶ ، این‌طور مشخص شده است که دایرهٔ K به دایرهٔ K' تبدیل می‌شود. تنها باید در ذهن خود، با انتقال موازی، دایرهٔ K را روی دایرهٔ K' قرار داد).

تبدیل تصویری را در بخش سوم، بند ۱۲ (جلد اول) بررسی کرده‌ایم؛ در این‌جا از قضیهٔ مهمی که در آن‌جا ثابت شده است، استفاده می‌کنیم. تبدیل تصویری به این ترتیب به‌طور کامل معین می‌شود که بدانیم، چهار نقطه‌ای که هر سه‌تای آن‌ها بر یک خط راست نیستند، به کجا رفته‌اند!

به «حرکتی» که بررسی می‌کردیم، برمی‌گردیم. این «حرکت» پاره‌ای از وتر a را به a' منجر می‌کند و بنابراین نقطهٔ B را به B' می‌رساند. چون این «حرکت»، وتر را به وتر تبدیل می‌کند، پس نقطهٔ C را هم به C' می‌رساند.

سپس، از آن‌جا که «حرکت» به‌طور کلی خط راست را به خط راست و دایرهٔ مفروض را به خودش (و در نمایش ما، دایرهٔ K را به دایرهٔ K') تبدیل می‌کند، بنابراین، مماس‌های در نقطه‌های B و C را، به مماس‌های در نقطه‌های B' و C' می‌رساند؛ درنتیجه، نقطهٔ D، محل برخورد دو مماس اول، به نقطهٔ D'، محل برخورد مماس‌های دوم می‌رود[2] (شکل ۱۶).

چون درعین‌حال، نقطهٔ A به نقطهٔ A' و خط‌های راست به خط‌های راست تبدیل می‌شود، پس خط راست AD هم روی خط راست $A'D'$ قرار می‌گیرد. خط راست AD، محیط دایره را در نقطه‌های E و F و خط راست $A'D'$ آن را در نقطه‌های E' و F' قطع می‌کند (روی شکل ۱۶ ، این نقطه‌ها روی محیط دایره‌های K و K' واقع‌اند). چون دایره به خودش تبدیل می‌شود، بنابراین نقطه‌های E و F به نقطه‌های E' و F' منجر می‌شوند. فرض کنید، نقطهٔ E روی کمانی باشد که بخش α را محدود کرده است و E' روی کمان محدودکنندهٔ α' باشد. بنابراین، چون بنابر شرط α به α' می‌رسد، پس نقطهٔ E درست روی

---

۱. می‌توان ثابت کرد، تبدیلی که دایره را به خودش و خط‌های راست را به خط‌های راست منجر می‌کند، را می‌توان با حفظ همین ویژگی به تمامی صفحه تصویری، یعنی صفحه‌ای که شامل نقطه‌های بی‌نهایت هم باشد، تعمیم داد.

۲. از جمله، اگر مماس‌های در نقطه‌های B و C با هم موازی باشند، آن‌وقت نقطهٔ D' به «نقطهٔ بی‌نهایت دور» می‌رود.

E' و نقطهٔ F درست روی F' قرار می‌گیرد.

به این ترتیب، در این «حرکت»، نقطه‌های B، C، E و F از محیط دایره، به نقطه‌های B'، C'، E' و F' منجر می‌شوند. در بین نقطه‌های B، C، E و F و همچنین در بین نقطه‌های B'، C'، E' و F'، سه نقطه‌ای وجود ندارد که روی یک خط راست باشند - بنابراین، طبق قضیه‌ای که از آن یاد کردیم، تبدیل تصویری، نقطه‌های B، C، E و F را، به‌صورت یگانه‌ای به نقطه‌های B'، C'، E' و F' تبدیل می‌کند. ولی «حرکت» هم، تبدیلی تصویری است. به این ترتیب، «حرکت» به‌صورتی یگانه، A، a و α را به A'، a' و α' تبدیل می‌کند، چیزی که اثبات آن را لازم داشتیم.

به این ترتیب، ثابت کردیم مُدل ما، با همهٔ اصل موضوع‌های هندسهٔ اقلیدسی، به‌جز اصل موضوع مربوط به خط‌های راست موازی، سازگار است، به زبان دیگر، همهٔ اصل موضوع‌های هندسهٔ *لباچوسکی*، در این مُدل صدق می‌کنند. این هندسه، که به‌عنوان نتیجه‌ای از هندسهٔ اقلیدسیِ درون دایره به‌دست می‌آید، هندسهٔ *لباچوسکی* را با واقعیت تطبیق می‌دهد که البته براساس اصطلاح‌های خاصی برای «خط راست» و «حرکت» شکل گرفته است. این وضع، اجازه می‌دهد هندسهٔ *لباچوسکی* را روی مُدل مشخص و مفروضی توجیه کنیم که، در بسیاری حالت‌ها، ساده‌تر و راحت‌تر است.

از دیدگاه تجزیه و تحلیل منطقی، به این‌جا می‌رسیم که: ۱) هندسهٔ *لباچوسکی* متناقض نیست و ۲) پوستولای مربوط به خط‌های راست موازی را نمی‌توان از بقیهٔ اصل موضوع‌ها نتیجه گرفت.

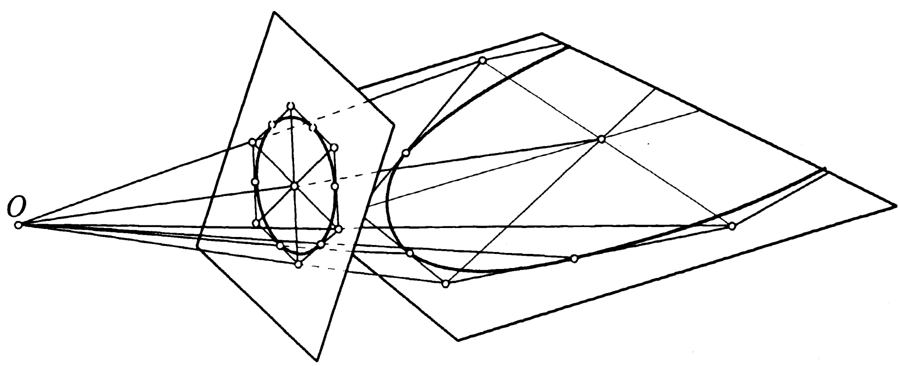
۶. جدا کردن نظریه‌های هندسی مستقل، از هندسهٔ اقلیدسی

۱. درکنار هندسهٔ *لباچوسکی*، تکامل اصولی هندسه، در مسیر دیگری ادامه یافت. برای غنی کردن همهٔ ویژگی‌های هندسی فضا، مطالعهٔ مستقل گروه‌های جداگانه‌ای از ویژگی‌ها، که از نظر پایداری و به‌هم‌پیوستگی ممتازند، درمعرض توجه قرار گرفت. این بررسی‌ها، که از نظر روش‌های خود تازگی داشت، فصل‌های تازه‌ای را در هندسه گشود: دانش مربوط به فُرم‌های فضایی، شبیه کالبدشکافی و فیزیولوژی که فصل مختلفی را دربارهٔ اندام انسان تشکیل می‌دهند. در آغاز هندسه بخش‌بندی نشده بود و به‌طور عمده، هندسهٔ متری بود و

به اندازه‌گیری شکل‌ها و ویژگی‌های فضا مربوط می‌شد. البته موقعیت‌هایی هم بررسی می‌شد که به اندازه‌گیری مربوط نمی‌شد و بستگی به خصلت کیفی وضع متقابل شکل‌ها داشت. درضمن مدت‌ها بود به این مطلب پی برده بودند که بخشی از این‌گونه ویژگی‌ها، نوعی‌پایداری را مشخص‌می‌کنند که با تغییرشکل‌هاوتغییر موضع‌های شکل حفظ می‌شوند.

برای نمونه، تصویر یک شکل را از یک صفحه بر صفحهٔ دیگر در نظر می‌گیریم (شکل ۱۷). ضمن این تصویر طول پاره‌خط‌های راست تغییر می‌کند، زاویه‌ها فرق می‌کنند و نحوهٔ ترسیم موضوع‌ها، به‌روشنی تحریف می‌شود. با وجود این، از جمله ردیف نقطه‌های واقع بریک خط راست تغییر نمی‌کند، ویژگی خط راست مماس برمنحنی‌ها حفظ می‌شود وغیره.

دربارهٔ طرح‌ریزی و تبدیل‌های تصویری در بخش سوم (جلد اول) هم صحبت کرده‌ایم و در آن‌جا بستگی روشنِ آن را با پرسپکتیو ـ نمایش شکل‌های فضایی بر صفحه نشان داده‌ایم. بررسی‌های مربوط به ویژگی‌های پرسپکتیو از زمان‌های باستان و حتی پیش از ***اقلیدس*** و در کارهای معماران باستانی آغاز شده است؛ هنرمندانی مثل ***دیورِر*** و ***لئوناردو داوینچی*** هم به مسألهٔ پرسپکتیو پرداخته‌اند؛ ***دزارگ*** ریاضی‌دان و مهندس (سدهٔ هفدهم) هم دربارهٔ ویژگی‌های پرسپکتیو بحث‌کرده است. سرانجام ***پُونسله*** در آغاز سدهٔ نوزدهم، برای نخستین بار، به‌طور منظم دربارهٔ ویژگی‌هایی از هندسه که ضمن هر تبدیل تصویری صفحه (یا فضا)، محفوظ می‌ماند، بررسی کرد و دانش تازه‌ای را به‌نام ***هندسهٔ تصویری*** بنیان گذاشت[1].



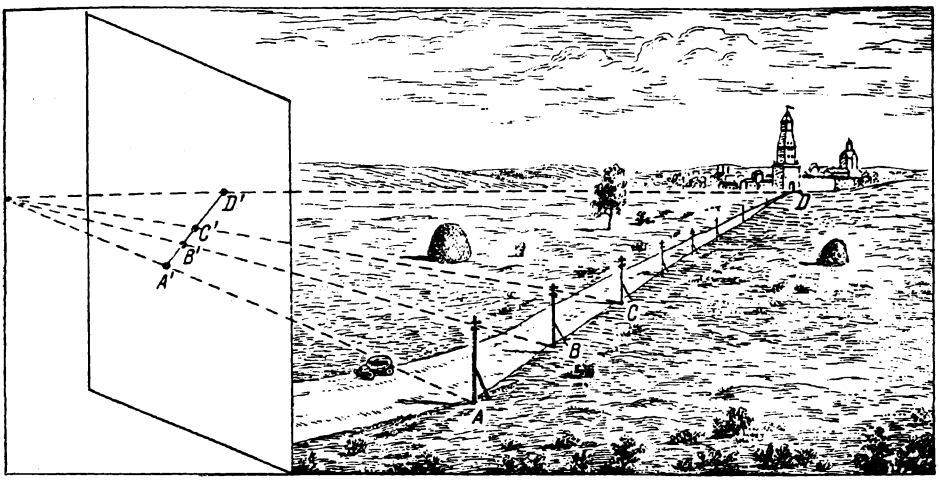
شکل ۱۷

۱. ***پُونسله***، مهندس نظامی فرانسوی، بررسی‌های هندسی خود را در زمان اسارت در روسیه بعد از ۱۸۱۲ انجام داد. نوشتهٔ او با عنوان «رساله‌ای دربارهٔ ویژگی‌های تصویری شکل‌ها» در سال ۱۸۲۲ چاپ شد.

ممکن است گمان رود، ویژگی‌هایی که ضمن هر تبدیل تصویری به قوت خود باقی می‌مانند، چندان زیاد نیست و درضمن ویژگی‌هایی مقدماتی‌اند؛ ولی درواقع این‌طور نیست. برای نمونه به این مطلب توجه کنیم: در هندسه قضیه‌ای وجود دارد که بنا بر آن، اگر ضلع‌های روبه‌رو در یک شش‌ضلعی محاط در دایره را امتداد دهیم تا به هم برسند، سه نقطه‌ای که از برخورد دوبه‌دوی این ضلع‌های روبه‌رو به‌دست می‌آید، روی یک خط راست‌اند؛ این قضیه برای شش‌ضلعی محاط در بیضی، یا سهمی یا هذلولی هم درست است. این قضیه، دربارهٔ ویژگی‌های تصویری صحبت می‌کند، چرا که این منحنی‌ها از راه تصویر دایره به‌دست می‌آیند. همچنین، این قضیه (که کمتر روشن است) که قطرهای یک شش‌ضلعی محیط بر دایره، در یک نقطه به‌هم می‌رسند، شباهت خاصی به قضیهٔ قبلی دارد؛ این دو قضیه در هندسهٔ تصویری، بستگی عمیقی با هم دارند. این مطلب هم کمتر شناخته شده است که در تصویر چهار نقطهٔ A، B، C و D که بر یک خط راست واقع‌اند، با وجود تغییر فاصله‌ها (شکل ۱۸)، نسبت دوگانهٔ $\frac{AC}{CB}:\frac{AD}{DB}$ بی‌تغییر می‌ماند:

$$\frac{AC}{CB}:\frac{AD}{DB}=\frac{A'C'}{C'B'}:\frac{A'D'}{D'B'}$$

و این منجر به رعایت بسیاری از بستگی‌ها در پرسپکتیو می‌شد. از جمله در عکس‌برداری از جاده‌های دور (شکل ۱۸)، وقتی که فاصله تیرهای تلگراف A'، B'، C' را از هم بدانیم،



شکل ۱۸

می‌توانیم، با استفاده از فاصله‌بین آن‌ها در عکس، فاصلهٔ از آن‌ها تا نقطهٔ D' را به‌دست آورد.

دربارهٔ هندسهٔ تصویری و استفاده از نتیجه‌های آن در کارهای عکس‌برداری هوایی در بخش سوم (جلد اول) صحبت کرده‌ایم. یادآوری می‌کنیم، قانون‌های هندسهٔ تصویری، در معماری، در ساختمان منظره چشم‌انداز، در دِکوراسیون وغیره هم کاربرد دارد.

پیدایی و پیشرفت هندسهٔ تصویری، در تکامل خود هندسه هم، نقشی اساسی به‌عهده داشت.

۲. یکی دیگر از هندسه‌های مستقل، ***هندسهٔ آفین*** است. در هندسهٔ آفین، ویژگی‌هایی از شکل بررسی می‌شود که به‌ازای هر تبدیلی تغییر نمی‌کنند و در آن‌ها، مختصات دکارتیِ نخستین (x, y, z) و مختصات جدید (x', y', z')، معرف هر نقطه، با این معادله‌های خطی به هم بستگی دارند:

$$
\begin{aligned}
x' &= a_۱x + b_۱y + c_۱z + d_۱ \ , \\
y' &= a_۲x + b_۲y + c_۲z + d_۲ \ , \\
z' &= a_۳x + b_۳y + c_۳z + d_۳
\end{aligned}
$$

با این فرض که دترمینان زیر مخالف صفر باشد:

$$
\begin{vmatrix}
a_۱ & b_۱ & c_۱ \\
a_۲ & b_۲ & c_۲ \\
a_۳ & b_۳ & c_۳
\end{vmatrix}
$$

هر تبدیل آفین برابر است با حرکتی در صفحه (که ممکن است یک انعکاس باشد) و سپس به‌هم فشردن یا باز شدن فضا در سه جهت دوبه‌دو عمود بر هم.

به‌ازای هریک از این تبدیل‌ها، بسیاری از ویژگی‌های شکل باقی می‌ماند. خط‌های راست به خط‌های راست تبدیل می‌شوند (و به‌طور کلی، همهٔ ویژگی‌های تصویری به قوت خود باقی می‌مانند)؛ به‌جز این خط راست موازی، بعد از تبدیل به‌صورت خط‌های راست موازی درمی‌آیند؛ نسبت حجم‌ها، نسبت مساحت‌های شکل‌هایی که روی یک صفحه یا صفحه‌های موازی‌اند، نسبت طول پاره‌خط‌های راستی که روی یک خط راست یا خط‌های راست موازی‌اند، محفوظ می‌ماند وغیره. بسیاری از قضیه‌ها، که درواقع مربوط به هندسهٔ آفین هستند، برای همه روشن است. از این جمله‌اند گزاره‌های: میانه‌های مثلث در یک نقطه به هم می‌رسند، قطرهای متوازی‌الاضلاع یکدیگر را نصف می‌کنند، نقطه‌های وسط

وترهای موازی در بیضی روی یک خط راست قرار دارند وغیره.

نظریهٔ خم‌ها (و رویه‌ها) ی درجه دوم، رابطهٔ بسیار نزدیکی با هندسهٔ آفین دارد. خود تقسیم‌بندی این منحنی‌ها، به بیضی‌ها، سهمی‌ها و هذلولی‌ها، درواقع بر اساس ویژگی‌های آفین این شکل‌ها است: در تبدیل آفین، بیضی به بیضی تبدیل می‌شود، نه به سهمی یا هذلولی؛ همچنین سهمی می‌تواند به هر سهمی دیگری تبدیل شود نه به بیضی وغیره.

بررسی تفصیلی هندسهٔ آفین و ویژگی‌های کلی آفین شکل‌ها، به‌خصوص از این جهت اهمیت دارد که، در مقایسه با تبدیل‌های پیچیده‌تر، وقتی با مقدارهای بسیار کوچک سر و کار داشته باشیم، درواقع می‌توانند تبدیل خطی به‌حساب آیند، یعنی تبدیل‌های آفین و استفاده از روش‌های محاسبهٔ دیفرانسیلی به‌ویژه ضمن بررسی حوزه‌های بسیار کوچک فضا، اهمیت زیادی پیدا می‌کنند.

۳. در سال ۱۸۷۲ ، **کلاین** در درس‌های خود در دانشگاه *ارلانگن*، که به *برنامهٔ ارلانگن* معروف شده است، نتیجه‌گیری‌های مربوط به هندسهٔ تصویری، هندسهٔ آفین و دیگر «هندسه‌ها» را جمع‌بندی و به‌روشنی نظام کلی ساختمان آن‌ها را تنظیم کرد: می‌توان هر گروه از تبدیل‌های فضا را که نسبت به هم یک‌ارزشی‌اند، بررسی کرد و آن ویژگی‌های شکل‌ها را درمعرض مطالعه قرار داد که ضمن تبدیل‌های این گروه، حفظ می‌شوند[۱].

از این دیدگاه، ویژگی‌های فضا، بنابر عمق و پایداری خود طبقه‌بندی می‌شوند. هندسهٔ اقلیدسی عادی از راه جدا شدن از همهٔ ویژگی‌های جسم واقعی، به‌جز ویژگی‌های هندسی ساخته می‌شود؛ این‌جا در شاخه‌های اختصاصی هندسه، باز هم انتزاعی در درون هندسه انجام می‌گیرد، انتزاع از یک‌رشته ویژگی‌های هندسی، به‌جز ویژگی‌های معینی که در شاخهٔ مربوط به آن‌ها علاقه‌مندیم.

بنا بر نظامی که کلاین پیشنهاد می‌کند، می‌توان تعداد زیادی هندسه ساخت. از جمله می‌توان تبدیل‌هایی را در نظر گرفت که، ضمن انجام آن‌ها، زاویهٔ بین خط‌ها ثابت بماند (تبدیل‌های همدیس ـ کُنفُرم ـ فضا) و دربارهٔ ویژگی‌هایی از شکل که در این تبدیل‌ها، یعنی

---

۱. واژهٔ «گروه» در این‌جا، به‌طور ساده و به‌مفهوم «گِرد آمدن» به‌کار نرفته است. وقتی دربارهٔ گروه تبدیل‌ها صحبت می‌شود (بخش دهم را ببینید)، چنان مجموعه‌ای از تبدیل‌ها در نظر است که شامل تبدیل همانی باشد (تبدیلی که همهٔ نقطه‌ها را بر جای خود باقی می‌گذارد) و که همراه هر تبدیلی شامل تبدیل وارون آن باشد (تبدیلی که همهٔ نقطه‌ها را به جای قبلی خود برگرداند) و که، به همراه هر دو تبدیل این مجموعه، تبدیلی هم‌ارز آن‌ها، وقتی یکی بعد از دیگری انجام شود، وجود داشته باشد.

تبدیل‌های همدیس، محفوظ می‌ماند، بررسی کرد (هندسهٔ همدیس‌پذیر). می‌توان به بررسی تبدیل‌هایی پرداخت که در بخشی از فضا، و نه همهٔ آن، انجام شود. برای نمونه، با بررسی نقطه‌ها و وترهای درون دایره، برای همهٔ تبدیل‌هایی که دایره را به خودش و وترها را به وترها منجر کند و به بررسی ویژگی‌هایی پرداخت که ضمن این تبدیل‌ها ثابت می‌مانند که در این صورت، به هندسه‌ای می‌رسیم که در بند ۴ و بند ۵ دربارهٔ آن صحبت کردیم و همان هندسهٔ ***لباچوسکی*** است.

۴. تکامل نظریه و تقسیم‌بندی آن، ولو از جهت اصولی (در این‌جا صحبت از محتوای واقعی نیست)، به همین‌جا پایان نمی‌پذیرد.

از جمله، اگر تنها به ویژگی‌های آفین شکل‌ها توجه داشته باشیم، می‌توان از میان همهٔ ویژگی‌های ممکنِ فضا و شکل‌های هندسی آن، تنها ویژگی‌هایی را جدا کرد که به آن‌ها علاقه‌مندیم و به‌دیگر ویژگی‌های دیگر فضا و شکل بی‌توجه بود. در این «فضا»، شکل‌ها به‌طور کلی، هیچ ویژگی دیگری، به‌جز ویژگی‌های آفین ندارند. درضمن، طبیعی است تلاش کنیم چنین هندسه‌ای را اصل موضوعی کنیم، یعنی فرض کنیم صحبت بر سر موضوع‌های انتزاعی خاصی است: از «نقطه‌ها»، «خط‌های راست» و «صفحه‌هایی صحبت کنیم که ویژگی‌های آن‌ها (و روشن است، این ویژگی‌ها، کمتر از حالتی است که در هندسهٔ اقلیدسی کار می‌کنیم) با اصل موضوع‌هایی سازگارند و درضمن، نتیجه‌هایی که از این اصل موضوع‌ها به‌دست می‌آید، متناظر با ویژگی‌های آفین شکل در فضای معمولی باشند.

درواقع هم می‌توان به چنین انتزاعی دست یافت؛ آن‌وقت چنین مجموعه‌ای از «نقطه‌ها»، «خط‌های راست»، «صفحه‌ها» و نتیجه‌های ناشی از آن‌ها را ***فضای آفین*** نامید.

به همین ترتیب می‌توان نظامی انتزاعی از موضوع‌ها را در نظر گرفت که تنها دارای ویژگی‌هایی باشند که متناظر با ویژگی‌های تصویری شکل‌ها در فضای اقلیدسی هستند (که در این‌جا هم، برای اصل موضوعی کردن، باید اصل موضوع‌هایی را، غیر از اصل موضوع‌های هندسهٔ معمولی در نظر گرفت).

۵. اگر در طبیعت و ماهیت نگاره‌های هندسی دقیق‌تر شویم، متوجه می‌شویم که در بسیاری از مسأله‌ها، با ویژگی‌هایی عمیق‌تر از حالت مربوط به تبدیل‌های تصویری سروکار داریم؛ این ویژگی‌های شکل، چنان پایدارند که با هر تحریفی که در شکل پیش آید، به‌شرطی که این تغییر منجر به پارگی شکل یا روی هم گذاشتن قطعه‌هایی از آن نشود، تغییر

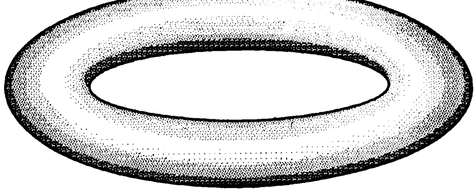
نمی‌کنند (تصور دربارهٔ این‌گونه تغییرشکل پیوسته را، به‌جز شرح عینی که در این‌جا آوردیم، می‌توان براساس تعریفی هم که در آنالیز از تابع‌های پیوسته داریم، روشن کرد و گفت: صحبت بر سر هرگونه تبدیلی از همهٔ نقطه‌های شکل و رساندن آن‌ها به موقعیت تازه‌ای است، به‌نحوی که مختصات دکارتی نقطه‌ها در موقعیت تازه، با تابع‌های پیوسته‌ای با مختصات نخستینِ آن‌ها مربوط باشند؛ همچنین مختصات قدیم را هم، به‌نوبهٔ خود، بتوان به‌کمک تابع‌های پیوسته‌ای برحسب مختصات تازه نوشت).

ویژگی‌هایی از شکل را که ضمن هر تبدیلی از این‌گونه حفظ می‌شود، ***ویژگی‌های توپولوژیک*** و دانشی را که به این‌گونه تبدیل می‌پردازد، ***توپولوژی*** می‌نامند (بخش هجدهم را در همین جلد ببینید).

ویژگی‌های توپولوژیک برای شکل‌های ساده، به‌طور مستقیم و به‌صورت عینی قابل توضیح است. از جمله به‌تقریب روشن است، اگر در روی صفحه، محیط دایره را به هر صورت دل‌خواهی کج و تحریف کنیم و به هر صورتی آن را بپیچانیم، به یک خم می‌رسیم که صفحه را به دو بخش تقسیم می‌کند: بخش درونی و بخش بیرونی؛ بنابراین، این که محیط دایره صفحه را تقسیم می‌کند، یک ویژگی توپولوژیک است. این هم به‌احتمالی روشن و عینی است که سطح چنبره (*torus* ـ شکل ۱۹) را به هیچ صورتی با تبدیل‌های پیوسته، نمی‌توان به سطح کره منجر کرد، بنابراین ویژگی سطحی که بتوان آن را با تبدیل پیوسته تغییر داد، مثل سطح چنبره، ویژگی توپولوژیک آن است که با ویژگی بسیاری از سطح‌های دیگر متفاوت است.

استدلالِ براساس پیوستگی، استدلالی شهودی است ولی اغلب چنان خوب ماهیت کار را نشان می‌دهد که بسیار جذاب است اگر بتوان این اثبات را کامل کرد و درضمن، از آن برای مسأله‌های بغرنج‌تر هم استفاده کرد.

از این جمله است، استدلالی که برای قضیهٔ اصلی جبر می‌توان به‌کار برد: هر معادلهٔ



شکل ۱۹

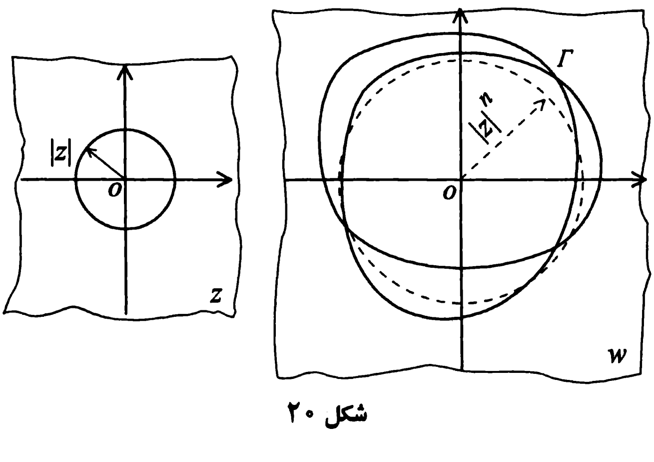
به‌صورت:

$$z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_{n-1} z + a_n = 0 \qquad (۷)$$

دست‌کم یک ریشهٔ حقیقی یا یک ریشهٔ مختلط دارد.

z را نقطه‌ای از یک صفحهٔ مختلط و $w=f(z)$ را نقطهٔ متناظر آن از صفحهٔ دیگر مختلط w می‌گیریم؛ درضمن $f(z)$ را به‌معنای بخش سمت چپ معادلهٔ (۷) به‌حساب می‌آوریم. اگر z را از لحاظ قدرمطلق خیلی بزرگ فرض کنیم، آن‌وقت تابع $f(z)$، خیلی کم با مقدار z^n فرق خواهد داشت؛ z^n هم، تابعی بسیار ساده است. به‌ویژه، می‌توان تحقیق کرد، اگر نقطهٔ z به‌صورتی پیوسته، محیط دایره‌ای به‌مرکز مبداء مختصات را در صفحهٔ مختلط بپیماید، آن‌وقت نقطهٔ $w_1=z^n$، درست n مرتبه دایره‌ای‌بامحیط $|z|^n$ را درصفحهٔ w دورخواهد زد[1]. نقطهٔ $w=f(z)$ هم، n طوقه را که مسیری مثل Γ تشکیل می‌دهد و به منحنی $w_1=z^n$ نزدیک است دور می‌زند (شکل ۲۰).

اکنون اگر محیط دایره‌ای را که نقطهٔ z می‌پیماید، مرتب به سمت یک نقطه کوچک کنیم، آن‌وقت n بار حرکت در مسیر Γ هم، که نقطهٔ $w=f(z)$ می‌پیماید، به‌طور پیوسته تغییرشکل



شکل ۲۰

۱. درواقع اگر $z=\rho(\cos\varphi + i\sin\varphi)$، آن‌وقت [بخش چهارم (جلد اول)، بند ۳ را ببینید]

$$z^n=\rho^n(\cos\varphi n + i\sin\varphi n)$$

و بنابراین، با معلوم بودن آوند φ از عدد z در فاصلهٔ ۰ تا 2π، آوند عدد z^n از ۰ تا $2\pi n$ تغییر می‌کند، یعنی نقطهٔ z^n، n دور کامل می‌زند.

می‌دهد و درضمن به یک نقطه نزدیک می‌شود. ولی روشن است، این مسیر Γ نمی‌تواند به یک نقطه نزدیک شود، مگر این که دست‌کم یک بار مبداء مختصات O را ــ که این مسیر در آغاز آن را در بر گرفته بود ــ قطع کند؛ بنابراین، دست‌کم یک بار به O می‌رسد و به‌ازای این مقدار z خواهیم داشت: $w = f(z) = 0$. همین مقدار z، ریشه‌ای از معادلهٔ (۷) است. درضمن روشن است که تعداد ریشه‌ها ـ ریشه به‌مفهوم معلومِ آن ــ باید برابر n باشد، زیرا هریک از n طوقهٔ مسیر Γ، ضمن فشرده شدن خود، نقطهٔ O را قطع می‌کند.

البته این استدلال وقتی دقیق می‌شود که براساس توپولوژی، ویژگی‌های مسیرها و تغییرشکل آن‌ها ــ که در این‌جا از آن‌ها استفاده کردیم ــ بنیان گذاشته شود.

مثال‌های زیادی از شاخه‌های مختلف ریاضیات می‌توان آورد ـ که گاه به‌کلی دور از هندسه‌اند، ولی به‌یاری توپولوژی به نتیجه می‌رسند.

با بررسی ویژگی‌های توپولوژیک، دوباره در برابر ما این امکان به‌وجود می‌آید که مجموعه‌ای از اشیاء را که دارای ویژگی‌هایی از این‌گونه هستند، جدا کنیم (بند ۷ بخش بیستم را ببینید). این مجموعهٔ اشیاء، نام ***فضای توپولوژیک انتزاعی*** را بر خود دارد.

این دیدگاه، به‌مراتب گسترده‌تر از آن است که تنها ویژگی‌های توپولوژیک را روی شکل‌های هندسی بررسی کنیم. فضاهای توپولوژیک، بسیار متنوع‌اند؛ از جمله، همهٔ نقطه‌های سطح چنبره با اضافه کردن خصوصیات خاص مربوط به نزدیکی نقطه‌های آن، یک فضای توپولوژیک را تشکیل می‌دهند، همهٔ نقطه‌های صفحه، یا همهٔ فضای اقلیدسی، یا همهٔ سطح‌های چندبرگی ریمانی، که دربارهٔ آن‌ها در بند ۵ بخش نهم (جلد دوم) صحبت کرده‌ایم و متعلق به نظریهٔ تابع‌های متغیر مختلط‌اند، هرکدام یک فضای توپولوژیک و درضمن متفاوت با هم تشکیل می‌دهند. ولی جالب‌تر از همه، آن فضاهایی است که موضوع‌های آن هیچ شباهتی با تصور ما دربارهٔ نقطه‌های هندسی ندارد، ولی دربارهٔ آن‌ها می‌توان از مفهوم‌های نزدیکی و فشردگی استفاده کرد. از جمله، برای همهٔ موقعیت‌های یک سازوکار مکانیکی، می‌توان به‌روشنی معنای موقعیت‌های «نزدیک به هم» داد و روشن کرد، نسبت به دیگرانی که موقعیتی نزدیک به موقعیت مفروض دارند، کدام موقعیت‌ها «بی‌نهایت» چسبیده به آن است.

می‌بینیم، مفهوم فضای توپولوژیک، مفهومی بسیار کلی است؛ و ما یک بار دیگر در بند ۸ به آن برمی‌گردیم.

در این بند می‌خواستیم، بیش از آن‌چه دربارهٔ هندسه‌های مختلف تصوری به خواننده

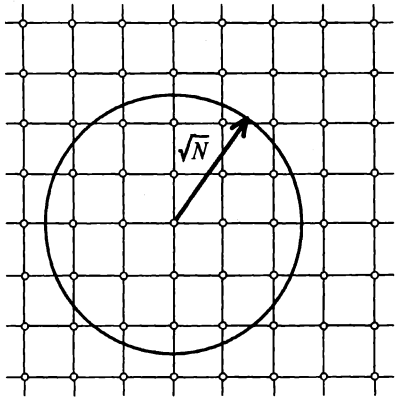
بدهیم، روشن کنیم که مشخص‌ترین مسأله‌ها، منجر به تقسیم و بررسی گروه ویژگی‌های هندسی جداگانه‌ای می‌شود و این بررسی‌ها، موجب پیدایی چنان اشیاء هندسی انتزاعی می‌شوند که تنها شامل همین ویژگی‌ها هستند، یعنی جدا کردن این ویژگی‌ها به‌صورت خالص خود به تصور دربارهٔ فضای انتزاعی خاصی می‌رساند.

در بند بعد، دربارهٔ مسیرهای دیگری که منجر به ساختن فضاهای انتزاعیِ مختلف می‌شود، صحبت خواهیم کرد.

۷. فضای چندبُعدی

۱. ساختن هندسه‌های فضای چندبُعدی، مسیر مهمی را در تکامل اندیشه‌های هندسی تازه تشکیل می‌دهد. یکی از علت‌های به‌وجود آمدن هندسهٔ چندبُعدی، تمایل به استفاده از ملاحظه‌های هندسی، ضمن حل مسأله‌های جبر و آنالیز بود. نزدیکی هندسه به حل مسأله‌های تحلیلی، براساس روش مختصاتی است. مثال ساده‌ای می‌آوریم.

می‌خواهیم تعداد جواب‌های درست را در نابرابری $x^2+y^2<N$ پیدا کنیم. اگر x و y را مختصات دکارتی روی صفحه بگیریم، مسأله به این‌جا منجر می‌شود که ببینیم: چند نقطه با مختصات درست در درون دایرهٔ به‌شعاع $\sqrt{N}$ وجود دارد. نقطه‌های با مختصات درست، رأس‌های مربع‌های با ضلع به طول واحدند که صفحه را پوشانده‌اند (شکل ۲۱). تعداد این



شکل ۲۱

نقطه‌ها در درون دایره به‌تقریب برابر است با تعداد مربع‌های واقع در درون دایره، یعنی به‌تقریب برابر است با مساحت دایرهٔ با شعاع به‌طول $\sqrt{N}$. بنابراین، تعداد جواب‌هایی که برای نابرابری باید به‌دست آوریم، به‌تقریب برابر است با πN. درضمن به‌سادگی ثابت می‌شود، وقتی $N \to \infty$، آن‌وقت، اشتباه نسبی در این‌جا به سمت صفر میل می‌کند. بررسی دقیق‌تر این خطا، یکی از مسأله‌های دشوار نظریهٔ عددها را تشکیل می‌دهد که در همین گذشتهٔ نه‌چندان دور، موضوع بررسی‌های عمیقی را تشکیل می‌داد.

در بررسی این مثال روشن شد، اگر آن را به مسأله‌ای با زبان هندسی تبدیل کنیم، بلافاصله نتیجه‌ای به‌دست می‌آید که از دیدگاه «جبر خالص» خیلی روشن نیست و به‌سادگی حاصل نمی‌شود. نابرابری‌های شامل سه مجهول هم، به‌صورت یک مسألهٔ تحلیلی، با همین روش حل می‌شوند. ولی اگر با تعدادی بیش از سه مجهول سر و کار داشته باشیم، این روش را نمی‌توان به‌کار برد، زیرا فضای ما سه‌بُعدی است، یعنی موقعیت هر نقطه در آن، با سه مختص معین می‌شود. برای این که بتوانیم از چنین روش هندسی، در حالت‌های مشابه استفاده کنیم، «فضای nبُعدی» انتزاعی را ساخته‌اند که هر نقطهٔ آن، با n مختص x_1، x_2، ... ، x_n معین می‌شود. درضمن مفهوم‌های اصلی هندسه به‌نحوی تعمیم داده می‌شوند که بتوان از ملاحظه‌های هندسی، برای حل مسأله‌های شامل n متغیر استفاده کرد؛ این وضع، موضوع پیدا کردن جواب و رسیدن به نتیجه را، بسیار ساده می‌کند. امکان چنین تعمیمی براساس یک قانون‌مندی جبری قرار دارد که به‌یاری آن، بسیاری از مسأله‌ها، با هر تعداد متغیر، شبیه هم حل می‌شوند. این وضع به ما امکان می‌دهد، از آن ملاحظه‌های هندسی که برای سه متغیر به‌کار بردیم، در هر مسأله‌ای با هر چند متغیر استفاده کنیم.

۲. سرچشمهٔ مفهوم‌های مربوط به فضای چهاربُعدی را، در کارهای لاگرانژ هم می‌توان پیدا کرد. لاگرانژ در زمینهٔ مکانیک، زمان را به‌طور صوری به‌عنوان «مختص چهارم» در کنار مختصات فضایی به‌کار برده است. ولی طرح منظم هندسهٔ چندبُعدی در سال ۱۸۴۴، به‌وسیلهٔ ***گراسمان*** ریاضی‌دان آلمانی و ***کالی*** ریاضی‌دان انگلیسی، بدون این که از کار یکدیگر آگاه باشند، داده شد. آن‌ها در این مسیر از همان راه هندسهٔ تحلیلی عادی رفتند. این شبیه‌سازی با هندسهٔ تحلیلی عادی را می‌توان به این ترتیب توضیح داد.

نقطه در فضای nبُعدی با n مختص x_1، x_2، ... ، x_n معین می‌شود. یک شکل، در فضای nبُعدی عبارت است از مکان هندسی، یا مجموعهٔ نقطه‌هایی که با این یا آن شرط سازگار باشد. به‌عنوان نمونه، «مکعب nبُعدی»، مکان هندسی نقطه‌هایی از فضای nبُعدی است که

مختصات آن‌ها، در نابرابری

$$a \leq x_i \leq b \ (i = ۱ , ۲ , \dots , n)$$

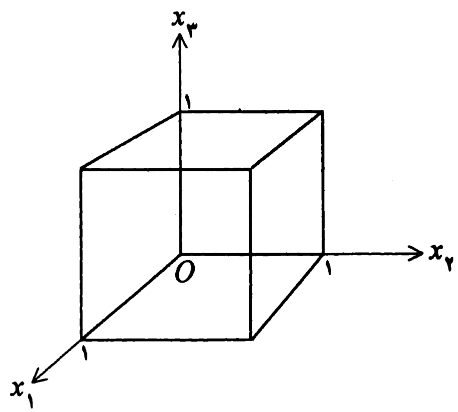
صدق کنند. شباهت این مکعب با مکعب معمولی آشکار است: در حالت $n = ۳$، یعنی در فضای سه‌بُعدی، این نابرابری معرف مکعبی است که یال‌های آن موازی محورهای مختصات و طول هر یال آن برابر $b - a$ است (در شکل ۲۲، حالت $a = ۰$، $b = ۱$ نشان داده شده است).

فاصلهٔ بین دو نقطه را می‌توان به‌عنوان ریشهٔ دوم از مجموع مجذورهای تفاضل مختصات معین کرد:

$$d = \sqrt{(x'_۱ - x''_۱)^۲ + (x'_۲ - x''_۲)^۲ + \cdots + (x'_n - x''_n)^۲}$$

و این درواقع، تعمیم دستوری است که برای تعیین فاصلهٔ بین دو نقطه در صفحه یا در فضای سه‌بعدی، یعنی به‌ازای $n = ۲$ یا $n = ۳$ می‌شناسیم.

اکنون می‌توان مسألهٔ برابری شکل‌ها را در فضای nبُعدی حل کرد. دو شکل را وقتی برابر به‌حساب می‌آوریم که بتوان بین نقطه‌های آن‌ها، چنان تناظری برقرار کرد که، به‌ازای آن، فاصلهٔ بین هر دو نقطه از یک شکل با فاصلهٔ بین دو نقطهٔ متناظر آن در شکل دوم، برابر باشد. تبدیلی را که در آن، فاصله‌ها محفوظ می‌ماند، می‌توان حرکت نامید[1]. در این صورت،



شکل ۲۲

۱. تعمیم تنها مربوط به این نیست که با n اندازه‌گیری سر و کار داریم، بلکه در این هم هست که در بین حرکت‌ها، انعکاس‌های صفحه را هم قرار داده‌ایم، زیرا یک انعکاس هم فاصلهٔ بین نقطه‌ها را تغییر نمی‌دهد.

شبیه هندسهٔ اقلیدسی معمولی می‌توان گفت: موضوع هندسهٔ n بُعدی، ویژگی‌هایی از شکل است که ضمن حرکت (به مفهوم عام آن) تغییر نمی‌کنند. این تعریف، برای موضوع هندسهٔ n بُعدی، در سال‌های ۷۰ سدهٔ نوزدهم داده شد و مبنای کار با آن قرار گرفت. از این زمان، هندسهٔ n بُعدی، موضوع بسیاری از بررسی‌ها در همهٔ جهت‌ها قرار گرفت، درست شبیه جهت‌هایی که در هندسهٔ اقلیدسی وجود دارد (هندسهٔ مقدماتی، نظریهٔ کلی منحنی‌ها وغیره).

مفهوم فاصلهٔ بین نقطه‌ها، به ما امکان می‌دهد به مفهوم‌های دیگری از فضای n بُعدی برسیم، مفهوم‌هایی مثل پاره‌خط راست، کره، طول، زاویه، حجم وغیره. از جمله، کرهٔ n بُعدی را می‌توان به‌عنوان مجموعهٔ نقطه‌هایی تعریف کرد که فاصلهٔ آن‌ها از نقطهٔ مفروض، بیشتر از مقدار مفروض R نباشد. بنابراین، کره را به‌صورت تحلیلی می‌توان این‌طور نشان داد:

$$(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2 \leq R^2$$

که در آن a_1، ...، a_n مُعرف مختصات مرکز آن است. سطح کره هم، با این معادله داده می‌شود:

$$(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2 = R^2$$

پاره‌خط راست AB را می‌توان به‌عنوان مجموعهٔ نقطه‌های X، تعریف کرد، به‌نحوی که مجموع فاصله‌های X از A و B، برابر با فاصلهٔ A از B باشد (طول پاره‌خط راست AB، یعنی فاصلهٔ بین دو نقطهٔ انتهایی آن A و B).

۳. دربارهٔ صفحه‌های با بُعدهای مختلف، اندکی بیشتر صحبت می‌کنیم.

در فضای سه‌بُعدی، با «صفحه‌های» یک‌بُعدی، یعنی خط راست و صفحه‌های معمولی (دوبُعدی) سر و کار داریم. در فضای n بُعدی، برای $n > 3$، صفحه‌های چندبُعدی با تعداد بُعدهای از ۳ تا $n - 1$ هم وجود دارند.

می‌دانیم در فضای سه‌بُعدی، صفحه با یک معادلهٔ خطی و خط راست با دو معادلهٔ خطی بیان می‌شود.

با تعمیم مستقیم، به این تعریف می‌رسیم: **صفحهٔ k بُعدی** در فضای n بُعدی به مکان هندسی نقطه‌هایی گفته می‌شود که مختصات آن‌ها، در دستگاهی شامل $n - k$ معادلهٔ خطی

صدق کند:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n-k,1}x_1 + a_{n-k,2}x_2 + \cdots + a_{n-k,n}x_n + b_n &= 0 \end{aligned} \tag{8}$$

درضمن، همهٔ معادله‌ها با هم سازگار و مستقل از یکدیگرند (یعنی هیچ‌کدام از آن‌ها را نمی‌توان از بقیه نتیجه گرفت). هریک از این معادله‌ها مُعرف صفحه‌ای $(n-1)$بُعدی و همهٔ آن‌ها با هم، مُعرف نقطه‌های مشترک این $(n-k)$ صفحه است.

این که همهٔ معادله‌های (۸) با هم سازگارند، به این معناست که، در حالت کلی، نقطه‌هایی وجود دارد که در همهٔ آن‌ها صدق می‌کند، یعنی $(n-k)$ صفحهٔ $(n-1)$ بُعدی، متقاطع‌اند. این که، هیچ‌کدام از این معادله‌ها را نمی‌توان از بقیه نتیجه گرفت، به این معناست که هیچ‌کدام از آن‌ها را نمی‌توان کنار گذاشت. در غیر این صورت، دستگاه منجر به تعداد کمتری معادله می‌شود که مُعرف صفحه‌ای با بُعد بیشتر است. بنابراین، اگر به زبان هندسی صحبت کنیم، کار به این‌جا می‌رسد که، صفحهٔ kبُعدی، به‌عنوان برخورد $(n-k)$ صفحهٔ $(n-1)$بُعدی تعریف می‌شود که با معادله‌های مستقل از یکدیگر نشان داده شده‌اند. در حالت خاص، اگر $k = 1$، آن‌وقت $(n-1)$ معادله داریم که یک «صفحهٔ یک‌بُعدی»، یعنی خط راست را معرفی می‌کند. به این ترتیب، تعریف صفحهٔ k بُعدی، تعمیم طبیعی نتیجه‌های آشنایی است که از هندسهٔ تحلیلی می‌شناسیم. این تعمیم، این سود را هم دارد که برای نتیجه‌گیری‌های مربوط به دستگاه معادله‌های خطی، تفسیری هندسی پیدا می‌کنیم که این نتیجه‌گیری‌ها را با روشنی بیشتر در برابر ما قرار می‌دهد. برای آشنایی با این نتیجه‌گیری‌ها و تعبیر هندسی مربوط به جبر خطی، می‌توانید به بخش شانزدهم مراجعه کنید.

یکی از ویژگی‌های مهم صفحهٔ k بُعدی این است که می‌توان خودِ آن را، به‌عنوان یک فضای k بُعدی به‌حساب آورد. از جمله، صفحهٔ سه‌بُعدی، همان فضای سه‌بُعدی معمولی است. این وضع به ما امکان می‌دهد به فضایی با بُعد بیشتر منتقل شویم و از نتیجه‌های بسیاری که در فضاهای با بُعد کمتر به‌دست آورده‌ایم، استفاده کنیم، شبیه استدلالی که برای عبور از n به $n+1$ انجام می‌دهیم.

اگر معادله‌های (۸) سازگار با یکدیگر و مستقل از هم باشند، آن‌وقت همان‌طور که در جبر ثابت می‌شود، می‌توان از n متغیر x_i، k متغیر را طوری انتخاب کرد که بقیهٔ $(n-k)$ متغیر برحسب آن‌ها بیان شوند[1]؛ مثل

$$x_{k+1} = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1k}x_k + d_1$$
$$x_{k+2} = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2k}x_k + d_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$x_n = c_{n-k,1}x_1 + c_{n-k,2}x_2 + \cdots + c_{n-k\,k}x_k + d_k$$

در این‌جا x_1، x_2، ... ، x_k را می‌توان هر عدد دل‌خواهی گرفت و بقیهٔ x_iها را برحسب آن‌ها به‌دست آورد. و این به‌معنای آن است که جای نقطه در صفحهٔ k بُعدی با k مختص معین می‌شود که می‌توانند هر مقدار دل‌خواهی باشند؛ و درست، به همین مفهوم است که صفحه k بُعد دارد.

از تعریف صفحهٔ با بُعدهای مختلف، می‌توان این قضیه‌های اصلی را با روش خالص جبری نتیجه گرفت.

۱) از هر $(k+1)$ نقطه‌ای که بر یک صفحهٔ $(k-1)$ بُعدی واقع نباشند، یک صفحهٔ k بُعدی می‌گذرد و درضمن تنها یکی.

شباهت کاملی که در این‌جا با حقیقت‌های معلوم هندسهٔ مقدماتی وجود دارد، به‌روشنی دیده می‌شود. اثبات این قضیه، بر نظریهٔ دستگاه‌های معادله‌های خطی تکیه دارد، ولی چنان پیچیده است که بهتر می‌دانیم در این‌جا به آن نپردازیم.

۲) اگر دو صفحهٔ l بُعدی و k بُعدی در فضای n بُعدی، دست‌کم یک نقطهٔ مشترک داشته باشند و درضمن $l+k \geq n$، آن‌وقت آن‌ها در صفحه‌ای که بُعد آن کمتر از $l+k-n$ نیست، یکدیگر را قطع می‌کنند.

به‌عنوان حالت خاص، از این‌جا نتیجه می‌شود، دو صفحهٔ دوبعدی از فضای سه‌بعدی، به‌شرطی که بر هم منطبق یا با هم موازی نباشند، روی یک خط راست با هم برخورد دارند ($n=3$، $l=2$، $k=2$، $l+k-n=1$). ولی در فضای چهاربعدی هم، دو صفحهٔ دوبعدی

---

۱. این k متغیر x_i را، در حالت کلی، نمی‌توان به‌دل‌خواه انتخاب کرد. برای نمونه از دستگاه $x_1+x_2+x_3=0$، $x_1-x_2-x_3=0$، برای x_1 به‌دست می‌آید: $x_1=0$ و درنتیجه x_2 و x_3 را نمی‌توان برحسب آن بیان کرد. ولی می‌توان ثابت کرد، k متغیر x_i را همیشه می‌توان پیدا کرد.

ممکن است تنها یک نقطهٔ مشترک داشته باشند. برای نمونه، صفحه‌هایی که با این دستگاه معادله‌ها داده شده‌اند:

$$\left.\begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array}\right\} \quad , \quad \left.\begin{array}{l} x_3 = 0 \\ x_4 = 0 \end{array}\right\}$$

تنها در یک نقطه با مختصات $x_1 = 0$، $x_2 = 0$، $x_3 = 0$، $x_4 = 0$ یکدیگر را قطع می‌کنند.

اثبات قضیه‌ای که در این‌جا تنظیم کردیم، بی‌اندازه ساده است: صفحهٔ l بُعدی با $(n-l)$ معادله داده می‌شود؛ صفحهٔ k بُعدی با $(n-k)$ معادله داده می‌شود؛ مختصات نقطه‌های برخورد این دو صفحه، باید به‌طور هم‌زمان در همهٔ

$$(n-l) + (n-k) = n - (l+k-n)$$

معادله صدق کنند. اگر هیچ‌یک از معادله‌ها، نتیجه‌ای از دیگر معادله‌ها نباشد، آن‌وقت بنابر تعریف صفحه، محل برخورد یک صفحهٔ $(l+k-n)$ بُعدی است؛ در غیر این صورت صفحه‌ای با بُعد بیشتر به‌دست می‌آید.

به دو قضیه‌ای که در این‌جا آوردیم، می‌توان دو قضیهٔ دیگر هم اضافه کرد.

۳) روی هر صفحهٔ k بُعدی، دست‌کم $(k+1)$ نقطه وجود دارد که نمی‌توانند بر صفحه‌ای با تعداد بُعد کمتر قرار گیرند. در فضای n بُعدی، دست‌کم $n+1$ نقطه وجود دارد که بر هیچ صفحه‌ای قرار ندارند.

۴) اگر خط راستی با یک صفحه (هرچند بُعد که داشته باشد) دو نقطهٔ مشترک داشته باشد، آن‌وقت خط راست به‌طور کامل بر این صفحه واقع است. به‌طور کلی، اگر صفحه‌ای با l بُعد، با صفحه‌ای که k بُعد دارد، $(l+1)$ نقطهٔ مشترکی که بر صفحهٔ $(l-1)$ بُعدی قرار ندارند، داشته باشد، آن‌وقت، این صفحه به‌طور کامل بر این صفحهٔ k بُعدی قرار دارد.

یادآوری می‌کنیم، هندسهٔ n بُعدی را هم می‌توان با تعمیم اصل موضوع‌هایی که در بند ۵ تنظیم کردیم، اصل موضوعی کرد. اگر چنین کنیم، آن‌وقت این چهار قضیه را می‌توان به‌عنوان اصل موضوع‌های وقوع پذیرفت. و این ثابت می‌کند که مفهوم اصل موضوع، مفهومی نسبی است؛ یک گزاره می‌تواند در یک ساختمانِ نظریه، به‌عنوان قضیه در نظر گرفته شود و در ساختمان دیگری از نظریه، به‌عنوان اصل موضوع.

۴. تا این‌جا تصوری کلی دربارهٔ مفهوم‌های ریاضی فضای چندبُعدی پیدا کردیم. برای

این که معنای واقعی و فیزیکی این مفهوم‌ها را بهتر درک کنیم، دوباره به مسألهٔ نمایش نموداری برمی‌گردیم. فرض کنید بخواهیم بستگی بین فشار گاز با حجم آن را نشان دهیم. دستگاه محورهای مختصات را در نظر می‌گیریم و روی یکی از محورها، حجم v و روی دیگری فشار p گاز را نشان می‌دهیم. در این صورت، بستگی فشار به حجم، با شرط‌هایی که داده شده است، به‌وسیلهٔ یک منحنی روی صفحهٔ محورها نشان داده می‌شود (در درجهٔ حرارت مفروض و برای گاز ایده‌آل، این نمودار بنابر قانون ***بویل ـ ماریوت***، یک سهمی است). ولی اگر با دستگاه فیزیکی پیچیده‌تری سروکار داشته باشیم که به‌جای دو متغیر (مثل حجم و فشار در حالت گاز)، در مثل پنج متغیر را به هم مربوط کرده باشد، آن‌وقت نمودار رفتار آن منجر به نمایشی در فضای پنج‌بُعدی می‌شود.

فرض کنید، صحبت بر سر آلیاژی از سه فلز یا مخلوطی از سه گاز باشد. چگونگی مخلوط به چهار عامل بستگی دارد: درجهٔ حرارت T، فشار p و درصد c_1 و c_2 دو گاز (مقدار درصد گاز سوم از این‌جا به‌دست می‌آید که مجموع کل درصدها برابر است با ۱۰۰، به‌نحوی که $c_3 = 100 - c_1 - c_2$). بنابراین، چگونگی این مخلوط با توجه به چهار فرض معین می‌شود. برای نمودار نمایش این چگونگی، باید یا چند نمودار را به‌هم ترکیب کرد و یا پیش خود تصور کرد که، این چگونگی به‌صورت نقطه‌ای در فضای چهاربُعدی و با چهار مختص T، p، c_1 و c_2 مشخص می‌شود. از چنین تصوری در شیمی در عمل استفاده می‌شود؛ به‌کار گرفتن روش‌های هندسهٔ چندبُعدی در مسأله‌های این دانش، به‌وسیلهٔ ***گیبس*** دانشمند امریکایی و ***کورناکوف*** عضو فرهنگستان علوم و مکتب «فیزیک ـ شیمی» اتحاد شوروی، تنظیم شد. در این‌جا، وارد کردن فضای n بُعدی به‌معنای نگه‌داشتن شباهت‌ها و تصورهای هندسی سودمند، با استفاده از روش سادهٔ نمایش نموداری است.

مثال دیگری می‌آوریم که به هندسه مربوط می‌شود. کره به‌یاری چهار فرض معین می‌شود: شعاع و سه مختص مرکز آن. بنابراین، کره را می‌توان با نقطه‌ای در یک فضای چهاربُعدی نشان داد. بنابراین، هندسهٔ خاص کره‌ها، که در میانه‌های سدهٔ نوزدهم به‌وسیلهٔ برخی ریاضی‌دانان ساخته شد، می‌تواند همچون نوعی هندسهٔ چهاربُعدی بررسی شود.

از آن‌چه گفته شد، می‌توان به سرچشمه‌های کلی و عینی فضای چندبُعدی پی برد. اگر شکلی یا حالتی از یک موضوع به‌وسیلهٔ n مقدار داده شده باشد، آن‌وقت این شکل یا این حالت را می‌توان همچون نقطه‌ای در فضای n بُعدی در نظر گرفت. استفاده از این تصور، به‌تقریب شبیه استفاده از نمودارهای معمولی است، یعنی امکان به‌کار گرفتن روش‌ها و

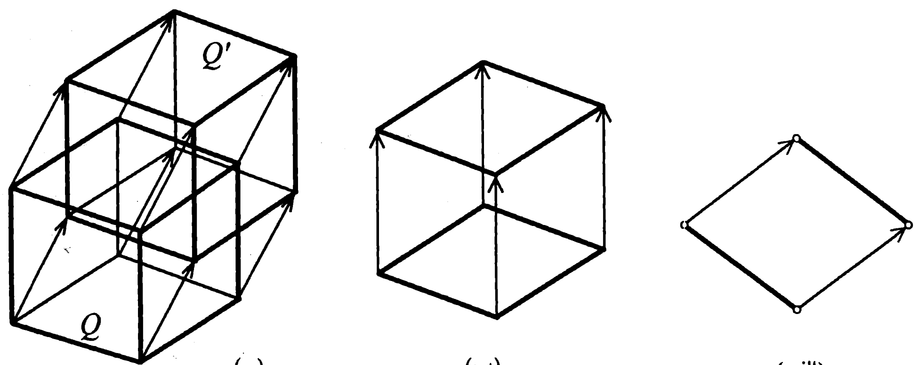
شباهت‌های هندسی معلوم، برای مطالعه پدیده‌ای که بررسی می‌کنیم.

بنابراین، در مفهوم‌های فضای n بُعدی، هیچ‌گونه معمایی وجود ندارد. فضای چندبُعدی گونه‌ای مفهوم انتزاعی است که برای توضیح موضوع‌هایی که به‌یاری هندسهٔ ساده و به‌معنای عادی آن قابل توضیح نیستند، به‌زبان هندسی، به‌وسیلهٔ ریاضی‌دانان پدید آمده است. این مفهوم انتزاعی مبنا و سرچشمه‌ای در واقعیت دارد، واقعیت را بازتاب می‌دهد و نیاز دانش‌ها را برطرف می‌کند و نمی‌توان آن را، بازی بیهوده‌ای با تصورهای ذهنی دانست. مفهوم انتزاعی فضای n بُعدی، این حقیقت را منعکس می‌کند که، موضوع‌هایی مثل کُره یا مخلوطی از سه گاز وجود دارند که بستگی به چند متغیر دارند، که در نتیجه مجموعهٔ همهٔ این موضوع‌ها، چندبُعدی هستند. در هر حالت، تعداد بُعدها، همان تعداد عامل‌هایی است که موضوع ما به آن‌ها بستگی دارد. نقطه‌ای که در فضا حرکت می‌کند، سه مختص دارد. ولی کُره‌ای که کوچک و بزرگ می‌شود، ضمن حرکت خود در فضا، به چهار «مختص» بستگی دارد، یعنی چهار مقداری که آن را مشخص می‌کنند.

در بندهای بعدی، باز هم دربارهٔ هندسهٔ چندبعدی صحبت خواهیم کرد. در این‌جا دوباره روی این مطلب تأکید می‌کنیم که، هندسهٔ چندبُعدی روشی است برای توضیح موضوع‌ها و پدیده‌های واقعی. تصوری که دربارهٔ یک فضای چهاربُعدی به‌وسیلهٔ داستان‌سرایان و تخیل‌گرایان مطرح و چیزی دور از جهان واقع در نظر گرفته می‌شود، هیچ ربطی به مفهوم ریاضیِ فضای چهاربُعدی ندارد. این گمان‌های ذهنی، تنها به کار سرگرمی و، اگر گذشت داشته باشیم، برخی داستان‌های شبه‌علمی می‌خورد.

۵. همان‌طور که گفتیم، هندسهٔ فضای چندبُعدی، در آغاز، از راه تعمیم صوری هندسهٔ تحلیلی معمولی، برای حالت‌هایی که با چند متغیر سر و کار داریم، پدید آمد. ولی این مسیر آغازین، نمی‌توانست ریاضی‌دانان را راضی کند. درواقع هدف، بالاتر از تعمیم مفهوم‌های هندسی و بیشتر در این سمت کشش داشت که روش‌های هندسی بررسی، تعمیم داده شود. بنابراین مهم این بود که طرح خالص هندسی، برای هندسهٔ n بُعدی داده شود، به‌نحوی که ربطی به ابزار تحلیلی نداشته باشد. این تلاش، برای نخستین بار، در سال ۱۸۵۲ میلادی، به‌وسیلهٔ *شِلِفلی* ریاضی‌دان سویسی انجام گرفت که در کارهای خودش، مسألهٔ مربوط به چندوجهی‌های منتظم در فضای چندبُعدی را بررسی کرد. البته، به کار *شِلِفلی* به‌وسیلهٔ هم‌عصرانش توجه زیادی نشد، زیرا برای درک آن لازم بود تا اندازهٔ کم‌وبیش زیادی، دیدگاه انتزاعی به هندسه را پذیرفت. تنها تکامل بعدیِ ریاضیات بر این کارها پرتوی افکند و به

برکت بستگی مفهوم‌های تحلیلی و هندسی به آن روشنی بخشید. بدون این که به‌صورتی جدی وارد این بحث شویم، به چند مثال از طرح هندسیِ هندسهٔ n بُعدی اکتفا می‌کنیم. از تعریف هندسیِ مکعب n بُعدی آغاز می‌کنیم. اگر پاره‌خط راستی را در صفحه در جهت عمود بر خودش تا فاصله‌ای برابر با طول آن حرکت دهیم، یک مربع رسم می‌شود، یعنی یک مکعب دوبُعدی (شکل ۲۳ـالف). به همین ترتیب، اگر مربع را در جهت عمود بر صفحهٔ آن و به فاصله‌ای برابر ضلع آن حرکت دهیم، به مکعب سه‌بعدی می‌رسیم (شکل ۲۳ـب). برای به‌دست آوردن مکعب چهاربعدی هم، ساختمان مشابهی را در نظر می‌گیریم: صفحهٔ سه‌بعدی را در فضای چهاربعدی و در این صفحهٔ سه‌بعدی، مکعب سه‌بعدی را در نظر می‌گیریم؛ سپس این مکعب سه‌بعدی را در جهت عمود بر صفحهٔ سه‌بعدی به فاصله‌ای برابر یال مکعب حرکت می‌دهیم (بنابر تعریف، خط راستی را عمود بر صفحهٔ k بُعدی می‌نامیم که بر همهٔ خط‌های راست واقع بر این صفحه عمود باشد). این ساختمان به‌صورتی مشروط در شکل ۲۳ـج نشان داده شده است. در این‌جا دو مکعب سه‌بعدی Q و Q' ــ وضع مکعب در حالت اولیه و در حالت پایانی ــ دیده می‌شود. خط‌هایی که رأس‌های این مکعب‌ها را به هم پیوسته است، مُعرف پاره‌خط‌های راستی‌اند که رأس‌ها را ضمن جابه‌جایی مکعب در طول خود حرکت می‌دهند. می‌بینیم، مکعب چهاربعدی روی هم ۱۶ رأس دارد: ۸ رأس در مکعب Q و ۸ رأس در مکعب Q'. این مکعب دارای ۳۲ یال است: ۱۲ یال مربوط به مکعب متحرک Q در وضع اولیهٔ خود، ۱۲ یال مربوط به وضع پایانی آن Q' و ۸ یال «جانبی». این مکعب چهاربعدی، ۸ وجه سه‌بعدی دارد که هرکدام از آن‌ها، یک



(ج) (ب) (الف)

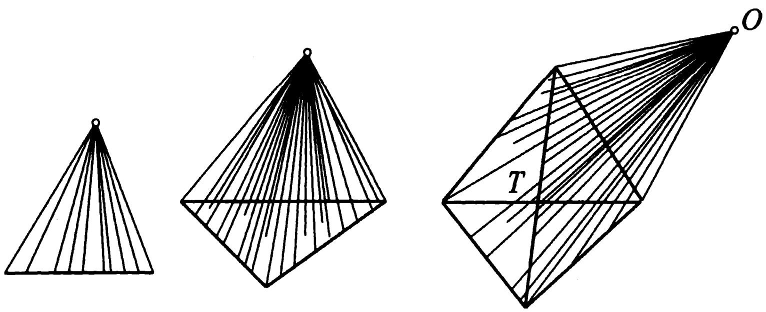
شکل ۲۳

مکعب است. ضمن حرکت مکعب سه‌بعدی، هر وجه آن یک مکعب سه‌بعدی رسم می‌کند، به‌نحوی که ۶ مکعب به‌دست می‌آید (وجه‌های جانبی مکعب چهاربعدی)؛ به‌جز این‌ها، دو وجه دیگر هم داریم: «وجه جلوی» و «وجه عقبی»، متناظر با موقعیت نخستین و موقعیت پایانی مکعب متحرک. سرانجام، مکعب چهاربعدی، دارای وجه‌های مربعی دوبعدی هم به‌تعداد ۲۴ مربع است. هریک از مکعب‌های Q و Q' شش‌تا، و دوازده مربع هم از رسم یال‌های مربع Q، ضمن جابه‌جایی آن، به‌دست می‌آید.

به این ترتیب، مربع چهاربعدی دارای ۸ وجه سه‌بعدی، ۲۴ وجه دوبعدی، ۳۲ وجه یک‌بعدی (یال‌ها) و سرانجام ۱۶ رأس است؛ هر وجه مکعبی با تعداد بُعدهای مربوط به خود است: مکعب سه‌بعدی، مربع، پاره‌خط راست، رأس (رأس با نقطه را هم می‌توان مکعبی به‌حساب آورد که تعداد بُعدهای آن برابر صفر است).

به همین ترتیب، می‌توان با جابه‌جا کردن مکعب چهاربعدی «در بعد پنجم»، مکعبی پنج‌بعدی به‌دست آورد و با ادامهٔ این ساختمان، می‌توان به مکعبی با هر تعداد بُعد رسید. همهٔ وجه‌های مکعب n بعدی، خود مکعب‌هایی هستند که تعداد بُعدهای آن‌ها از n کمتر است: $(n-1)$ بعدی، $(n-2)$ بعدی، ... و سرانجام یک‌بعدی، یعنی یال. مسألهٔ جالب، که چندان هم دشوار نیست، پیدا کردن تعداد وجه‌هایی است که بعدهای معینی دارند. به‌سادگی معلوم می‌شود که یک مکعب n بعدی دارای $2n$ وجه $(n-1)$ بعدی و 2^n رأس است. ولی، از جمله، تعداد یال‌های آن چقدر است؟

نمونهٔ دیگری از چندوجهی در فضای n بعدی می‌آوریم. ساده‌ترین چندضلعی در روی صفحه، مثلث است که آن را سه‌گوشه هم می‌گویند و در بین چندضلعی‌های روی صفحه کمترین تعداد راس را دارد. برای این که یک چندوجهی با کمترین تعداد رأس به‌دست آوریم، کافی است نقطه‌ای را که بر صفحهٔ مثلث واقع نیست، انتخاب و آن را با پاره‌خط راستی به‌هر نقطهٔ این مثلث وصل کنیم. این پاره‌خط‌های راست، سه وجه هرم را پر می‌کنند و یک ***چهاروجهی*** پدید می‌آورند (شکل ۲۴). برای این که ساده‌ترین چندوجهی را در فضای چهاربعدی به‌دست آوریم، به این ترتیب استدلال می‌کنیم. روی صفحه‌ای سه‌بعدی، یک چهاروجهی مثل T در نظر می‌گیریم. سپس نقطه‌ای انتخاب می‌کنیم که روی این صفحهٔ سه‌بعدی نباشد و آن را با پاره‌خط‌های راست به همهٔ نقطه‌های چهاروجهی T می‌پیوندیم. در سمت راست شکل ۲۴، به‌طور مشروط، این ساختمان را نشان داده‌ایم. هریک از پاره‌خط‌های راستی که نقطهٔ O را به نقطه‌ای از چهاروجهی وصل کرده است، با چهاروجهی



شکل ۲۴

نقطه‌های مشترک دیگری ندارد، زیرا در غیر این صورت، چنین پاره‌خط راستی به‌طور کامل در فضای سه‌بعدی قرار می‌گیرد که شامل T است. همهٔ این پاره‌خط‌های راست، «در بُعد چهارم»اند. آن‌ها ساده‌ترین چندوجهی چهاربعدی را پر می‌کنند که آن را **سادک چهاربعدی** (یا **سیمپلکس چهاربعدی**) می‌نامند. وجه‌های سه‌بعدی آن، چهاروجهی هستند: یک وجه در قاعده و چهار وجه جانبی، که به وجه‌های دوبعدی قاعده تکیه دارند؛ روی‌هم ۵ وجه. وجه‌های دوبعدی آن مثلث‌اند؛ روی هم ۱۰ مثلث: ۴ وجه در قاعده و ۶ وجه جانبی. سرانجام، این ساده‌ترین چندوجهی چهاربعدی، ۱۰ یال و ۵ رأس دارد.

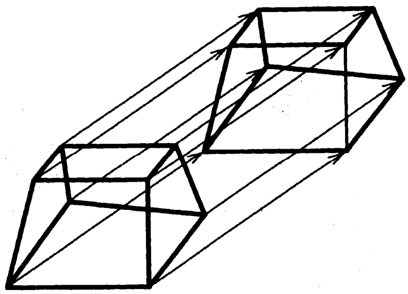
با تکرار همین ساختمان، سرانجام ساده‌ترین چندوجهی n بعدی به‌دست می‌آید که آن را سادَک n بعدی (سیمپلکس n بعدی) گویند. همان‌طور که از ساختمان دیده می‌شود، این ساده‌ترین چندوجهی n بعدی $(n+1)$ رأس دارد. می‌توان متوجه شد که همهٔ وجه‌های آن هم، سادَک‌هایی با تعداد بُعد کمترند: $(n-1)$ بعدی، $(n-2)$ بعدی و غیره[1].

به همین سادگی می‌توان مفهوم‌های منشور و هرم را هم تعمیم داد. اگر یک چندضلعی را موازی با خودش، از صفحه به بُعد سوم ببریم، یک منشور را رسم می‌کند. به همین ترتیب، می‌توان چندوجهی سه‌بعدی را به بُعد چهارم برد و منشور چهاربعدی به‌دست آورد (در شکل ۲۵ ، این منشور چهاربعدی، به‌طور مشروط نشان داده شده است). روشن است که

---

۱. هر m رأس سادک، سادک $(m-1)$ بعدی را که «از آن‌ها پدید آمده است» معین می‌کند، یعنی وجه $(m-1)$ بعدی سادک n بعدی مفروض را. بنابراین، تعداد وجه‌های $(m-1)$ بعدی سادک n بعدی برابر است با تعداد ترکیب‌های m به m از همهٔ $(n+1)$ رأس: یعنی برابر است با

$$C_{n+1}^{m} = \frac{(n+1)!}{m!\,(n-m+1)!}$$



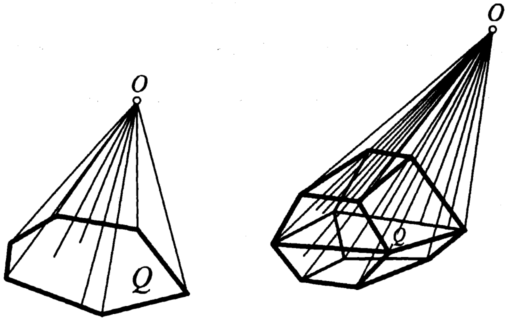
شکل ۲۵

مکعب چهاربعدی، حالت خاصی از منشور چهاربعدی است.

هرم را به این ترتیب می‌توان ساخت. چندضلعی Q و نقطهٔ O را، که بر صفحهٔ چندضلعی واقع نیست، انتخاب می‌کنیم. هر نقطهٔ چندضلعی Q با پاره‌خط راستی به نقطهٔ O وصل می‌شود؛ این پاره‌خط‌های راست هرم به قاعدهٔ Q را پر می‌کنند (شکل ۲۶). به همین ترتیب، اگر در فضای چهاربعدی، چندوجهی سه‌بعدی Q و نقطهٔ O (که همراه با Q در یک صفحهٔ سه‌بعدی نیستند) داده شده باشد، آن‌وقت پاره‌خط‌های راستی که نقطه‌های چندوجهی Q را به O وصل می‌کنند، هرم چهاربعدیِ با قاعدهٔ Q را پر می‌کنند. سادک چهاربعدی، چیزی جز هرم با قاعدهٔ یک چهاروجهی نیست.

درست به همین ترتیب، می‌توان با آغاز از چندوجهی $(n-۱)$ بعدی Q، منشور n بعدی و هرم n بعدی را به‌دست آورد.

به‌طور کلی، چندوجهی n بعدی، بخشی از فضای n بعدی است که به‌وسیلهٔ تعداد محدودی صفحه‌های $(n-۱)$ بعدی محصور شده است؛ چندوجهی k بعدی، بخشی از



شکل ۲۶

صفحهٔ k بعدی است که به‌وسیلهٔ تعداد معینی صفحهٔ $(k-1)$ بعدی محدود شده است. وجه‌های چندوجهی، خود چندوجهی‌هایی با بُعد کمتر هستند.

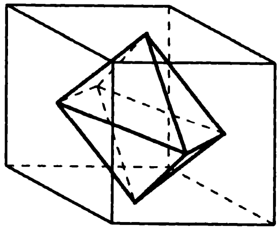
نظریهٔ چندوجهی‌های n بعدی، تعمیمی غنی از نظریهٔ چندوجهی‌های عادی سه‌بعدی عادی است. از جمله، بسیاری از قضیه‌های مربوط به چندوجهی‌های سه‌بعدی، بدون هیچ دشواری، برای هر تعداد بُعد تعمیم داده می‌شوند، ولی در این‌جا مسأله‌هایی وجود دارد که حل آن‌ها، برای چندوجهی‌های n بعدی، بی‌اندازه دشوار می‌شود. در این باره می‌توان از بررسی‌های **گ. ف. وُرونوف** (۱۸۶۸ـ۱۹۰۸) یاد کرد که درواقع، در رابطه با مسأله‌های مربوط به نظریهٔ عددها انجام گرفت و تا زمان ما هم، هندسه‌دانان دنبال آن را گرفته‌اند. یکی از این مسأله‌ها ـ که به «مسألهٔ وُرونوف» مشهور است ـ هنوز هم به‌طور کامل حل نشده است[۱].

نمونه‌ای از این مطلب، که بین فضاهای با بُعدهای مختلف، تفاوت وجود دارد، مربوط به چندوجهی‌های منتظم است. در صفحه، یک چندضلعی منتظم، می‌تواند هر تعداد ضلع داشته باشد؛ به زبان دیگر، بی‌نهایت گونهٔ مختلف از «چندوجهی‌های دوبعدی» منتظم وجود دارد. تعداد چندوجهی‌های منتظم سه‌بعدی، تنها پنج گونهٔ مختلف است: چهاروجهی، مکعب، هشت‌وجهی، دوازده‌وجهی و بیست‌وجهی. در فضای چهاربعدی شش گونه از چندوجهی‌های منتظم یافت می‌شود؛ ولی در هر فضایی که بیش از چهار بُعد داشته باشد، همه‌جا، تعداد چندوجهی‌های منتظم برابر است با سه. این‌ها عبارت‌اند از: ۱) شبیه چهاروجهی ـ سادک n بعدی منتظم، یعنی سادکی که همهٔ یال‌های آن برابر باشند؛ ۲) مکعب n بُعدی؛ ۳) شبیه هشت‌وجهی که به این ترتیب ساخته می‌شود: مرکز هر وجه مکعب، رأسی از این چندوجهی است؛ مثل این که چندوجهی از آن‌ها پدید آمده است. این ساختمان، برای فضای سه‌بعدی، در شکل ۲۷ داده شده است. می‌بینیم، در رابطه با چندوجهی‌های منتظم، هریک از فضاهای دو، سه و چهار بُعدی وضعی ویژهٔ خود دارند.

۶. به مسألهٔ مربوط به حجم جسم‌ها در فضای n بعدی می‌پردازیم. حجم یک جسم n بعدی، شبیه آن‌چه در هندسهٔ عادی انجام می‌شود، به‌دست می‌آید. حجم، به‌معنای نسبت دادن عددی به شکل است؛ درضمن، اگر دو جسم برابر باشند، حجم‌های برابر دارند، یعنی

---

۱. صحبت بر سر جست‌وجوی چنان چندوجهی‌های کوژی است که با آن‌ها بتوان فضا را پر کرد (چندوجهی‌ها باید موازی با هم و با وجه‌های چسبیده به هم باشند). حالت فضای سه‌بعدی این مسأله، در رابطه با نیازهای بلورشناسی، به‌وسیلهٔ **فه‌دوروف** حل شده است؛ وُرونوف همین مسأله را برای حالت n بعدی در برابر خود گذاشت، ولی تنها برای حالت دو، سه و چهار بُعدی حل شد.

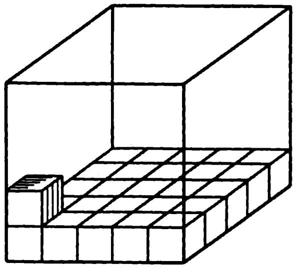


شکل ۲۷

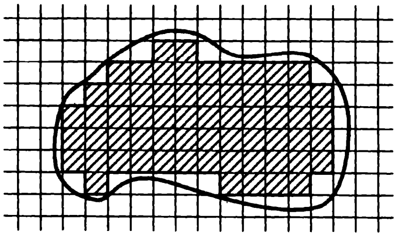
ضمن حرکت شکل، به‌عنوان یک جسم صُلب، حجم آن تغییر نمی‌کند؛ همچنین در حالتی که جسمی که شامل دو جسم جداگانه باشد، حجم آن برابر است با مجموع حجم‌های دو جسم. به‌عنوان واحد حجم، مکعبی در نظر گرفته می‌شود که یال آن، طولی برابر واحد داشته باشد. با این فرض، ثابت می‌شود، مکعبی که یال آن به‌طول a باشد، حجمی برابر a^n دارد. این اثبات شبیه اثباتی که در هندسهٔ مقدماتی و برای فضای سه‌بعدی معمول است، یعنی با پر کردن مکعب به‌وسیلهٔ لایه‌های از مکعب واحد، انجام می‌شود (شکل ۲۸). چون مکعب واحد را در n جهت می‌توان قرار داد، پس حجم مکعب n بعدی به یال a، برابر است با a^n.

برای تعیین حجم هر جسم n بعدی، به‌جای آن به‌تقریب جسمی را در نظر می‌گیریم که بتواند با مکعب‌های n بعدی پر شود، شبیه شکل ۲۹ که به‌جای شکل اصلی، شکلی پله‌ای در نظرگرفته‌ایم که با مربع‌ها پر شده است. حجم جسم به‌عنوان حد چنین جسم‌های پله‌ای، با کوچک و کوچکتر کردن مکعب‌هایی که در آن قرار داده‌ایم، محاسبه می‌شود.

حجم شکل k بعدی هم که در یک صفحهٔ k بعدی قرار دارد، درست به همین ترتیب



شکل ۲۸



شکل ۲۹

محاسبه می‌شود. از تعریف حجم، بلافاصله نتیجهٔ مهمی به‌دست می‌آید: اگر جسم را متشابه با خود بزرگ کنیم، به‌نحوی که همهٔ اندازه‌های خطی آن λ برابر شود، حجم جسم k بعدی، λ^k برابر می‌شود.

اگر جسمی را به لایه‌های موازی با هم تقسیم کنیم، آن‌وقت حجم آن، برابر مجموع حجم‌های این لایه‌هاست:

$$V = \sum V_i$$

حجم هر لایه را می‌توان به‌تقریب، حاصل‌ضرب ارتفاع آن Δh_i در حجم $(n-1)$ بعدی («مساحت») مقطع S_i متناظر با آن دانست. درنتیجه، حجم تمامی جسم، به‌تقریب برابر است با مجموع

$$V \approx \sum S_i \, \Delta h_i$$

که اگر همهٔ Δh_i را به سمت صفر میل دهیم، حجم جسم به‌صورت انتگرال

$$V = \int_0^H S(h)\, dh \qquad (۹)$$

درمی‌آید، که در آن، H عبارت است از امتداد عمود بر مقطع‌ها.

همهٔ این‌ها، شبیه محاسبهٔ حجم در فضای سه‌بعدی است. برای نمونه، در منشور همهٔ مقطع‌ها با هم برابرند و بنابراین، «مساحت» S مقطع، به ارتفاع h بستگی ندارد. درنتیجه برای منشور $V = S \cdot H$، یعنی حجم منشور برابر است با حاصل‌ضرب «مساحت» قاعده در ارتفاع. حجم هرم n بعدی را محاسبه می‌کنیم. هرمی را در نظر می‌گیریم که ارتفاع آن H و مساحت قاعدهٔ آن S باشد. آن را با صفحه‌ای موازی با قاعده و به‌فاصله h از رأس قطع

می‌کنیم. در این صورت، هرمی با ارتفاع h به‌دست می‌آید. مساحت قاعدهٔ این هرم به $s(h)$ نشان می‌دهیم. روشن است که این هرم کوچکتر، با هرم اصلی متشابه است: همهٔ اندازه‌های هرم کوچکتر نسبت به اندازه‌های متناظر خود در هرم اصلی، مثل نسبت h به H است، یعنی همهٔ آن‌ها در $\frac{h}{H}$ ضرب شده‌اند. بنابراین، حجم $(n-1)$ بعدی (یعنی «مساحت») قاعدهٔ آن برابر است با

$$s(h)=\left(\frac{h}{H}\right)^{n-1}S$$

زیرا اگر اندازه‌های خطیِ شکل $(n-1)$ بعدی λ برابر شود، حجم آن در λ^{n-1} ضرب می‌شود.

مساحت هرم اصلی، بنابر دستور (۹)، برابر است با

$$V=\int_0^H s(h)\,dh$$

و از آن‌جا

$$V=\int_0^H S\left(\frac{h}{H}\right)^{n-1}dh=\frac{S}{H^{n-1}}\int_0^H h^{n-1}dh=\frac{S}{H^{n-1}}\cdot\frac{H^n}{n}=\frac{1}{n}SH$$

یعنی حجم هرم n بعدی برابر با $\frac{1}{n}$ حاصل‌ضرب «مساحت» [حجم $(n-1)$ بعدی] قاعده در ارتفاع. وقتی n را برابر ۲ یا ۳ بگیریم، به همان دستورهای آشنا می‌رسیم: مساحت (حجم دوبعدی) مثلث برابر است با نصف حاصل‌ضرب قاعده در ارتفاع؛ حجم هرم سه‌بعدی برابر است با یک‌سوم مساحت قاعده در ارتفاع.

کره را می‌توان به‌تقریب، به‌عنوان مجموعه‌ای از هرم‌های بسیار نازک در نظر گرفت که رأس مشترک همهٔ آن‌ها در مرکز کره باشد. ارتفاع هریک از این هرم‌ها، برابر شعاع R است و از مجموع همهٔ σ_iها (قاعده‌های هرم‌ها)، به‌تقریب مساحت S از سطح کره به‌دست می‌آید. چون حجم هر هرم برابر است با $\frac{1}{n}R\sigma_i$، بنابراین با روی هم گذاشتن همهٔ این حجم‌ها، برای حجم کره به‌دست می‌آید:

$$V\approx\frac{1}{n}R\sum\sigma_i\approx\frac{1}{n}RS$$

که در حد به دستور دقیقِ $V=\frac{1}{n}RS$ می‌رسد، یعنی حجم کره برابر است با $\frac{1}{n}$ حاصل‌ضرب

شعاع در مساحت سطح آن. اگر n را برابر ۲ یا ۳ بگیریم، همان دستورهای مشهور به‌دست می‌آید[1].

ویژگی مهمی از کره را یادآوری می‌کنیم که برای فضای n بعدی، درست شبیه حالت فضای سه‌بعدی قابل اثبات است. بین همهٔ جسم‌های با حجم مفروض، کمترین مساحت سطح متعلق به کره، و تنها کره است.

۷. ما تنها به هندسهٔ مقدماتی فضای n بعدی پرداختیم، ولی واقع این است که می‌توان آن را تکامل داد و هندسهٔ «عالی» را از جمله برای نظریهٔ عمومی منحنی‌ها (یا خم‌ها) و سطح‌ها (یا رویه‌ها) مطرح کرد. سطح در فضای n بعدی می‌تواند بُعدهای مختلفی داشته باشد: «سطح» یک‌بعدی، یعنی منحنی‌ها، سطح دوبعدی، سطح سه‌بعدی، ... ، و سرانجام سطح $(n-۱)$ بُعدی. منحنی را می‌توان همچون مکان هندسی نقطه‌هایی تعریف کرد که، مختصات آن‌ها، به‌طور پیوسته به یک متغیر ـ پارامتر t ـ بستگی داشته باشند:

$$x_۱ = x_۱(t) \ \ , \ \ x_۲ = x_۲(t) \ \ , \ \ \dots \ \ , \ \ x_n = x_n(t)$$

منحنی، عبارت است از مسیر حرکت نقطه‌ای که با واسطهٔ پارامتر t تغییرمکان می‌دهد، در فضای n بعدی. وقتی فضا، برای نمایش حالت یک دستگاه فیزیکی به‌کار می‌رود (مثل آن‌چه در شماره ۴ آوردیم)، آن‌وقت منحنی معرف تغییر حالت‌ها یا حرکت‌های پیوسته و متوالی، در بستگی با پارامتر t است (که در مثل می‌تواند زمان باشد). و این، تعمیمی از نمایش نموداری معمولی روندهاست که حالت‌های مختلف آن‌ها به‌وسیلهٔ منحنی‌ها نشان داده می‌شود.

در هر نقطهٔ منحنی فضای n بعدی، نه‌تنها یک خط مماس («صفحهٔ یک‌بعدی مماس»)، بلکه «صفحه‌های مماس» در همهٔ بعدها از ۱ تا $n-۱$ وجود دارد. سرعت دوران هریک از صفحه‌ها، نسبت به سرعت طول منحنی، انحنای متناظر آن را معین می‌کند. بنابراین منحنی دارای $(n-۱)$ صفحهٔ مماس است، از یک‌بعدی تا $(n-۱)$ بعدی و درضمن $(n-۱)$ انحنای متناظر با آن‌ها. هندسهٔ دیفرانسیلی در فضای n بعدی بسیار پیچیده‌تر از فضای سه‌بعدی است. در این‌جا، مکان آن را نداریم که دربارهٔ نظریهٔ سطح‌ها صحبت کنیم.

---

۱. حجم کره را با استفاده از دستور (۹) هم می‌توان محاسبه کرد؛ مقطع‌های کرهٔ n بُعدی، کره‌های $(n-۱)$ بعدی است و بنابراین، حجم کرهٔ n بعدی، با روش عبور از $(n-۱)$ به n، به‌دست می‌آید.

تا این‌جا گفت‌وگو از هندسهٔ n بعدی بود که تعمیم مستقیمی از هندسهٔ معمولی اقلیدسی است. ولی می‌دانیم، به‌جز هندسهٔ اقلیدسی، هندسه‌های دیگری هم وجود دارند: هندسهٔ لباچوسکی، هندسهٔ تصویری وغیره. این هندسه‌ها را هم می‌توان به‌سادگی و برای هر تعداد بعد فضا تعمیم داد.

۸. تعمیم موضوع هندسه

۱. وقتی دربارهٔ مفهوم واقعی فضای n بعدی و تطبیق آن در جهان واقع صحبت می‌کردیم، به مسألهٔ تعمیم موضوع هندسه رسیدیم، یعنی به مسألهٔ مربوط به عام بودن مفهوم فضا در ریاضیات. ولی پیش از آن که به تعریف‌های کلی بپردازیم، چند مثال را بررسی می‌کنیم.

تجربه ثابت می‌کند که بینایی انسان عادی، همان‌طور که برای نخستین بار م. و. **لومونوسُف** یادآوری کرده است، «سه‌رنگی» است، یعنی هر احساس رنگ (C) ترکیبی از سه احساس اصلی است: قرمز (R)، سبز (G) و آبی (B)، با شدت‌های معین[1]. اگر این شدت‌ها را برحسب واحدهایی مثل x، y و z در نظر بگیریم، می‌توانیم بنویسیم:

$$C = xR + yG + zB$$

همان‌طور که نقطه می‌تواند در فضا، بالا و پایین، راست و چپ ویا جلو و عقب حرکت کند، همان‌طور هم، احساس رنگ ـ نور C می‌تواند، در سه جهت، با تغییر ترکیب آن از بخش‌های قرمز، سبز و آبی، پیوسته تغییر کند. بنابراین، می‌توان گفت که، مجموعهٔ همهٔ رنگ‌های ممکن «فضای رنگی» سه‌بعدی را تشکیل می‌دهند. در این فضا، شدت‌های x، y و z نقش مختصات نقطه (نور C) را به‌عهده دارند (اختلاف مهمی که نسبت به مختصات معمولی

---

۱. صحبت بر سر رنگ است نه روشنایی. احساس رنگ هم، یک پدیدهٔ عینی است که در نور منعکس می‌شود. همین احساس ممکن است به‌وسیلهٔ موج‌های نوری مختلف به‌وجود آید: برای نمونه، رنگ سبز می‌تواند، نه‌تنها از نور سبز خالصی که از طیف دیده می‌شود، بلکه همچنین از اختلاط قرمز و آبی به‌دست آید. از طرف دیگر، انسان‌هایی که از «کوری رنگ» [دالتونیسم، منسوب به جون دالتون (۱۷۶۶-۱۸۴۴)] رنج می‌برند، تنها دو احساس اصلی دارند: حالت‌هایی که نسبت به رنگ «کوری کامل» دارند و وقتی که تنها یک احساس اصلی رنگ و بی‌اندازه «کم‌رنگ» دارند.

وجود دارد، این است که شدت‌ها نمی‌توانند مقدارهای منفی را بپذیرند. در حالت $x=y=z=0$، با رنگ کامل سیاه سر و کار داریم که مُعرف نبودن هرگونه نوری است).

تغییر پیوستهٔ نور را می‌توان همچون یک منحنی در «فضای رنگی» نشان داد؛ این منحنی در مثل شامل رنگین‌کمان است؛ همچنین «منحنی» رنگ‌ها، معرف یک‌رشته احساس است که ضمن تغییر پیوستهٔ شفافیت روشنایی در اثر رنگ‌های مختلف پدید می‌آید، درحالی که «رنگی بودن» آن ثابت است.

سپس، اگر رنگ قرمز (R) و سفید (W) داده شده باشد، آن‌وقت با مخلوط کردن آن‌ها به نسبت‌های مختلف[1]، به‌طور پیوسته و پشت‌سرهم، رنگ‌های از R تا W به‌دست می‌آید که می‌توان آن را پاره‌خط راست RW نامید. این تصور که بین قرمز و سفید، رنگ‌های مختلفی وجود دارد، مفهومی روشن است.

به این ترتیب، مفهوم شکل‌های سادهٔ هندسی و نسبت‌های هندسی در «فضای رنگ‌ها» پدید می‌آید. «نقطه» یعنی رنگ، «پاره‌خط راست» AB، یعنی مجموعه‌ای از رنگ‌ها که از مخلوط کردن رنگ‌های A و B به‌دست می‌آید؛ معنای این جمله که «نقطهٔ D بر پاره‌خط راست AB واقع است»، این است که رنگ D، مخلوطی از رنگ‌های A و B است. با مخلوط کردن سه رنگ، تکه‌ای از صفحه به‌دست می‌آید که می‌توان آن را «مثلث رنگ‌ها» نامید. همهٔ این‌ها را هم می‌توان به‌صورت تحلیلی و با استفاده از مختصات x، y و z رنگ‌ها شرح داد؛ درضمن دستورهای مربوط به پاره‌خط‌های راست و صفحه‌ها در «فضای رنگ‌ها»، به‌طور کامل شبیه دستورهای هندسهٔ تحلیلی معمولی است[2].

در فضای رنگ‌ها، رابطه‌های هندسهٔ اقلیدسی دربارهٔ استقرار نقطه‌ها و پاره‌خط‌های راست، برقرار است. بررسی این رابطه‌ها، هندسهٔ آفین را تشکیل می‌دهد و می‌توان گفت،

---

۱. این مخلوط را می‌توان از این راه به‌دست آورد که گردهای رنگین نرم را به نسبت‌های مختلف به‌هم آمیخت، با این شرط که روشنایی تغییر نکند.

۲. برای نمونه، اگر رنگ C_0 و C_1 مربوط به شدت‌هایی با مختصات x_0، y_0، z_0 و x_1، y_1، z_1 باشند، آن‌وقت رنگ C، که بین C_0 و C_1 قرار دارد، مختصاتی برابر

$$x=(1-t)x_0+tx_1,\quad y=(1-t)y_0+ty_1,\quad z=(1-t)z_0+tz_1$$

خواهد داشت که در آن، t نمایندهٔ سهم رنگ C_1 و $(1-t)$ سهم رنگ C_0 در مخلوطی است که رنگ C را ساخته‌اند.

هندسهٔ آفین با مجموعهٔ همهٔ امکان‌هایی که برای احساس رنگ‌ها وجود دارد، سازگار است (البته، این مطلب به‌طور کامل و دقیق، درست نیست، زیرا همان‌طور که گفتیم، مختصات x، y، z رنگ‌ها نمی‌توانند منفی باشند. بنابراین فضای رنگ‌ها، تنها پاسخ‌گوی آن بخش از فضاست که در دستگاه محورهای مختصات، همهٔ نقطه‌ها با مختصات مثبت یا صفر باشند).

درضمن تصوری طبیعی دربارهٔ درجهٔ اختلاف رنگ‌ها پیدا می‌کنیم. برای نمونه، معلوم می‌شود که گُلی روشن، به رنگ سفید نزدیک‌تر است تا گُلی سیر، یا تمشکی به‌قرمز نزدیک‌تر است تا آبی وغیره. به این ترتیب، برای فاصلهٔ بین رنگ‌ها، به‌عنوان درجهٔ اختلاف آن‌ها، مفهوم کیفی به‌دست می‌آید. این مفهوم کیفی را می‌توان از روی اندازه‌های کمیتی تشخیص داد. ولی تعیین فاصلهٔ بین رنگ‌ها، آن‌طور که در هندسهٔ اقلیدسی با دستور

$$r=\sqrt{(x_{\circ}-x_{1})^{2}+(y_{\circ}-y_{1})^{2}+(z_{\circ}-z_{1})^{2}}$$

به‌دست می‌آید، غیرطبیعی به‌نظر می‌رسد، زیرا فاصلهٔ تعیین‌شده ممکن است با احساس ما تطبیق نکند: پیش می‌آید که برای دو رنگ با درجه‌های مختلف نسبت به رنگ مفروض، فاصله‌های برابر به‌دست آید. تعیین فاصله، باید رابطه‌های واقعیِ بین احساس رنگ‌ها را منعکس کند.

با توجه به این مطلب، در فضای رنگ‌ها، برای اندازه‌گیری فاصله‌ها، روش خاصی انتخاب شده است که به این ترتیب انجام می‌شود.

وقتی رنگ به‌صورتی پیوسته تغییر می‌کند، انسان بلافاصله این تغییر را احساس نمی‌کند؛ تغییر رنگ تنها وقتی احساس می‌شود که به مرحلهٔ خاصی به‌نام «آستانهٔ اختلاف» برسد. با توجه به این نکته، همهٔ رنگ‌هایی را که از رنگ مفروض در فاصلهٔ تا آستانهٔ اختلاف قرار دارند، هم‌فاصله از آن به‌حساب می‌آورند. از این‌جا، به‌خودی‌خود، به این نتیجه می‌رسیم که فاصلهٔ بین هر دو رنگ را باید به‌عنوان کمترین تعداد آستانه‌های اختلاف در نظر گرفت که می‌توان بین آن‌ها قرار داد. طول خط رنگی، عبارت است از تعداد آستانه‌های اختلافی که روی آن قرار دارد و فاصلهٔ بین دو رنگ، کوتاه‌ترین فاصله‌ای است که شامل منحنی بین این دو رنگ باشد. این شبیه به آن است که فاصلهٔ بین دو نقطه از سطح را، به‌عنوان طول کوتاه‌ترین مسیر بین آن‌ها در نظر می‌گیریم.

به این ترتیب، اندازه‌گیری طول و فاصله در فضای رنگ‌ها، منجر به گام‌های خیلی کوچک (بلکه بی‌نهایت کوچک) می‌شود.

بنابراین، در فضای رنگ‌ها، برخی ویژگی‌های فضای نااقلیدسی مخصوص به خودش تعریف می‌شود. این فضا، مفهومی واقعی دارد: این فضا می‌تواند ویژگی‌های مجموعه‌ای از رنگ‌های ممکن را شرح دهد، یعنی ویژگی‌هایی که در اثر تحریک نوری، موجب عکس‌العمل در چشم می‌شود.

مفهوم فضای رنگ‌ها در آغاز سدهٔ نوزدهم به‌وجود آمد. دربارهٔ هندسهٔ این فضا بسیاری از فیزیک‌دانان، از جمله *هِلم‌هولتس* و *ماکسوِل* بررسی‌هایی انجام دادند. پس از آن‌ها، این بررسی‌ها، چه در زمینهٔ نظری و چه در زمینهٔ عملی، همچنان ادامه یافت. درنتیجه، مبنایی ریاضی برای حل مسأله‌های مربوط به اختلاف رنگ‌های علامت‌ها (سیگنال‌ها)، دربارهٔ رنگ‌ها در صنعت نساجی وغیره به‌دست آمد.

۲. به مثال دیگری می‌پردازیم که پیش از این هم دربارهٔ آن، اندکی صحبت کرده‌ایم.

فرض کنید بخواهیم دربارهٔ دستگاهی «فیزیکی ـ شیمیایی» مثل اختلاط گازها، آلیاژ یا چیز دیگری مطالعه کنیم. فرض کنیم حالت این دستگاه به‌وسیلهٔ n مقدار معین شود (مثل حالت مخلوط گازی که با فشار، درجهٔ حرارت و تراکم بخش‌های سازندهٔ آن معین می‌شود). در این صورت گویند، دستگاه دارای n درجه آزادی است، یعنی چگونگی آن را می‌توان در n جهت مستقل از یکدیگر، با تغییر هریک از این مقدارها، تغییر داد. این مقدارهایی که حالت دستگاه را معین می‌کنند، نقش مختصات آن را به‌عهده دارند. بنابراین، مجموعهٔ همهٔ حالت‌های آن، همچون یک فضای n بعدی بررسی می‌شود که آن را، «فضای فازی (یا مرحله‌ای) دستگاه» گویند.

تغییر پیوستهٔ چگونگی‌ها، یعنی روندهایی را که در دستگاه پیش می‌آید، به‌وسیلهٔ منحنی‌هایی در این فضا نشان می‌دهند. حوزه‌های جداگانهٔ حالت‌ها، که آن‌ها را با نوعی نشانه از هم جدا می‌کنیم، معرف حوزه‌های فضای مرحله‌ای (فازی) هستند. حالتی که برای دو حوزه، مرز به‌حساب می‌آید، سطحی را در این فضا تشکیل می‌دهند.

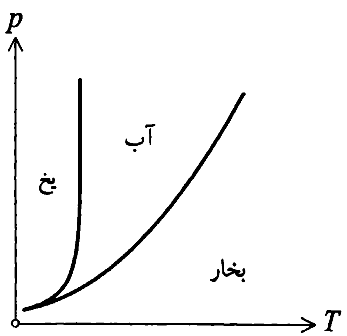
در شیمی به‌ویژه بررسی شکل و همجواری متقابل حوزه‌هایی از فضای مرحله‌ای دستگاه اهمیت دارد که متناظر با حالت‌های کیفی مختلف‌اند. سطح‌هایی که این حوزه‌ها را از هم جدا می‌کنند، متناظر با چنان سطح‌های کیفی مثل گداز، تبخیر، رسوب وغیره هستند. حالت‌های دستگاهی که دو درجهٔ آزادی داشته باشد، با نقطه‌های روی صفحه نشان داده

می‌شود. نمونهٔ این وضع، می‌تواند جسم متجانسی باشد که حالت آن به‌کمک فشار (p) و درجهٔ حرارت (T) معین می‌شود. همین p و T مختصات نقطه‌ای هستند که حالت را مشخص می‌کند. در این صورت کار منجر به تقسیم حوزه‌ها می‌رسد که پاسخ‌گوی حالت‌های مختلف کیفی‌اند. برای آب، این حوزه‌ها عبارت‌اند از حوزه‌های یخ، آب مایع و بخار (شکل ۳۰). منحنی‌های تقسیم پاسخ‌گوی حالت‌های گذر یخ به آب و از آن به بخار است.

وقتی با بررسی دستگاهی سر و کار داشته باشیم که تعداد بیشتری درجهٔ آزادی داشته باشد، باید از هندسهٔ چندبُعدی استفاده کنیم.

مفهوم فضای فازی (مرحله‌ای) نه‌تنها در دستگاه‌های شیمیایی، بلکه در دستگاه‌های مکانیکی و به‌طور کلی در هر دستگاهی که، حالت‌های ممکن آن، مجموعهٔ پیوسته‌ای را تشکیل دهند، کاربرد دارد. برای نمونه، در نظریهٔ سینتیک گازها، دستگاه فضای فازیِ ذره‌های مادی (مولکول‌های گاز) بررسی می‌شود. چگونگی حرکت یک ذره در هر لحظه، از روی موضع و سرعت آن معین می‌شود که روی هم شش مقدارند: سه مختص و سه مؤلفهٔ سرعت (روی سه محور مختصات). چگونگی N ذره با $6N$ مقدار داده می‌شود و چون مولکول‌ها خیلی زیادند، $6N$ عدد بزرگی از آب درمی‌آید. این، همان فضای فازی $6N$ بعدی دستگاه مولکول‌هاست.

در این فضا، نقطه، مُعرف حالت همهٔ تودهٔ مولکول‌ها، به‌وسیلهٔ مختصات و سرعت‌های آن‌هاست. حرکت نقطه، تغییرحالت را نشان می‌دهد. این تصور انتزاعی، برای همهٔ نتیجه‌گیری‌های نظری عمیق، بسیار سودمند است. سخن کوتاه، مفهوم فضای فازی به‌عنوان سلاحی برای دانش‌های دقیق طبیعی درآمده و برای حل بسیاری از دشواری‌ها به‌کار می‌رود.



شکل ۳۰

۳. اکنون، با مثال‌هایی که آوردیم، می‌توانیم به نتیجه‌گیری دربارهٔ تعمیم موضوع هندسه بپردازیم.

فرض کنید، بخواهیم مجموعهٔ پیوسته‌ای از برخی موضوع‌ها، یا روندها و یا حالت‌ها و در مثل، مجموعهٔ همه گونه‌های مختلف و ممکنِ رنگ‌ها یا مجموعهٔ حالت‌های گروهی از مولکول‌ها را بررسی کنیم. رابطه‌هایی که در چنین مجموعه‌ای وجود دارد، ممکن است با رابطه‌های فضای عادی شباهت داشته باشند، مثل «فاصلهٔ» بین رنگ‌ها یا «وضع متقابل» حوزه‌های فضای فازی. در چنین حالتی، با جدا کردن ویژگی‌های کیفی موضوع‌هایی که بررسی می‌کنیم و تنها توجه به رابطه‌های بین آن‌ها، می‌توانیم مجموعهٔ مفروض را همچون یک فضای خاص مطالعه کنیم. «نقطه‌های» این «فضا» عبارت‌اند از خود موضوع‌ها: پدیده‌ها یا چگونگی‌ها. «شکل» در این «فضا»، به هر مجموعه‌ای از «نقطه‌ها» گفته می‌شود، مثل «منحنی» رنگ‌های رنگین‌کمان یا «حوزهٔ» بخار در «فضای» حالت‌های آب. «هندسهٔ» چنین فضایی به‌عنوان رابطه‌هایی که بین موضوع‌ها، پدیده‌ها و یا حالت‌ها وجود دارد، تعریف می‌شود. از جمله، «هندسهٔ» فضای رنگ‌ها با قانون‌های ترکیب رنگ‌ها و فاصله‌های بین رنگ‌ها تعریف می‌شود.

معنای واقعی این دیدگاه و سازگاری آن با جهان واقع در این است که به ما امکان می‌دهد، از مفهوم‌ها و روش‌های هندسهٔ انتزاعی برای بررسی پدیده‌های گوناگون استفاده کنیم. به این ترتیب، میدان کاربرد مفهوم‌ها و روش‌های هندسی، بی‌اندازه گسترده می‌شود. درنتیجهٔ تعمیم مفهوم فضا، اصطلاح «فضا» دو معنای متفاوت در دانش پیدا می‌کند: از یک‌سو، با فضایی سر و کار داریم که فضای واقعی عادی، یعنی صورت عام وجود ماده را بازتاب می‌دهد؛ از سوی دیگر، یک «فضای انتزاعی»، یعنی مجموعه‌ای از موضوع‌های همگن، مثل پدیده‌ها، حالت‌ها وغیره است که در آن رابطه‌هایی مشابه رابطه‌های فضایی وجود دارد.

یادآوری کنیم، فضای عادی هم، که تصور آن برای ما تا اندازه‌ای ساده‌تر است، مجموعه‌ای از حالت‌های همگن است. از جمله، می‌توان آن را به‌عنوان مجموعهٔ همهٔ وضع‌های همگنی که یک جسم بسیار کوچک ــ یعنی «نقطهٔ مادی» ــ می‌تواند داشته باشد، در نظر گرفت. این یادآوری را به این معنا نگرفته‌ایم که توانسته‌ایم تعریفی برای فضا بیاوریم، بلکه هدف این بود که بستگی روشن‌تری بین دو مفهوم مختلف فضا برقرار کنیم. دربارهٔ مفهوم فضای انتزاعی و دربارهٔ رابطهٔ هندسهٔ انتزاعی با هندسهٔ عادیِ فضای واقعی باز هم

گفت‌وگو خواهیم کرد.

۴. گسترده‌ترین کاربرد فضای انتزاعی، در خود ریاضیات است. در هندسه، می‌توان از «فضایی» صحبت کرد که به این یا آن شکل بستگی دارد: «فضای کره‌ها» که پیش از این از آن یاد کردیم، «فضای خط‌های راست» وغیره.

این روش کار، به‌ویژه در بررسی‌های مربوط به نظریهٔ چندوجهی‌ها، بی‌اندازه سودمند است. برای نمونه، در بند ۵ بخش هفتم (جلد دوم) دربارهٔ قضیهٔ مربوط به وجود چندوجهی کوژ، به‌شرط مفروض بودن گستردهٔ آن، صحبت کرده‌ایم. اثبات این قضیه برپایهٔ بررسی دو فضا قرار دارد: «فضای چندوجهی‌ها» و «فضای گسترده‌ها». مجموعهٔ چندوجهی‌های کوژ، که تعداد مفروضی رأس داشته باشند، همچون فضایی در نظر گرفته می‌شود که در آن، هر نقطه، نمایندهٔ یک چندوجهی است؛ به همین ترتیب، مجموعهٔ گسترده‌های ممکن هم، به‌عنوان فضایی در نظر گرفته می‌شود که هر نقطهٔ آن، نمایندهٔ یک گسترده است. ساختن چندوجهی‌ها به‌کمک گستردهٔ آن‌ها، یعنی پدید آوردن تناظر بین چندوجهی‌ها و گسترده‌ها، به زبان دیگر، متناظر کردن نقطه‌های «فضای چندوجهی‌ها» با نقطه‌های «فضای گسترده‌ها». مسأله به این‌جا می‌رسد که ثابت کنیم، هر گسترده‌ای پاسخ‌گوی یک چندوجهی است، یعنی هر نقطهٔ یک فضا، پاسخ‌گوی نقطه‌ای از فضای دیگر است. و این هم، به‌طور مستقیم با استفاده از توپولوژی ثابت می‌شود.

به همین ترتیب، می‌توان یک‌رشته از قضیه‌های مربوط به چندوجهی‌ها را ثابت کرد؛ درضمن، این «روش فضای انتزاعی»، در بسیاری حالت‌ها (مثل قضیهٔ مربوط به وجود چندوجهی با مفروض بودن گستردهٔ آن)، بسیار ساده‌تر از روش‌های شناخته‌شدهٔ دیگری است که برای اثبات قضیه وجود دارد. ولی با کمال تأسف، این روش به‌اندازهٔ کافی پیچیده است و ما نمی‌توانیم در این‌جا تصور دقیق‌تری دربارهٔ آن بدهیم.

مفهوم تعمیم‌یافتهٔ فضا، در آنالیز، جبر و نظریهٔ عددها هم، کاربردهای گسترده‌ای پیدا کرده است. آغاز این کاربرد را باید در تصور عادی تابع‌ها به یاری منحنی‌ها دانست. یک متغیر x را، می‌توان به‌طور ساده، متناظر با نقطهٔ واقع بر یک خط راست دانست. به همین ترتیب، دو متغیر متناظر با نقطه‌ای واقع بر صفحه و n متغیر متناظر با نقطه‌ای از فضای n بعدی است: n متغیر x_1، x_2، ... ، x_n متناظر است با نقطه‌ای با مختصات $(x_1, x_2, \ldots, x_n)$. در این‌جا صحبت بر سر «دامنهٔ» تابع یا «بُرد» تابع $f(x_1, x_2, \ldots, x_n)$، نقطه‌ها یا منحنی‌ها سطح‌های ناپیوستگی تابع‌ها وغیره است. این زبان هندسی همه‌جا به‌کار می‌رود و تنها

مربوط به روش بیان نیست؛ تصور هندسی، بسیاری از حقیقت‌های آنالیز را، به‌صورت «عینی» و شبیه تصورهایی که از فضای عادی داریم، در برابر ما می‌گذارد و به ما اجازه می‌دهد، برای اثبات آن‌ها، از روش‌های هندسی در فضای n بعدی استفاده کنیم.

در جبر هم، وقتی با معادله‌های شامل n مجهول یا تابع‌های جبری شامل n متغیر کار می‌کنیم، با همین موقعیت روبه‌رو می‌شویم. پیش از این هم دیده‌ایم، معادلهٔ خطی با n مجهول، در فضای n بعدی، با یک صفحه تعریف می‌شود؛ m معادله از این‌گونه، معرف m صفحه است و هر جواب یک دستگاه m معادلهٔ n مجهولی با نقطه‌ای بیان می‌شود که در همهٔ این صفحه‌ها مشترک است. صفحه‌ها ممکن است در یک نقطه مشترک نباشند، بلکه روی یک خط راست یا روی صفحه‌ای دوبُعدی یا، به‌طور کلی، صفحه‌ای k بُعدی یکدیگر را قطع کنند. به این ترتیب، مسألهٔ حل یک دستگاه معادله‌های خطی، منجر به مسألهٔ پیدا کردن تقاطع صفحه‌ها می‌شود. این روش هندسی، برتری‌هایی دارد. به‌طور کلی، مسأله‌های جبر خطی که به‌طور عمده دربارهٔ معادله‌های خطی بحث می‌کند، و تبدیل‌های خطی، اغلب و همان‌طور که در بخش پانزدهم دیدیم، تا اندازه زیادی، با روش‌های هندسی حل می‌شوند.

۵. در همهٔ مثال‌هایی که آوردیم، صحبت بر سر مجموعهٔ پیوسته‌ای از این یا آن موضوع‌هاست که به‌عنوان یک «فضا» در نظر گرفته می‌شود. موضوع‌های این مجموعه، می‌تواند رنگ‌ها، حالت‌های یک دستگاه، شکل‌ها، مجموعهٔ مقدارهای متغیرها و... باشد. در هر حالت، موضوع به‌وسیلهٔ تعداد معین و محدودی فرض داده می‌شود و بنابراین، با فضایی متناظر می‌شود که بُعد معین و محدودی، برابر با تعداد این فرض‌ها، دارد.

ولی از آغاز سدهٔ بیستم، «فضای بی‌نهایت‌بُعدی» هم، نظر ریاضی‌دانان را به خود جلب کرد که متناظر با چنان موضوع‌هایی است که، هرکدام از آن‌ها را، نمی‌توان با تعداد معین و محدودی فرض داد. و این، قبل از همه، «فضاهای تابعی» است.

اندیشهٔ تفسیر مجموعهٔ تابع‌های این یا آن گونه به‌عنوان یک فضای خاص، یکی از اندیشه‌های اصلی شاخهٔ تازه‌ای از آنالیز، یعنی آنالیز تابعی بود، که برای حل و بررسی بسیاری از پرسش‌ها بی‌اندازه سودمند افتاد. خواننده می‌تواند تصوری از این شاخه را در بخش نوزدهم، که به آنالیز تابعی اختصاص دارد، به‌دست آورد.

می‌توان مجموعهٔ تابع‌های پیوستهٔ شامل یک یا چند متغیر را در نظر گرفت. همچنین می‌توان کلاس گوناگونی از تابع‌های ناپیوسته را به‌عنوان «فضا» در نظر گرفت و بسته به نوع

مسأله‌ای که مطرح است، فاصلهٔ بین آن‌ها را تعریف کرد. درنتیجه، تعداد «فضاهای تابعی» ممکن، بی‌پایان است، و در عمل بسیاری از این گونه‌ها را بررسی می‌کنند.

به همین ترتیب می‌توان «فضای منحنی‌ها»، «فضای جسم‌های کوژ»، «همهٔ گونه‌های حرکت یک دستگاه مکانیکی» وغیره را در نظر گرفت. از جمله در بند ۵ بخش هفتم (جلد دوم) از قضیه‌هایی یاد کردیم که، روی هر سطح کوژ بسته، دست‌کم سه خط بستهٔ ژئودزیک وجود دارد و هر دو نقطه‌ای از آن را می‌توان با بی‌نهایت خط ژئودزیک به هم وصل کرد. برای اثبات این قضیه‌ها از فضای منحنی‌های واقع بر سطح استفاده می‌کنند: در قضیهٔ اول از فضای منحنی‌های بسته و در قضیهٔ دوم از فضای منحنی‌هایی که دو نقطهٔ مفروض را به هم می‌پیوندند. در مجموعهٔ همهٔ منحنی‌های ممکن، که دو نقطه مفروض را به هم می‌پیوندند، تعریف خاصی از فاصله را در نظر می‌گیرند که از این راه، مجموعه به فضا تبدیل می‌شود. اثبات قضیه، براساس کاربرد برخی نتیجه‌گیری‌های عمیق توپولوژی در این فضاست.

اکنون نتیجهٔ کلی را تنظیم می‌کنیم.

در ریاضیات، منظور از «فضا»، هر مجموعه‌ای از موضوع‌های همگن است (مثل پدیده‌ها، حالت‌ها، تابع‌ها، شکل‌ها، مقدارهای متغیر وغیره) که بین آن‌ها رابطه‌هایی شبیه رابطه‌های فضای معمولی وجود دارد (پیوستگی، فاصله وغیره). درضمن برای بررسی چنین مجموعه‌ای از موضوع‌ها به‌عنوان فضا، همهٔ ویژگی‌های این موضوع‌ها، به‌جز رابطه‌هایی که در این فضا، مشابه فضای معمولی، قبول کرده‌ایم، کنار گذاشته می‌شود. این رابطه‌ها به‌گونه‌ای تعریف شده‌اند که بتوان آن‌ها را روی هم، ساختمان یا «هندسهٔ» فضا نامید. در این فضا، خود موضوع‌ها، نقش «نقطه‌ها» را به‌عهده دارند؛ «شکل‌ها» به‌عنوان مجموعه‌ای از نقطه‌ها در نظر گرفته می‌شوند.

در هندسهٔ این فضای انتزاعی، موضوع بحث عبارت است از چنان ویژگی‌هایی از فضا و شکل‌های آن، که براساس تعریف رابطه‌های درونی این فضا به‌دست می‌آیند. از جمله، ضمن بررسی فضای تابع‌های پیوسته، هیچ کاری به ویژگی‌های تابع‌های جداگانه نداریم. در این‌جا، هر تابع نقش یک نقطه را به‌عهده دارد و درنتیجه «دارای جزئی نیست»، به این مفهوم هیچ ساختمانی و هیچ ویژگی در بیرون از بستگی با دیگر نقطه‌ها ندارد، به زبان دقیق‌تر، جدا از همهٔ این‌ها در نظر گرفته می‌شود. در فضای تابعی، ویژگی‌های تابع‌ها، تنها برحسب بستگی آن‌ها با سایر تابع‌ها ــ برحسب فاصله‌های آن‌ها و برحسب رابطه‌های دیگری که ناشی از فاصله است ــ معین می‌شود.

امکان‌های گوناگونی که برای مجموعه‌های ممکنِ موضوع‌ها و رابطه‌های مختلف بین آن‌ها وجود دارد، پاسخ‌گوی گونه‌های متفاوت فضاهاست که تعدادی بی‌پایان را تشکیل می‌دهند و در ریاضیات به بررسی گذاشته می‌شوند. فضاها را می‌توان، برحسب آن رابطه‌های فضایی که پایهٔ تعریف فضا را تشکیل داده‌اند، تقسیم‌بندی کرد. بدون این که بخواهیم از همهٔ گونه‌های مختلف فضاهای انتزاعی یاد کنیم، پیش از همه، به دو گونهٔ آن توجه می‌کنیم: فضای توپولوژیک و فضای متری.

۶. ***فضای توپولوژیک*** (بخش هجدهم را ببینید) به هر مجموعه‌ای از نقطه‌ها گفته می‌شود که در آن، رابطهٔ همسایگی یک نقطه در مجموعهٔ نقطه‌ها و متناظر با آن رابطهٔ همسایگی یا چسبیدگی دو مجموعه (شکل) نسبت به هم تعریف شده باشد. و این تعمیمی از همسایگی شکل‌ها در فضای معمولی است.

لباچوسکی هم، با فراست بی‌اندازهٔ خود، یادآوری می‌کند که اساسی‌ترین رابطه از میان همهٔ رابطه‌های شکل‌ها، رابطهٔ همسایگی است. او می‌نویسد: «***همسایگی و نزدیکی***، که دربارهٔ جسم‌ها وجود دارد، وقتی **هندسی** نامیده می‌شود که در استدلال‌های خود آن را نگه داریم و همهٔ ویژگی‌های دیگر را، چه همیشگی و چه تصادفی، کنار بگذاریم». برای نمونه، هر نقطه از محیط دایره با مجموعهٔ همهٔ نقطه‌های درونی آن چسبیده و همسایه است؛ دو بخش یک جسم یک‌پارچهٔ مربوط به هم، چسبیده به یکدیگرند (یا در همسایگی هم قرار دارند). ضمن پیشرفت توپولوژی، همین ویژگی «همسایگی»، اساس همهٔ دیگر ویژگی‌های توپولوژیک بوده است.

مفهوم «همسایگی»، به‌معنای تصور دربارهٔ نزدیکی بی‌نهایتِ نقطه به مجموعه است. بنابراین، هر مجموعه‌ای از موضوع‌ها که در آن، مفهوم طبیعی مربوط به پیوستگی و دربارهٔ نزدیکی و تقریب بی‌پایان وجود داشته باشد، یک فضای توپولوژیک است.

فضای توپولوژیک، مفهومی بی‌اندازه کلی دارد و مطالعهٔ چنین فضاهایی ــ توپولوژی انتزاعی ــ چیزی جز این نیست که بررسی ریاضی کلی‌تری دربارهٔ مفهوم پیوستگی داشته باشیم.

تعریف ریاضی دقیق فضای توپولوژیک را می‌توان به این ترتیب داد[1].

فضای توپولوژیک کلی به مجموعه‌ای مثل R، شامل عضوهایی («نقطه‌هایی»)، گفته

---

۱. تعریفی که در اینجا آمده، تعریف عادی و کنونی فضای توپولوژیک نیست (ویراستار).

می‌شود، به شرطی که در این مجموعه، برای هر زیرمجموعهٔ M از آن، نقطه‌های همسایگی وجود داشته باشد که با شرط‌های (اصل موضوع‌های) زیر بسازند:

۱) هر نقطهٔ مجموعهٔ M جزو نقطه‌های همسایگی آن باشد (طبیعی است بگوییم، مجموعه به هر نقطهٔ خود چسبیده است).

۲) اگر مجموعهٔ M_1 شامل مجموعهٔ M_2 است، آن‌وقت مجموعهٔ نقطه‌های همسایگی M_1، شامل همهٔ نقطه‌های همسایگی M_2 باشد (اگر کوتاه، ولی نه‌چندان دقیق بگوییم: در مجموعهٔ بزرگتر، تعداد نقطه‌های همسایگی کمتر نباشد).

به این‌ها، اصل موضوع‌های دیگری هم، که بستگی به نوع فضای توپولوژیک دارند، افزوده می‌شود.

با دستگیرهٔ مفهوم همسایگی، به‌سادگی می‌توان یک‌رشته از مهم‌ترین مفهوم‌های توپولوژیک را نتیجه گرفت. این مفهوم‌ها، درضمن، کلی‌ترین و اساسی‌ترین مفهوم‌های هندسه‌اند و تعریف آن‌ها به‌مراتب عینی‌تر و روشن‌ترند. مثال‌هایی می‌آوریم.

۱) ***چسبیدگی مجموعه‌ها.*** مجموعه‌های M_1 و M_2 را چسبیده بر یکدیگر گوییم، به‌شرطی که یکی از آن‌ها، دست‌کم شامل یک نقطه در همسایگی دیگری داشته باشد (به این مفهوم، از جمله، محیط دایره، چسبیده به درون دایره است).

۲) ***پیوستگی یا آن‌طور که ریاضی‌دانان می‌گویند، همبند بودن شکل.*** شکل، یعنی مجموعهٔ نقطه‌های M را، همبند گوییم، به‌شرطی که نتوان آن را به‌صورت دو بخش غیرچسبیده در نظر گرفت (از جمله، پاره‌خط راست همبند و پاره‌خط راست بدون نقطهٔ میانِ آن، غیرهمبند است).

۳) **مرز.** مرز مجموعهٔ M در فضای R، به مجموعهٔ همهٔ نقطه‌هایی گفته می‌شود که هم به خود مجموعهٔ M و هم به مجموعهٔ متمم آن $R-M$، یعنی بخش باقی‌ماندهٔ فضای R، چسبیده باشد (و این همان تعریف طبیعی مفهوم مرز است).

۴) ***نقطه‌های درونی.*** نقطه‌ای از مجموعهٔ M را درونی گوییم، وقتی که متعلق به مرز آن، یعنی چسبیده به مجموعهٔ متممِ $R-M$ نباشد.

۵) ***نگاشت یا تبدیل پیوسته.*** تبدیل مجموعهٔ M را پیوسته گوییم، وقتی که، ضمن تبدیل، همسایگی‌ها به‌هم نریزند (احتمال نمی‌رود، بتوان تعریف طبیعی‌تری برای تبدیل پیوسته پیدا کرد).

به این سیاهه، می‌توان تعریف‌های مهم دیگری را هم افزود، از جمله تعریف مفهوم

هم‌گرایی دنباله‌ای از شکل‌ها به سمت شکل مفروض، یا تعریف مفهوم تعداد بُعدهای فضا.

می‌بینیم، بسیاری از مفهوم‌های اساسی هندسه را می‌توان به یاری مفهوم همسایگی تعریف کرد. ارزش توپولوژی به‌ویژه در این است که، برای این مفهوم‌ها، تعریف‌های کلی و دقیقی ارائه می‌دهد و اساسی را پایه می‌گذارد که کاربرد دقیق ملاحظه‌های مربوط به درک عینی پیوستگی را پیدا کنیم.

توپولوژی عبارت است از آموزش چنان ویژگی‌هایی از فضاها، شکل‌های درون آن‌ها و تبدیل‌های آن‌ها، که با رابطهٔ همسایگی تعریف می‌شوند.

این رابطه (یعنی رابطهٔ همسایگی) چنان کلی و عمیق است که توپولوژی را به کلی‌ترین نظریهٔ هندسی تبدیل و در شاخه‌های مختلف ریاضیات ــ هرجا که صحبتی از پیوستگی باشد ــ نفوذ کرده است. ولی درست به‌دلیل همین نیروی کلی بودن، توپولوژی در بیشتر شاخه‌های انتزاعی خود، از چارچوب خاص هندسه بیرون رفته است. با وجود این، پایهٔ آن بر ویژگی‌های فضای واقعی قرار دارد و سودمندی و نیروی نتیجه‌گیری‌های آن به روش‌هایی مربوط می‌شود که سرچشمه‌ای در تصورهای هندسی عینی دارند. از جمله می‌توان از روش کلی تقریب شکل‌ها به وسیلهٔ چندوجهی‌ها نام برد که به‌وسیلهٔ پ. س. *آلکساندرف* تکامل یافت و، گرچه به‌صورت انتزاعی، درگونه‌های بی‌اندازهٔ کلی فضای توپولوژیک راه پیدا کرد.

امروز، هر ویژه‌کاری که به بررسی موضوع‌هایی مشغول است، متوجه می‌شود که به‌طور طبیعی به سمت نزدیکی و چسبیدگی کشیده می‌شود و درنتیجه با ابزاری به‌نام توپولوژی سر و کار پیدا می‌کند و از این راه نتیجه‌هایی به‌دست می‌آورد که برای جنبه‌های خاص کار او، بی‌معنی نیست.

۷. *فضای متری*، به مجموعه‌ای از عضوها (نقطه‌ها) گفته می‌شود که برای آن‌ها، مفهوم فاصله تعریف شده باشد، یعنی به هر دو نقطهٔ X و Y، عدد $r(X,Y)$[1]، به‌نحوی نسبت داده شده باشد که با این شرط‌ها (اصل موضوع‌های فضای متری) سازگار باشد:

۱) $r(X,Y)=0$، وقتی و تنها وقتی که نقطه‌های X و Y بر هم منطبق باشند.

۲) برای هر سه نقطهٔ X، Y و Z داشته باشیم:

$$r(X,Y)+r(Y,Z)\geq r(Z,X)$$

به این شرط «نابرابری مثلثی» گویند، زیرا شبیه ویژگی فاصله‌های معمولی بین نقطه‌های

---

۱. نویسنده به طور ضمنی فرض کرده است که $r(X,Y)=r(Y,X)$ (ویراستار).

A_1 ، A_2 و A_3 در فضای اقلیدسی است (شکل ۳۱):

$$r(A_1, A_2) + r(A_2, A_3) \geq r(A_3, A_1)$$

به‌عنوان فضای متری، می‌توان این نمونه‌ها را آورد:

۱) فضای اقلیدسی با هر بُعدی؛

۲) فضای *لباچوسکی*؛

۳) هر سطحی با متریک درونی خودش (جلد دوم، بخش هفتم، بند ۴)؛

۴) فضای C از تابع‌های پیوسته که فاصلهٔ بین آن‌ها با دستور زیر بیان می‌شود:

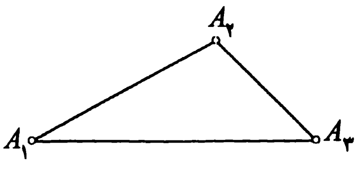
$$r(f_1, f_2) = \max \left| f_1(x) - f_2(x) \right|$$

۵) فضای هیلبرتی (بخش نوزدهم را ببینید) که چیزی جز فضای «اقلیدسی بی‌نهایت بُعدی» نیست.

فضای هیلبرتی، مهم‌ترین فضایی است که در آنالیز تابعی کاربرد دارد؛ این فضا رابطهٔ تنگاتنگی با نظریهٔ رشته‌های فوریه و به‌طور کلی، با نظریهٔ بسط تابع‌ها به رشته به‌وسیلهٔ تابع‌های متعامد دارد (در این‌جا، x_1، x_2، x_3، ... ، ضریب‌های این رشته‌ها هستند). این فضا، در فیزیکِ ریاضی نقشی اساسی دارد و در مکانیک کوانتایی اهمیت زیادی پیدا کرده است. معلوم شده است، مجموعهٔ همهٔ حالت‌های ممکن در یک دستگاه اتمی، مثل اتم هیدروژن، از دیدگاه انتزاعی، می‌تواند همچون فضای هیلبرتی بررسی شود.

تعداد فضاهای متری را، که در ریاضیات بررسی می‌شود، می‌توان به‌مراتب افزایش داد؛ در بند بعد با یکی از مهم‌ترین فضاهای متری به‌نام فضای ریمانی آشنا خواهیم شد، ولی همین نمونه‌هایی هم که در این‌جا آوردیم، برای درک اهمیت مفهوم فضای متری، می‌تواند کافی باشد.

در فضای متری، همیشه می‌توان همهٔ مفهوم‌های توپولوژیک را تعریف کرد، ولی به‌جز



شکل ۳۱

آن‌ها، مفهوم‌های دیگری هم (مفهوم‌های «متری») وجود دارد، مثل مفهوم طول منحنی‌ها. در هر فضای متری، طول را به همان صورت معمولی تعریف می‌کنند که همان ویژگی‌های اصلی طول معمولی را هم دارد. طول منحنی به‌معنای حد مجموع فاصلهٔ بین نقطه‌هایی که پشت‌سرهم روی منحنی قرار دارند (X_1، X_2، ...، X_m) فهمیده می‌شود، به‌شرطی که این نقطه‌ها، مرتب انبوه و انبوه‌تر شوند.

۸. در ریاضیات، به‌جز فضای توپولوژیک و فضای متری، فضاهای بسیار دیگری هم بررسی می‌شود. با یک‌رشته از این فضاها، در بند ۶ آشنا شدیم. در هریک از این فضاها، مسألهٔ گروهی از تبدیل‌ها مطرح است (مثل فضای تصویری یا فضای آفین). در این‌گونه فضاها، می‌توان «برابری» شکل‌ها را تعریف کرد. دو شکل وقتی «برابرند» که ضمن تبدیل از گروه مفروضی، بتوان یکی را به دیگری منجر کرد.

نمی‌خواهیم در تعریف همهٔ گونه‌های ممکنِ فضا وارد شویم؛ این فضاها بسیار زیاد و مختلف‌اند و خواننده می‌تواند برای آشنایی با آن‌ها، به کتاب خاصی که دربارهٔ شاخه‌های مختلف هندسهٔ امروزی نوشته شده است، مراجعه کند.

ولی ببینیم چه عامل‌هایی موجب چنین گسترشی در مفهوم‌های هندسی شده است؟ به چه مناسبت، از جمله، باید فضای تابع‌های پیوسته را در ریاضیات وارد کرد و به بررسی آن پرداخت؟ آیا نمی‌توان مسأله‌های آنالیز را با امکان‌های عادی و بدون توسل به چنین فضاهای انتزاعی، حل کرد؟

پاسخ کوتاه و کلی به این پرسش این است که با وارد کردن فضاهای تازه در بررسی‌های ریاضی، مسیرهایی را کشف می‌کنیم تا بتوانیم از مفهوم‌ها و روش‌های هندسی، که بسیار زیادند، استفاده کنیم.

ویژگی مفهوم‌ها و روش‌های هندسی در این است که، سر آخر، همراه با تصورهای عینی هستند و این برتری را، حتی برای انتزاعی‌ترین موضوع‌ها حفظ می‌کنند. آن‌چه یک تحلیل‌گر می‌تواند بعد از انجام عمل‌های طولانی به‌دست آورد، هندسه‌دان کم‌وبیش بلافاصله نتیجه می‌گیرد. نمونهٔ ساده و مقدماتیِ این برتری را می‌توان در نمودارها دید که به‌روشنی بستگی‌های بین کمیت‌ها را به ما نشان می‌دهد. روش هندسی، روشی ترکیبی است، که برخلاف روش تحلیلی، قادر است همه‌چیز را با هم در برابر ما بگذارد. البته در نظریه‌های انتزاعی هندسی، نمی‌توانیم بلافاصله به عینی بودن موضوع دست یابیم، ولی تصورهای

عینی دربارهٔ موضوع، و هم ویژگی ترکیبی روش هندسی، به قوت خود باقی می‌ماند.

خواننده با بسیاری از کاربردهای مفهوم‌ها و روش‌های هندسی آشناست: کاربرد طرح‌های هندسی در آنالیز، تصور هندسی عددهای مختلط و تابع‌های با متغیرهای مختلط، استدلال و اثبات برای قضیهٔ اصلی جبر و کاربردهای دیگر مفهوم‌ها و روش‌های هندسی. هندسه در همه‌جا موضوع کلی و چگونگی کلی بحث را به ما می‌شناساند. در آغاز بند ۷ و در همین بند در شمارهٔ ۵، نمونه‌هایی از قضیه‌هایی دیدیم که اثبات آن‌ها، با استفاده از هندسهٔ چندبُعدی بلافاصله به‌دست می‌آمد. در این‌جا مثال دیگری از آنالیز می‌آوریم که با به‌کار گرفتن مفهوم فضای تابعی حل می‌شود.

در توپولوژی ثابت می‌شود، اگر حوزه‌ای از صفحهٔ معمولی را در نظر بگیریم که شکلی تحریف‌شده از دایره داشته باشد، سپس آن را به‌صورتی پیوسته تا آن‌جا که می‌خواهیم تحریف کنیم و تغییرشکل دهیم، به‌شرطی که به‌هرحال در درون دورهٔ نخستینِ خود باقی بماند، آن‌وقت دست‌کم یک نقطه پیدا می‌شود که، بعد از تبدیل، در جای نخستینِ خود قرار دارد. این، حقیقتی است خاص توپولوژی.

اکنون مسأله‌ای را در نظر می‌گیریم که به‌کلی دور از هندسه است: می‌خواهیم تابع $y(x)$ را پیدا کنیم که در این معادلهٔ دیفرانسیلی صدق کند[1]:

$$y' = f(x, y) \tag{۱۰}$$

درضمن می‌دانیم، برای $x = 0$ داریم $y = 0$.

روشن است، به‌جای معادلهٔ (۱۰) می‌توان این معادله را در نظر گرفت:

$$y = \int_0^x f(t, y(t))\,dt \tag{۱۱}$$

پرسشی طبیعی پیش می‌آید: آیا تابع $y(x)$ وجود دارد که در این معادله صدق کند؟

به مسأله با دید دیگری نگاه می‌کنیم. هر تابع پیوستهٔ $y(x)$ را، نقطه‌ای از فضای انتزاعی می‌گیریم. نتیجهٔ محاسبهٔ انتگرال

$$\int_0^x f(t, y(t))\,dt = z(x)$$

---

۱. $f(x,y)$ را تابعی پیوسته از متغیرهای x و y فرض کرده‌ایم.

تابعی پیوسته از x است، یعنی نقطه‌ای از فضای انتزاعی ما. با انتخاب «نقطه‌های» مختلف y، یعنی تابع‌های مختلف $y(x)$، در حالت کلی، نقطه‌های مختلفی برای z به‌دست می‌آید. به این ترتیب، مجموعهٔ نقطه‌های فضای ما، دوباره بر همان نقطه‌ها نگاشته می‌شود. مسألهٔ مربوط به وجود جواب برای معادلهٔ (۱۱) منجر به این مسأله می‌شود که: آیا در فضای انتزاعی ما «نقطه‌ای» پیدا می‌شود که بعد از چنین تبدیلی، بر «جای» پیشین خودش قرار گیرد؟

طبیعی‌ترین پرسش در نظریهٔ معادله‌های دیفرانسیلی، به پرسشی دربارهٔ ویژگی «فضای تابعی» انتزاعی منجر می‌شود. قضیه‌ای که در بالا آوردیم، به ما تلقین می‌کند که صحبت بر سر ویژگی توپولوژیک فضای متناظر با آن است.

در این مسیر است که به یاری بررسی‌های لازم توپولوژیک، اثبات بسیاری از قضیه‌ها دربارهٔ وجود جواب در معادله‌های دیفرانسیلی به‌دست می‌آید که ممکن است خیلی ساده‌تر و کوتاه‌تر باشد. به‌ویژه روشن می‌شود که معادلهٔ (۱۰)، برای هر تابع پیوستهٔ $f(x, y)$ به‌واقع دارای جواب است.

۹. هندسهٔ ریمانی

۱. این اندیشه که هر مجموعهٔ پیوسته‌ای از پدیده‌های همگن را می‌توان همچون فضای خاصی تفسیر کرد، برای نخستین بار به‌وسیلهٔ ***ریمان*** در سخن‌رانی‌های او «دربارهٔ فرضیه‌های مربوط به بنیان‌های هندسه»، که در سال ۱۸۵۴ در دانشگاه گوتینگن ایراد شده بود، مطرح شد. این، یک سخن‌رانی آزمایشی، از نوع یک گزارش یا رسالهٔ علمی بود که هر دانشیار یا استادی باید در برابر دانشکده ارائه دهد. ***ریمان***، دیدگاه خود را دربارهٔ نظریهٔ گسترده‌ای از هندسه ــ که امروز هندسهٔ ریمانی نامیده می‌شود ــ در خط کلی خود و بدون اثبات اندیشه‌های اصلی، بیان کرد. می‌گویند، هیچ‌کدام از شنوندگان، چیزی از صحبت‌های او نفهمیدند، به‌جز ***گوس*** که در آن زمان در سال‌های پیری به‌سر می‌برد. ***ریمان*** در نوشتهٔ دیگری، ابزار صوری هندسهٔ خود را، در مسألهٔ مربوط به رسانایی گرما به‌کار برد، به‌نحوی که هندسهٔ انتزاعی ریمانی، در بستگی تنگاتنگ با فیزیکِ ریاضی متولد شد. اندیشهٔ ***ریمان***، گامی جدی بعد از لباچوسکی، در مسیر تکامل هندسه بود. با وجود این، اندیشهٔ ***ریمان***،

خیلی زود نتوانست از طرف دیگران، به‌صورت شایسته‌ای ارزیابی شود، سخن‌رانی او و کار او دربارهٔ رسانایی گرما، تنها در سال ۱۸۶۸ و پس از مرگ او چاپ شد. یادآوری این مطلب لازم است که در سال ۱۸۶۸ ، نخستین تفسیر از هندسهٔ *لباچوسکی* به‌وسیلهٔ بلترام و در سال ۱۸۷۲ تفسیر دیگری به‌وسیلهٔ *کلاین* انجام گرفت. *کلاین* دیدگاهی کلی را دربارهٔ هندسه‌های گوناگون تنظیم کرد: هندسهٔ اقلیدسی، هندسهٔ *لباچوسکی*، هندسهٔ تصویری، هندسهٔ آفین وغیره، همچون بررسی ویژگی‌هایی از شکل که با تبدیل‌هایی از این یا آن گروه تغییر نمی‌کنند، در نظر گرفته شد. در همین سال‌ها بود که مفهوم هندسهٔ چندبُعدی، جای خود را به‌قطعی محکم کرد. به این ترتیب، سال‌های ۷۰ سدهٔ نوزدهم را باید، سال‌های دگرگونی در تاریخ هندسه دانست، وقتی که اندیشه‌های تازهٔ هندسی، که در جریان پنجاه سال روی هم جمع شده بود، به‌وسیلهٔ گروه گسترده‌ای از ریاضی‌دانان پذیرفته شد و درضمن به‌طور جدی وارد در دانش‌ها شد. در چنین اوضاع و احوالی بود که کارهای *ریمان* ادامه پیدا کرد و در پایان سدهٔ نوزدهم، هندسهٔ ریمانی به شکوفایی خود رسید و در مکانیک و فیزیک، کاربردهای خود را پیدا کرد. وقتی که در سال ۱۹۱۵ ، *اینشتین* در کار خودش دربارهٔ نظریهٔ عمومی نسبیت، از هندسهٔ ریمانی در نظریهٔ جاذبهٔ عمومی استفاده کرد، توجه به هندسهٔ ریمانی بیشتر شد و تکامل توفانی خود را آغاز کرد و در جهت‌های مختلف تعمیم یافت.

۲. اندیشه‌های *ریمان*، در ادامهٔ درخشان خود، به‌اندازهٔ کافی ساده و روشن شد، به‌شرطی که از استدلال‌ها و دستورهای ریاضی جدا شویم و توجه خود را، تنها به ماهیت اصلی آن معطوف کنیم. این سادگی، خاص همهٔ اندیشه‌های بزرگ است. آیا اندیشهٔ *لباچوسکی*، که با نفی پوستولای پنجم *اقلیدس*، به‌عنوان یکی از هندسه‌های ممکن، آغاز شد، ساده نبود؟ آیا اندیشهٔ اصلی تکامل موجودات زنده یا اندیشهٔ اصلی ساختمان اتمی ماده، ساده نیست؟ همهٔ این‌ها ساده‌اند، ولی درعین‌حال بسیار پیچیده، زیرا اندیشه‌های تازه، در درجهٔ اول باید مسیر خود را به پهنای لازم باز کنند، نه این که در چارچوب خود باقی بمانند، در درجهٔ دوم، تعمیم همه‌جانبهٔ یک اندیشهٔ تازه، تکامل آن و یافتن راه‌های کاربرد آن، به کاری سترگ نیازمند است که بدون یاری ابزارهای دانش‌های اختصاصی ممکن نیست. برای هندسهٔ ریمانی، این ابزار عبارت‌اند از دستورهای آن؛ این دستورها بغرنج‌اند و بنابراین، تنها برای ویژه‌کاران قابل دستیابی است. ولی در این‌جا، بدون این که دربارهٔ این دستورهای بغرنج بیندیشیم، به جوهر و ماهیت اندیشهٔ *ریمان* توجه

می‌کنیم. همان‌طور که گفتیم، **ریمان** از بررسی مجموعهٔ پیوستهٔ دل‌خواهی از پدیده‌ها، به‌عنوان یک فضا، آغاز کرد. در این فضا، مختصات هر نقطه عبارت است از مقدارهایی که پدیدهٔ متناظر را از پدیده‌های دیگر جدا می‌کند، مثل شدت x، y، z که تعیین‌کنندهٔ رنگ $C=xR+yG+zB$ است. اگر تعداد این مقدارها n باشد (x_1، x_2، ... ، x_n)، آن‌وقت با فضای n بعدی سر و کار داریم. در این فضا می‌توان منحنی‌ها را بررسی کرد و اندازهٔ طول آن‌ها را با گام‌های کوچک (بی‌نهایت کوچک)، شبیه اندازه‌گیری طول منحنی در فضای معمولی، به‌دست آورد.

برای این که طول گام‌های بی‌نهایت کوچک را اندازه بگیریم، کافی است قانونی را بدهیم که به یاری آن، فاصلهٔ هر نقطهٔ مفروض، تا هر نقطهٔ بی‌نهایت نزدیک به آن را معین کند. این قانونی که فاصله را معین می‌کند، ***اندازهٔ حدی*** یا **متریک** نامیده می‌شود. ساده‌ترین حالت، زمانی است که این قانون، به همان قانونی تبدیل می‌شود که در فضای اقلیدسی از آن استفاده می‌کنیم؛ و این، یک فضای اقلیدسی در مقیاس بی‌نهایت کوچک است. به زبان دیگر، رابطه‌های هندسیِ هندسهٔ اقلیدسی در آن صدق می‌کند، تنها به‌شرطی که با حوزهٔ بی‌نهایت کوچکی از آن سر و کار داشته باشیم؛ راستش این است که این رابطه‌ها برای هر حوزهٔ بی‌نهایت کوچک درست‌اند، ولی نه با دقت کامل؛ هرچه حوزه کوچکتر باشد، این دقت بیشتر است. فضایی که در آن فاصله‌ها را با این قانون اندازه می‌گیرند، ***فضای ریمانی*** و هندسهٔ این فضا را، ***هندسهٔ ریمانی*** گویند. بنابراین، فضای ریمانی، فضایی که «در بی‌نهایت کوچک» خود، یک فضای اقلیدسی است.

ساده‌ترین مثال برای فضای ریمانی، هر سطح صاف و همواری با هندسهٔ ذاتی (یا درونی) آن است. هندسهٔ ذاتی سطح، عبارت است از هندسهٔ ریمانی دوبُعدی. درواقع، هر سطحی در نزدیکی هر نقطهٔ خود، نسبت به صفحهٔ مماس، تفاوت اندکی دارد و هرچه سطح کوچکتری در بررسی ما واقع باشد، این اختلاف هم کوچکتر می‌شود. بنابراین، هندسه در حوزهٔ کوچکی از سطح، با هندسهٔ روی صفحه اختلاف اندکی دارد و هرچه این حوزه کوچکتر باشد، اختلاف هم کمتر می‌شود. ولی در حوزه‌های بزرگ، هندسهٔ سطح خمیده با هندسهٔ اقلیدسی متفاوت است، همان‌طور که در بند ۴ بخش ششم (جلد دوم) دیدیم و همان‌طور که به‌سادگی در نمونه‌های کره و شبه‌کره دیده می‌شود. هندسهٔ ریمانی چیزی نیستِ جز تعمیم طبیعی هندسهٔ ذاتی سطح دوبُعدی، برای سطح‌های n بُعدی. شبیه سطح‌ها، در این‌جا هم، تنها به هندسهٔ ذاتی آن‌ها نظر داریم، فضای سه‌بُعدی ریمانی، در

حوزه‌های کوچک، همان هندسهٔ اقلیدسی است، ولی در حوزه‌های بزرگ ممکن است با فضای اقلیدسی متفاوت باشد. از جمله، طول محیط دایره، ممکن است با شعاع آن متناسب نباشد؛ این محیط، تنها وقتی با تقریبی خوب، متناسب با شعاع است که با دایره‌های کوچک سر و کار داشته باشیم. مجموع زاویه‌های مثلث ممکن است برابر با دو قائمه نباشد (درضمن، در ساختمان مثلث‌ها، نقش پاره‌خط‌های راست، به‌عهدهٔ کوتاه‌ترین فاصله‌هاست، یعنی خط‌هایی که دو نقطه را با کمترین فاصله به هم وصل می‌کنند).

می‌توان فضای واقعی را تنها در حوزه‌هایی اقلیدسی دانست که در مقایسه با مقیاس‌های اخترشناسی، بزرگ نباشد. هرچه این حوزه کوچکتر باشد، سازگاری آن با هندسهٔ اقلیدسی بیشتر است، ولی می‌توان اندیشید (و درواقع هم چنین است) که، در مقیاس‌های خیلی بزرگ، با هندسه‌ای سر و کار داریم که تا اندازه‌ای با هندسهٔ اقلیدسی متفاوت است. و می‌دانیم، ***لباچوسکی*** هم، این اندیشه را مطرح کرد. ***ریمان*** این اندیشه را تکامل داد و به‌صورت عام‌تری درآورد؛ اندیشهٔ او این بود که، هر هندسه‌ای در مقیاس بی‌نهایت کوچک‌ها، اقلیدسی است، و این ویژگیِ همهٔ هندسه‌هاست، نه‌تنها هندسهٔ ***لباچوسکی*** که حالت خاصی از هندسهٔ ریمانی است.

از آن‌چه گفتیم، روشن می‌شود که هندسهٔ ریمانی، برپایهٔ ترکیب و تعمیم سه اندیشه پدید آمد، سه اندیشه‌ای که تا زمان او هم، موجب تکامل هندسه شده بود. اندیشهٔ اول دربارهٔ امکان وجود هندسه‌هایی بود که با هندسهٔ اقلیدسی متفاوت باشند؛ اندیشهٔ دوم به مفهوم مربوط به وجود هندسهٔ درونی سطح‌ها و اندیشهٔ سوم، مربوط به مفهوم فضاهایی بود که بُعد دل‌خواهی داشته باشند.

۳. برای این که با تعریف ریاضی هندسهٔ ریمانی آشنا شویم، قبل از هر چیز، قانون اندازه‌گیری فاصله‌ها را در هندسهٔ اقلیدسی به‌یاد بیاوریم.

اگر دستگاه مختصات قائم x، y را روی صفحه در نظر بگیریم، آن‌وقت با توجه به قضیهٔ ***فیثاغورس*** فاصلهٔ بین دو نقطه‌ای که مختصات آن‌ها، به‌اندازهٔ Δx و Δy با هم اختلاف دارند، با این دستور بیان می‌شود:

$$s=\sqrt{\Delta x^{۲}+\Delta y^{۲}}$$

که در فضای سه‌بعدی، به این صورت درمی‌آید:

$$s=\sqrt{\Delta x^{۲}+\Delta y^{۲}+\Delta z^{۲}}$$

در فضای اقلیدسی n بعدی هم، فاصله با این دستور بیان می‌شود:

$$s=\sqrt{\Delta x_1^2+\cdots+\Delta x_n^2}$$

از همین‌جا، می‌توان به‌سادگی، قانون اندازه‌گیری فاصله را در فضای ریمانی داد. این قانون باید، در حالتی که با حوزه‌های بی‌نهایت کوچکِ نزدیک به یک نقطه سر و کار داریم، با قانون اقلیدسیِ اندازه‌گیری فاصله، یکی شود. این موضوع، ما را به تنظیم این قانون راهنمایی می‌کند.

فضای n بعدی ریمانی دارای این ویژگی است که، در نزدیکی هر نقطهٔ A بتوان مختصات x_1، x_2، ... ، x_n را طوری وارد کرد که فاصلهٔ از نقطهٔ A تا نقطهٔ X، که بی‌نهایت نزدیک به A است، با این دستور بیان شود:

$$ds=\sqrt{dx_1^2+\cdots+dx_n^2} \qquad (۱۲)$$

که در آن dx_1، ... ، dx_n، اختلاف‌های بی‌نهایت کوچک مختصات نقطه‌های A و X است. به زبان دقیق‌تر می‌توان گفت: فاصلهٔ نقطهٔ A تا هر نقطهٔ X نزدیک به آن، با همان دستورِ هندسهٔ اقلیدسی بیان می‌شود، ولی با تقریب معینی؛ هرچه نقطهٔ X به نقطهٔ A نزدیک‌تر باشد، این تقریب دقیق است، یعنی

$$s(AX)=\sqrt{\Delta x_1^2+\cdots+\Delta x_n^2}+\varepsilon$$

که در آن، ε در مقایسه با جملهٔ اول، مقدار کوچکی است و هرچه اختلاف مختصاتی Δx_1، ... ، Δx_n کمتر باشد، ε مقداری کوچکتر می‌شود[۱].

و این، تعریف ریاضی دقیق متریک ریمانی و فضایِ ریمانی است. تفاوت متریک ریمانی با متریک اقلیدسی در این است که چنین قانونی تنها برای هر نقطهٔ مفروض نزدیک به نقطهٔ A درست است. به‌جز آن، مختصاتی را که به این سادگی بیان می‌شود، باید برای نقطه‌های مختلف، متفاوت انتخاب کرد[۲]. باز هم دربارهٔ متریک ریمانی صحبت خواهیم کرد و

۱. مفهوم دقیق دستور (۱۲) چنین است. فرض کنید یک منحنی از نقطهٔ A طوری آغاز شده باشد که مختصات نقطه‌های X آن، به‌صورت تابع‌های $x_1(t)$، $x_2(t)$، ... ، $x_n(t)$، نسبت به متغیری مثل t، داده شده باشد. در این صورت، دیفرانسیل ds، طول‌های کمان‌های این منحنی در نقطهٔ A، با دستور (۱۲) بیان می‌شود.

۲. اگر بتوان در تمامی فضا، مختصات را طوری انتخاب کرد که، برای هر دو نقطهٔ نزدیک به هم، این قانون اندازه‌گیری فاصله درست باشد، به‌معنای آن است که با فضای اقلیدسی سر و کار داریم.

اختلاف دقیق‌تر آن را با متریک اقلیدسی خواهیم آورد.

این که فضای ریمانی در بخش بی‌نهایت کوچک خود، بر فضای اقلیدسی منطبق است، به ما اجازه می‌دهد کمیت‌های هندسی اساسی مربوط به آن را، شبیه هندسهٔ ذاتی سطح، که در بخش بی‌نهایت کوچک خود به صفحه نزدیک است، تعریف کنیم (جلد دوم، بخش هفتم، بند ۴ را هم ببینید). از جمله، حجم بی‌نهایت کوچک، به‌همان صورتی که در فضای اقلیدسی وجود دارد، بیان می‌شود. حجم یک حوزهٔ محدود، از راه جمع‌بندی حوزه‌های بی‌نهایت کوچک، یعنی انتگرال‌گیریِ دیفرانسیل حجم، به‌دست می‌آید. طول منحنی، از راه جمع‌بندی فاصله‌های بی‌نهایت کوچک بین نقطه‌های بی‌نهایت نزدیک به هم در آن، یعنی انتگرال‌گیری ds در طول منحنی محاسبه می‌شود. و این، بیان تحلیلی ساده‌ای از آن است که، طول با روی هم گذاشتن مقیاس بی‌نهایت کوچک، در طول منحنی به‌دست آید. زاویهٔ بین دو منحنی، که از یک نقطه آغاز شده‌اند، درست به همان صورت فضای اقلیدسی تعریف می‌شود. در هندسهٔ n بعدی ریمانی، می‌توان سطح‌هایی را مشخص کرد که بُعد آن‌ها از ۲ تا $(n-1)$ باشد. درضمن به‌سادگی ثابت می‌شود، هر سطحی از این‌گونه، به‌نوبهٔ خود، یک فضای ریمانیِ متناظر با بُعد خود است، درست شبیه آن که، سطح در فضای اقلیدسی معمولی، یک سطح ریمانی دوبُعدی از آب درمی‌آید.

همچنین می‌توان ثابت کرد که، فضای ریمانی را همیشه، به‌عنوان سطحی در فضای اقلیدسی در نظر گرفت، ولی با بُعدی به‌اندازهٔ کافی بیشتر: از جمله، برای هر فضای ریمانی n بعدی، در فضای $\frac{n(n+1)}{2}$ بُعدی اقلیدسی، می‌توان سطح n بُعدی‌ای یافت که، از دیدگاه هندسهٔ ذاتی آن، اختلافی با این فضای ریمانی ندارد (دست‌کم، در بخش محدود مفروضی از آن).

۴. برای این که در هندسهٔ ریمانی، عبارت‌های تحلیلی مقدارهای مختلف هندسی را به‌دست آوریم، باید قبل از همه، عبارت کلی برای قانون اندازه‌گیری طول را در فضای ریمانی بدهیم، بدون این که ارتباطی به مختصات نقطه داشته باشد. دستور (۱۲) برای هر نقطهٔ A مربوط به مختصاتی است که به‌ویژه برای این نقطه انتخاب کرده‌ایم، به‌نحوی که برای عبور از یک نقطه به نقطهٔ دیگر، باید خودِ مختصات را هم تغییر داد که البته، چندان مناسب نیست. آزاد شدن از این قید دشوار نیست. به این گزاره توجه کنید.

فرض کنید، در حوزه‌ای از فضای ریمانی، مختصاتی مثل y_1، y_2، ...، y_n را وارد کرده باشیم. در این صورت «فاصلهٔ بی‌نهایت کوچک» یا آن که می‌گویند، «عنصر طول» ds از نقطهٔ

A به مختصات y_1، y_2، ... ، y_n تا نقطهٔ X به مختصات y_1+dy_1، y_2+dy_2، ... ، y_n+dy_n، با این دستور بیان می‌شود:

$$ds=\sqrt{\sum_{i,k=1}^{n} g_{ik}dy_i\,dy_k} \quad \text{یا} \quad ds^2=\sum g_{ik}dy_i\,dy_k \qquad (۱۳)$$

که در آن، ضریب‌های g_{ik}، تابع‌هایی از مختصات y_1، y_2، ... ، y_n مربوط به نقطهٔ A هستند.

عبارتی که در سمت راست دستور (۱۳) قرار دارد، نسبت به دیفرانسیل‌های مختصات y_1، y_2، ... ، y_n، صورت درجهٔ دوم (یا فرم کوادراتیک) نامیده می‌شود[1]: در واقع، این صورت درجهٔ دوم چنین است:

$$\sum g_{ik}dy_i dy_k=g_{11}dy_1^2+g_{12}dy_1dy_2+g_{21}dy_2dy_1+g_{22}dy_2^2+\cdots$$

چون $dy_1dy_2=dy_2dy_1$، پس جمله‌های دوم و سوم را می‌توان یکسان دانست: $g_{12}=g_{21}$ و به‌طور کلی $g_{ik}=g_{ki}$؛ این کار از این جهت ممکن است که تنها مجموع آن‌ها $(g_{ik}+g_{ki})dy_idy_k$ مهم است.

صورت درجه دوم (۱۳) معین مثبت است، زیرا روشن است که $ds^2>0$، به‌جز حالتی که همهٔ دیفرانسیل‌های dy_i برابر صفر باشند.

عکس این گزاره هم درست است. اگر در فضای nبعدی که با مختصات y_1، y_2، ... ، y_n داده شده است، عنصر طول با دستور (۱۳) داده شده باشد، به‌شرطی که صورت درجه دوم آن معین مثبت باشد (یعنی همیشه بزرگتر از صفر باشد، مگر وقتی که همهٔ y_i ها برابر صفر باشند) ، آن وقت با یک فضای ریمانی سروکار داریم. به‌زبان دیگر، در نزدیکی هر نقطه‌ای مثل A، می‌توان مختصات خاص جدید x_1، x_2،...، x_nرا طوری وارد کرد که، در مختصات جدید این نقطه، عنصر طول به‌صورت ساده دستور (۱۲) بیان شود:

$$ds^2=dx_1^2+dx_2^2+\cdots+dx_n^2$$

به‌این ترتیب، متریک ریمانی (یعنی تعریف طول، که در قطعهٔ بی‌نهایت کوچک، اقلیدسی است)، می‌تواند به‌وسیلهٔ هر صورت درجه دوم معین مثبت (۱۳) با

۱. صورت درجه دوم از چند کمیت، به‌عبارتی جبری گفته می‌شود که نسبت به‌این کمیت‌ها، یک چند جمله‌ای همگن از درجهٔ دوم باشد.

ضریب‌های g_{ik} (یعنی تابع‌هایی از مختصات y) داده شود، چنین است روش کلی متریک ریمانی.

منحنی در فضای ریمانی به‌این‌ترتیب داده می‌شود که هر n مختص نقطه، نسبت به‌پارامتر t، که در بازهٔ معینی تغییر می‌کند، مشخص شود:

$$y_1 = y_1(t)\ ,\ y_2 = y_2(t)\ ,\ \dots\ ,\ y_n = y_n(t)\ (a \le t \le b) \qquad (۱۴)$$

طول منحنی هم با این انتگرال بیان می‌شود:

$$s = \int ds = \int \sqrt{\sum g_{ik} dy_i dy_k}$$

در حالتی که منحنی با معادله‌های (۱۴) داده شده باشد، داریم:

$$dy_1 = y'_1\, dt\ ,\ \dots\ ,\ dy_n = y'_n\, dt$$

$$s = \int_a^b \sqrt{\sum g_{ik}\, y'_i\, y'_k}\, dt \qquad (۱۵)$$

چون g_{ik}ها، تابع‌های معلومی از مختصات y_1 ، ... ، y_n هستند، که این‌ها هم، به‌نوبهٔ خود بنابر دستور (۱۴) به t بستگی دارند، بنابراین در دستور (۱۵) در زیر علامت انتگرال، تابع معینی از t برای منحنی مفروض وجود دارد، در نتیجه: انتگرال آن، معنای مشخصی دارد و طول منحنی هم، مقدار معینی است.

از بین منحنی‌هایی که دو نقطهٔ A و B را به‌هم وصل می‌کنند، آن که کمترین طول را دارد، به‌عنوان فاصلهٔ بین این نقطه‌ها در نظر گرفته می‌شود. خود این منحنی ــ از نظر هندسی ــ نقشی شبیه پاره خط راست AB به‌عهده دارد. می‌توان ثابت کرد، در یک حوزهٔ کوچک، هر دو نقطه، تنها به‌وسیلهٔ یک خط کوتاه‌ترین به‌هم وصل می‌شود، یعنی این کوتاه‌ترین خطی که دو نقطه را به‌هم می‌پیوندد، منحصر به‌فرد است. مسألهٔ پیدا کردن خط‌های ژئودزیک (کوتاه‌ترین خط‌ها)، عبارت از مسألهٔ مربوط به‌می‌نیمم انتگرال (۱۵)، و این، مسأله‌ای مربوط به‌حساب وردش‌ها است که در بخش هشتم (جلد دوم) با آن آشنا شده‌ایم. استفاده معمولی از روش‌های حساب وردش‌ها، امکان رسیدن به‌معادله‌ای دیفرانسیلی را می‌دهد که به‌کمک آن خط‌های ژئودزیک مشخص و، درضمن، ویژگی‌های کلی آن‌ها برای هر فضای ریمانی ثابت شود.

ثابت می‌کنیم، دستور کلی (۱۳) برای متریک ریمانی، در هر دستگاه برقرار است.

فرض کنیم، مختصات y_1 ، y_2 ، ... ، y_n را در حوزه‌ای از فضای ریمانی وارد کرده باشیم. نقطهٔ دل‌خواه A را در این حوزه انتخاب می‌کنیم و x_1 ، x_2 ، ... ، x_n را مختصات خاص نقطهٔ A می‌گیریم که برای آن‌ها، عنصر طول در این نقطه با دستور (۱۲) بیان شود. دستور (۱۲) را می‌توان این طور نوشت:

$$ds^2 = dx_1^2 + dx_2^2 + \cdots + dx_n^2 \tag{16}$$

مختصات x_i برحسب مختصات y_i $(i = 1 , \dots , n)$ با دستورهایی قابل بیان‌اند:

$$\begin{aligned} x_1 &= f_1(y_1 , y_2 , \dots , y_n) \\ x_2 &= f_2(y_1 , y_2 , \dots , y_n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= f_n(y_1 , y_2 , \dots , y_n) \end{aligned}$$

یا

$$dx_1 = \frac{\partial f_1}{\partial y_1} dy_1 + \frac{\partial f_2}{\partial y_2} dy_2 + \cdots + \frac{\partial f_n}{\partial y_n} dy_n$$

و به‌همین ترتیب برای dx_2 ، ... ، dx_n. این عبارت‌ها را در دستور (۱۶) می‌گذاریم. در این صورت، اگر سمت راست برابری را در عبارت حاصل، به‌مربع تبدیل و جمله‌های به‌صورت dy_1^2 ، $dy_1 dy_2$ ، dy_2^2 و غیره را باهم در نظر بگیریم، به‌این عبارت می‌رسیم:

$$ds^2 = g_{11} dy_1^2 + 2g_{12} dy_1 dy_2 + g_{22} dy_2^2 + \cdots + g_{nn} dy_n^2$$

که در آن، ضریب‌های g_{11} ، g_{12} ، ... ، g_{nn} برحسب مشتق‌های جزیی $\frac{\partial f_i}{\partial y_j}$ بیان می‌شوند (چگونگی این عبارت‌ها برای ما مهم نیست). ولی این، همان دستور (۱۳) است که به‌صورت باز شدهٔ خود نوشته شده است. به‌این ترتیب، گزارهٔ ما ثابت شد.

اکنون ثابت می‌کنیم که، برعکس، دستور (۱۳) معرف متریک ریمانی است، یعنی با انتخاب مختصات x_i، که خاص نقطه است، می‌توان آن را به‌صورت سادهٔ (۱۶) منجر کرد. فرض کنید

$$ds^2 = \sum g_{ik} dy_i dy_k$$

در ضمن، g_{ik}ها، تابع‌هایی از مختصات y_1 ، ... ، y_n هستند و صورت درجه دوم سمت راست برابری معین مثبت است. نقطه‌ای مثل A را در نظر می‌گیریم و این دستور را، تنها برای این نقطه بررسی می‌کنیم. در این صورت، ضریب‌های g_{ik} عددهای مفروضی‌اند و متغیرهایی که عبارت ما به آن‌ها بستگی دارند، عبارت‌اند از dy_1 ، ... ، dy_n. از جبر می‌دانیم، هر صورت درجه دوم معین مثبت را (با هرگونه ضریب‌های عددی)، می‌توان به‌یاری تبدیل خطی متغیرها، به‌مجموع مربع‌ها منجر کرد[1] (بخش شانزدهم را ببینید)، یعنی تبدیلی مثل

$$dy_1 = a_{11}dx_1 + \cdots + a_{1n}dx_n \qquad (17)$$

$$dy_n = a_{1n}dx_1 + \cdots + a_{nn}dx_n$$

وجود دارد که بعد از تبدیل این عبارت‌ها، به‌صورت (۱۳)، به‌دست می‌آید:

$$ds^2 = dx_1^2 + \cdots + dx_n^2$$

اگر متغیرهای y_1 ، y_2 ، ... ، y_n را به‌یاری دستورهای

$$y_1 = a_{11}x_1 + \cdots + a_{1n}x_n$$

$$y_n = a_{n1}x_1 + \cdots + a_{nn}x_n$$

به x_1، ...، x_n تبدیل کنیم، آن وقت دیفرانسیل‌های dy_i با توجه به‌دستورهای (۱۷)، به‌دیفرانسیل‌های dx_i مربوط می‌شوند. در نتیجه، این تغییر مختصات مسألهٔ ما را حل می‌کند: با انتخاب مختصات x_1، ...، x_n برای نقطه، مربع دیفرانسیل ds^2، به‌صورت سادهٔ «اقلیدس»، و با دستور (۱۶) بیان می‌شود. به‌این‌ترتیب ثابت شده، دستور کلی (۱۳)، در واقع، متریک ریمانی را می‌دهد.

۵. فضای اقلیدسی، حالت خاص و ساده‌ای از فضای ریمانی است[2]. مسألهٔ مهم در

---

۱. این که در این جا، متغیرها دیفرانسیل‌ها هستند، برای ما مهم نیست: می‌توانیم آن‌ها را به‌طور ساده همچون متغیرهای مستقل در نظر بگیریم.

۲. در فضای اقلیدسی، عنصر طول در دستگاه مختصات قائم با دستور (۱۶)، یعنی $ds^2 = \Sigma dx_i^2$ داده می‌شود. اگر به‌دستگاه دیگری برویم، بنابر نتیجه‌ای که در شمارهٔ ۴ گرفتیم، ds^2 به‌صورت درجه دوم (۱۳) درمی‌آید. ←

هندسهٔ ریمانی، بیان تحلیلی اختلاف فضای کلی ریمانی از فضای اقلیدسی و تعیین معیاری برای فضای نااقلیدسی ریمانی است. این معیار عبارت است از به‌اصطلاح انحنای فضا (یا خمیدگی فضا).

باید براین نکته تأکید کنیم که، مفهوم انحنای فضا، به‌هیچ به‌این معنا نیست که فضای ما در فضایی از درجهٔ بالاتر و جادارتر واقع است که در درون آن، به‌نوعی کج می‌شود و انحنا پیدا می‌کند. انحنای در درون خودِ فضای مفروض تعریف می‌شود و معرف اختلاف آن با فضای اقلیدسی و به‌معنای ویژگی‌های هندسی ذاتی آن است. این را باید به‌روشنی درک کرد که مفهوم انحنای فضا به‌چیزی بیگانه و خارج از فضا مربوط نیست. وقتی می‌گوییم، فضای واقعی ما انحنا دارد، تنها به‌این معنی است که ویژگی‌های هندسی آن با ویژگی‌های فضای اقلیدسی فرق دارد؛ ولی به‌هیچ وجه به‌این معنی نیست که فضای ما در فضای بزرگتری قرار دارد که در آن جا دچار خمیدگی می‌شود. تصور یک فضا در درون فضای دیگر، هیچ ربطی به‌کاربرد فضای ریمانی در فضای واقعی ندارد و مربوط به‌زمینهٔ داستان‌های تخیلی است.

مفهوم انحنای فضای ریمانی تعمیمی از مفهوم انحنای گوسی سطح برای n بعد است. همان‌طور که در بند ۴ بخش هفتم (جلد دوم) گفتیم، انحنای گوسی معیاری برای جدا کردن هندسهٔ ذاتی سطح از هندسهٔ روی صفحه است و می‌تواند از دیدگاه خاص هندسهٔ ذاتی بررسی شود. این انحنای هندسهٔ ذاتی سطح، چیزی جز همان انحنای فضای ریمانی دو بعدی نیست که عبارت است از سطح مفروض.

برای نمونه، دو دستور هندسهٔ ذاتی را به‌یاد آوریم که در آن‌ها انحنای گوسی وارد شده است. فرض کنید، روی سطح و در نزدیکی نقطهٔ O، مثلث کوچکی داشته باشیم که ضلع‌های آن، خط‌های ژئودزیک باشند؛ زاویه‌های آن را α، β، γ و مساحت آن را σ می‌گیریم. مقدار $\alpha+\beta+\gamma-\pi$، اختلاف مجموع زاویه‌های آن را از مجموع زاویه‌های مثلث روی صفحه نشان می‌دهند.

→

بنابراین، در فضای اقلیدسی، در هر دستگاه مختصات، همین دستور کلی (۱۴) برای عنصر طول برقرار است. ولی تفاوت فضای اقلیدسی از هر فضای دیگر در این است که در آن می‌توان مختصات (یعنی مختصات قائم) را طوری انتخاب کرد که دستور (۱۶) همه جا و برای یک مختصات ثابت برقرار باشد، نه‌این که مثل حالت کلی فضای ریمانی، تنها در نزدیکی این یا آن نقطه قابل استفاده باشد.

اگر مثلث به‌سمت نقطهٔ O جمع شود، آن وقت نسبت $\alpha+\beta+\gamma-\pi$ به‌مساحت آن σ به‌سمت انحنای گوسی K در نقطهٔ O میل می‌کند. به‌زبان دیگر، برای مثلث کوچک داریم:

$$\frac{\alpha+\beta+\gamma-\pi}{\sigma}=K+\varepsilon$$

که اگر مثلث به‌سمت نقطهٔ O نزدیک شود، آن وقت $\varepsilon\to 0$. به‌این‌ترتیب، انحنای گوسی K معیاری است برای اختلافی که بین مجموع زاویه‌های مثلث روی سطح، با مجموع زاویه‌های مثلث روی صفحه، وجود دارد.

اکنون در روی سطح، دایرهٔ کوچکی به‌مرکز نقطهٔ O را در نظر می‌گیریم (یعنی مکان هندسی نقطه‌هایی که، از O به‌یک فاصله‌اند؛ فاصله را به‌مفهومی که روی سطح دارد می‌گیریم). اگر r طول شعاع دایره و l طول محیط آن باشد، آن وقت

$$l=2\pi r-\frac{\pi}{3}Kr^{3}+\varepsilon$$

که در آن، K معرف انحنای گوسی در نقطهٔ O و ε عددی کوچک در مقایسه با r^{3} است.

در این‌جا، انحنای گوسی همچون معیاری برای انحراف طول محیط دایرهٔ کوچک واقع بر سطح، از $2\pi r$ است که در هندسهٔ اقلیدسی به‌دست می‌آید.

انحنای فضای ریمانی هم، در نقش مشابهی ظاهر می‌شود. این انحنا را، برای نمونه، به‌این ترتیب می‌توان تعیین کرد. در فضای مفروض ریمانی، سطح هموار F را که از خط‌های ژئودزیک نقطهٔ مفروض O گذشته است، در نظر گرفت. انحنای گوسی این سطح را به‌عنوان انحنای فضا در نقطهٔ O و در جهت سطح F می‌پذیریم. در حالت کلی، این انحنا نه‌تنها در نقطه‌های مختلف O، بلکه برای سطح‌های مختلف ژئودزیک F هم، که از یک نقطهٔ O می‌گذرند، فرق می‌کند. بنابراین، انحنای فضا در نقطهٔ مفروض، با یک عدد مشخص نمی‌شود. خود ریمان، قانونی داد که انحناهای سطح‌های مختلف F در یک نقطه را، به‌هم مربوط می‌کند. به‌یاری این رابطه، انحنای در نقطه به‌وسیلهٔ یک دستگاه عددی مشخص می‌شود که به‌آن ***تانسور انحنا*** (یا تانسور خمیدگی) گویند.

با همهٔ این‌ها، در این جا نمی‌توانیم بیشتر مطلب را روشن کنیم، زیرا برای این منظور یاضی زیادی نیاز داریم. تنها لازم است بگوییم، انحنا معیاری برای نااقلیدسی بودن انی است؛ این انحنا در درون فضا و به‌عنوان معیار انحراف متریک آن از متریک سی، تعریف می‌شود و به‌عنوان نمونه، اختلاف مجموع زاویه‌های مثلث با 2π و

اختلاف طول محیط دایره را با $2\pi r$ معلوم می‌کند. این انحنا در حالت کلی، در نقطه‌های مختلف، مقدارهای مختلف دارد؛ در ضمن در یک نقطه هم، نه با یک عدد، بلکه با دستگاهی از عددها داده می‌شود.

فضای ریمانی ممکن است نسبت به‌ویژگی‌های خود همگن نباشد و بنابراین، حرکت آزاد شکل در آن، بدون تغییر فاصله بین نقطه‌های آن ممکن نیست. پرسش مهمی پیش می‌آید: در کدام گونه از فضاهای ریمانی، حرکت آزاد شکل، و با همان درجه آزادی که در فضای اقلیدسی وجود دارد، ممکن است؟ چنین فضاهایی، همگن‌ترین فضای ریمانی به‌شمار می‌آیند.

معلوم شده است که فضای اقلیدسی، یک فضای همگن بدون انحنا است (فضای با انحنای صفر). گونهٔ دیگری از فضای همگن، فضای *لباچوسکی* است، به‌نحوی که هندسهٔ لباچوسکی هم، همچون هندسهٔ اقلیدسی، حالت خاصی از هندسهٔ کلی ریمانی است.

به‌طور کلی، فضاهای ریمانی که در آن‌ها امکان آزاد حرکت شکل وجود داشته باشد، فضاهای با انحنای ثابت‌اند. در این فضاها، انحنا در همهٔ نقطه‌ها و برای همهٔ سطح‌های ژئودزیک، مقدار ثابتی است (به‌جای «تانسور انحنا» که از نقطه‌ای به‌نقطهٔ دیگر تغییر می‌کند، در این‌جا با یک عدد سر و کار داریم که برای همهٔ نقطه‌ها یکی است). فضای با انحنای صفر، فضای اقلیدسی و فضای با انحنای منفی، فضای *لباچوسکی* است؛ فضای با انحنای مثبت دارای همان هندسهٔ کرهٔ n بعدی در فضای $(n+1)$ بعدی اقلیدسی است.

۶. هندسهٔ ریمانی، برای کاربرد خود، خیلی انتظار نکشید. همان‌طور که پیش از این هم گفتیم، خود ریمان، از این ابزار صوری برای حل مسألهٔ مربوط به‌رسانایی گرما استفاده کرد، ولی این هنوز کاربرد دستورهای آن بوده، نه اندیشهٔ فضای انتزاعی با اندازه‌گیری اقلیدسی فاصله‌ها در حوزه‌های بی‌نهایت کوچک. این کاربرد در فضای رنگ‌ها داده شد که در آن، با استفاده از متریک ریمانی، فاصلهٔ بین رنگ‌های نزدیک، امکان توضیح پیدا می‌کند؛ فضای رنگ‌ها، می‌تواند تعبیری عینی از فضای سه بعدی ریمانی باشد.

فضای ریمانی در مکانیک هم کاربرد مهمی پیدا کرد. برای این که ماهیت این کاربرد را بفهمیم، در آغاز حرکت نقطه برسطح را بررسی می‌کنیم. یک نقطهٔ مادی، در مثل یک سنگ‌ریزه را در نظر می‌گیریم که می‌تواند آزادانه روی سطح هموار حرکت کند، ولی از این سطح خارج نمی‌شود. نقطه درست روی سطح حرکت می‌کند. می‌توان روی سطح، مختصات x_1 و x_2 را وارد کرد؛ در این صورت، حرکت نقطه در بستگی مختصات آن با

زمان، و سرعت ــ سرعت تغییر مختصات ــ با مشتق‌های آن‌ها $\dot{x}_1$, $\dot{x}_2$ نسبت به‌زمان، معین می‌شود. به‌نظر می‌رسد، نقطه در سطحی دوبعدی حرکت می‌کند. ولی این، فضای اقلیدسی نیست، بلکه هندسهٔ خودش را دارد: هندسهٔ ذاتی سطح. می‌توان قانون‌های حرکت را طوری تبدیل کرد که در آن‌ها تنها مختصات x_1 و x_2 نقطه روی سطح و مشتق‌های اول و دوم آن‌ها دخالت داشته باشد.

اگر نیرویی، برنقطه اثر بگذارد، آن وقت مؤلفه‌هایی از آن که بر سطح عمود است، پایداری خود را از دست می‌دهد و تنها مؤلفهٔ مماس بر سطح می‌ماند؛ آن هم تنها در طول سطح عمل می‌کند. بنابراین نیروهایی راکه بر نقطه عمل می‌کنند، می‌توان به‌عنوان تأثیری که بر خود سطح دارند، به‌حساب آورد. هندسهٔ ذاتی سطح، حالت خاصی از هندسهٔ ریمانی است. بنابراین، حرکت نقطه بر سطح، حرکت در فضای دو بعدی ریمانی است. قانون‌های این حرکت، دارای همان خصلت‌های قانون‌های عادی حرکت است، تنها با این تفاوت که در این جا، هندسهٔ ذاتی سطح در نظر گرفته می‌شود. این وضع به‌ویژه با به‌یاد آوردن نمونه‌ای که در ۴ بخش هفتم (جلد دوم) آمده بود، روشن می‌شود: نقطه‌ای که روی سطح با اینرسی (ماند یا لختی) ولی بدون مالش حرکت کند. روی خط ژئودزیک و با سرعت ثابت حرکت می‌کند. از آن جا که خط‌های ژئودزیک در روی سطح، نقش خط‌های راست را به‌عهده دارند، بنابراین، وضعی که پیش آمده است، شبیه قانون اینرسی است؛ این همان قانون اینرسی است، ولی برای حرکت روی سطح یا، به‌صورت انتزاعی، برای حرکت در فضای دو بعدی ریمانی.

البته، در این تصور انتزاعی، هنوز هیچ‌گونه برتری دیده نمی‌شود، زیرا هنوز صحبت برسر حرکت در سطح عادی است.

فایدهٔ این دیدگاه انتزاعی، وقتی به‌روشنی دیده می‌شود که بر دستگاهی مکانیکی بپردازیم که موقعیت آن، با بیش از دو مقدار داده شده باشد. در این صورت، نمایش حرکت‌های آن، به‌وسیلهٔ حرکت نقطه روی سطح، ناممکن می‌شود. با این وضع در بند ۷ هم برخورد کرده بودیم، وقتی که در این‌باره صحبت می‌کردیم که چگونه روش‌های نموداری در فضای انتزاعی، دربارهٔ فضای چند بعدی، خود را نشان می‌دهند.

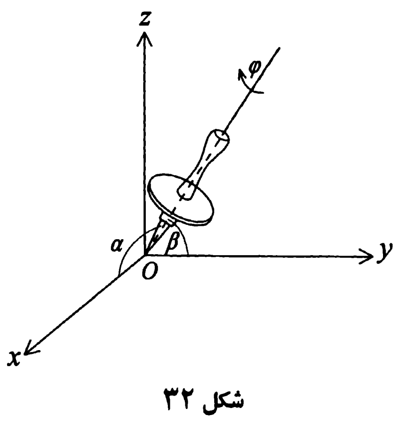
یک دستگاه مکانیکی را در نظر می‌گیریم که ترکیب آن، یعنی بخش‌های جزئی آن، با n مقدار x_1 ، x_2 ، ... ، x_n داده شده باشد. اگر صحبت بر سر دستگاهی از چند نقطهٔ مادی باشد، آن وقت جا و موقعیت آن‌ها به‌وسیلهٔ همهٔ مختصات، برای هر نقطه سه مختص، داده

می‌شود. نمونهٔ دیگر «فرفره» است (چرخی که دور محور خود می‌چرخد، در ضمن محور هم می‌تواند دور نقطهٔ ثابتی دوران کند). دوران فرفره دور محور با زاویهٔ دوران داده می‌شود، و انحراف محور با دو زاویه‌ای که نسبت به‌دو جهت مفروض ساخته است. روی هم سه مقدار وجود دارد که موقعیت فرفره را معین می‌کند (شکل ۳۲).

هر پیکربندی ــ هر موقعیت بخش‌های دستگاه ــ می‌تواند همچون یک «نقطه» از فضایی در نظر گرفته شود که شامل همهٔ پیکربندی‌های ممکن آن است. این فضا را، فضای پیکربندی گویند[1]. بُعد این فضا برابر است با تعداد مقدارهای x_1 ، x_2 ، ... ، x_n که پیکربندی و ترکیب دستگاه را معین می‌کنند. این مقدارها، مختصات نقطه را در فضای پیکربندی تشکیل می‌دهند. برای دستگاهی که شامل سه نقطهٔ مادی است، برای هر نقطه سه مختص و بنابراین روی هم، نه مختص به‌دست می‌آید. در حالت فرفره، سه مختص برای سه زاویه داریم. یعنی فضای پیکربندی فرفره، سه بعدی است.

حرکت دستگاه، به‌معنای حرکت نقطه در فضای پیکربندی است. سرعت این حرکت از روی سرعت‌های تغییر مختصات x_1 ، x_2 ، ... ، x_n معین می‌شود.

دربارهٔ این فضاها، در رابطه با ساختار توپولوژیک آن‌ها، در بخش هجدهم هم صحبت خواهیم کرد. در این جا براین مطلب تأکید می‌کنیم که در فضای پیکربندی، می‌توان قانون خاصی را وارد کرد که، بنابر آن، اندازه‌گیری فاصله‌ها، به‌ویژگی‌های مکانیکی دستگاه



شکل ۳۲

۱. نباید این فضا را با «فضای فازی» که در شمارهٔ ۲ در بند ۸ دیدیم، اشتباه کرد. در فضای فازی، نقطه نه تنها به‌موقعیت، بلکه به‌سرعت حرکت نقطه‌های دستگاه؛ در هر لحظه بستگی دارد.

بستگی داشته باشد. برای نمونه، اگر انرژی جنبشی دستگاه با دستور

$$T=\frac{1}{2}\sum_{i,k=1}^{n} a_{ik}\dot{x}_i\dot{x}_k$$

بیان شود، که در آن، $\dot{x}_i$ معرف سرعت تغییر مختص مربوط است، آن وقت مربع فاصله بی‌نهایت کوچک با دستور

$$ds^2=\sum_{i,k=1}^{n} a_{ik}dx_i dx_k$$

داده می‌شود.

بنابراین، فضای پیکربندی، یک فضای ریمانی است، در ضمن، حرکت دستگاه نه‌تنها حرکت نقطه در فضای پیکربندی است، بلکه خود معادله‌هایی که حرکت دستگاه را بیان می‌کنند، همان معادله‌های حرکت این نقطه‌اند؛ به‌زبان دیگر، مکانیک دستگاه همچون مکانیک نقطه در فضای پیکربندی نمایان می‌شود. در حالت خاص، وقتی حرکت دستگاه زیر تأثیر اینرسی، یعنی بدون تأثیر نیرو باشد (شبیه دوران آزاد فرفره)، آن وقت حرکت یکنواخت نقطه روی خط ژئودزیک این فضا به‌دست می‌آید.

چنین تصویری از حرکت از این جهت سودمند است که، همراه با برخی تعمیم‌ها و تغییر شکل‌های آن، در مکانیک نظری کاربرد دارد.

به‌این ترتیب، هندسهٔ ریمانی، به‌عنوان یک روش انتزاعی هندسی، کاربرد خود را در پدیده‌های فیزیکی پیدا می‌کند. این هندسهٔ انتزاعی به‌هیچ‌وجه دل‌خواه یا نوعی بازی فکری ریاضی نیست، بلکه قانون‌مندی‌های واقعی را دربارهٔ پدیده‌هایی که بررسی می‌کنیم، منعکس می‌کند؛ تنها این بازتاب قانون‌مندی‌های پدیده‌های واقعی، در این جا، شکلی انتزاعی دارند. راستش این است که هر تفسیر ریاضی از پدیده‌های فیزیکی، چنین طبیعتی دارد و طبیعت هر کاربردی از هندسهٔ انتزاعی (هر جا که باشد) به همین گونه است. درست است که در جایی مثل هندسهٔ ریمانی، از انتزاع‌های نیرومندتر وظریف‌تری استفاده شده است، ولی جوهر و ماهیت اصلی کار، بی‌تغییر می‌ماند.

هندسهٔ ریمانی، درخشان‌ترین کاربرد خود را در نظریهٔ نسبیت پیدا کرد. در این باره در بند بعد به‌تفصیل صحبت خواهیم کرد، زیرا در این جا، مسألهٔ مهم و در عین حال دشوارِ بستگی هندسهٔ انتزاعی با ویژگی‌های فضای واقعی مطرح است.

۷. در نیمهٔ سدهٔ بیستم، هندسه‌های فضاهای مختلف نااقلیدسی، پیشرفت زیادی کرده و در جهت مختلف تعمیم یافته‌اند. هندسه‌هایی پدید آمده‌اند که هندسهٔ ریمانی، به‌صورت حالت خاصی از آن‌ها درمی‌آید. نخستین این هندسه‌ها، هندسهٔ فینسلری است که اندیشهٔ اصلی آن را ریمان هم طرح کرده بود[۱]؛ سپس می‌توان از *ا.کارتان*، هندسه‌دان بزرگ فرانسوی نام برد که توانست هندسهٔ ریمانی را با برنامه ارلانگنِ *کلاین* و نظریه‌های دیگر تطبیق دهد. در این جا نمی‌توانیم دربارهٔ سمت‌گیری‌های این هندسه‌ها صحبت کنیم، تنها یادآوری می‌کنیم، این هندسه‌ها به‌گونه‌ای طرح شده‌اند که برای آن‌ها بتوان از ابزار تحلیلی استفاده کرد. برای تکامل هندسه، در سمت‌های تازه، گروهی از هندسه‌دانان شوروی هم نقش داشته‌اند. از جمله می‌توان از هندسه «چند متری» *پ.ک.راشوسکی*، و بررسی‌های *و.و.واگنر* (که در زمینه‌هایی از طرح‌کلی‌ترین نظریهٔ فضاهای خمیده تا کاربرد هندسهٔ نااقلیدسی در مکانیک به‌صورت گسترده‌ای کار کرده است) و دیگران نام برد.

۱۰. هندسهٔ انتزاعی و فضای واقعی

۱. تا این جا تکامل اندیشه‌های هندسی را با آغاز از هندسهٔ *لباچوسکی* دنبال کردیم، در فضاهای انتزاعی فرو رفتیم و از موضوع نخستین هندسه، یعنی فضای واقعی که همهٔ پدیده‌ها و روند در آن شکل گرفته‌اند، بسیار دور شدیم. اکنون، دوباره به‌این فضا، به‌مفهوم عادی آن، برمی‌گردیم و تلاش می‌کنیم به‌این پرسش پاسخ دهیم که: پیشرفت هندسهٔ انتزاعی، چه کاری برای درک ویژگی‌های فضای واقعی ما انجام داده است؟

می‌دانیم، هندسه از تجربه و از بررسی تجربی شکل‌های فضایی و بستگی‌های جسم‌ها با یکدیگر، برخاسته است: از اندازه‌گیری قطعه زمین‌ها، حجم ظرف‌ها و غیره. بنابراین، سرچشمهٔ پیدایی هندسه، از جمله شبیه مکانیک، چیزی جز نظریهٔ فیزیکی نیست. اصل موضوع‌های هندسهٔ اقلیدسی، نتیجه است از تجربهٔ طولانی بشر که به‌صورتی روشن و دقیق تنظیم شده‌اند؛ این اصل موضوع‌ها، معرف قانون‌های حاکم بر طبیعت‌اند و به‌عنوان قانون‌های حاکم بر هندسه پذیرفته شده‌اند، همان‌طور که قانون‌های بنیانی مکانیک را هم

۱. فینسلر، ریاضی‌دان آلمانی، کار مفصل خود را دربارهٔ این هندسه، از سال ۱۹۱۶ آغاز کرد.

اغلب، اصل موضوع‌های مکانیک می‌نامند[1]. ولی نباید گمان برد که، این اصل موضوع‌ها، دقت مطلق دارند و در رابطه با تجربه‌های تازه، نمی‌توانند دقیق‌تر و عام‌تر شوند؛ ویژگی‌های فضا، ممکن است در جاهایی و تا اندازه‌ای، با آن چه از هندسهٔ اقلیدسی به‌دست می‌آید، متفاوت باشد.

این بحث را پیش از این هم داشته‌ایم و، به‌احتمالی، به‌اندازهٔ کافی روشن شده باشد. ولی بیش از صدوپنجاه سال از آن زمان که اندیشهٔ *لباچوسکی* به‌طور کامل به‌رسمیت شناخته شد، نمی‌گذرد. تا پیش از *لباچوسکی و گوس*، به‌ذهن کسی نمی‌رسید که ممکن است هندسه اقلیدسی آن قدر دقیق نباشد که بتواند ویژگی‌های فضا را با دقت معین کند. *لباچوسکی* هندسهٔ خود را همچون نظریه‌ای دنبال کرد که می‌تواند ویژگی‌های فضای واقعی را شرح دهد. سپس *ریمان* و دیگران، پرسش مربوط به‌ویژگی‌های فضا، دربارهٔ امکان‌هایی که برای اندازه‌گیری طول وجود دارد و کشف امکان‌هایی برای دقیق‌تر کردن مفهوم اندازه‌گیری را در برابر خود گذاشتند. به‌طور کلی هندسهٔ انتزاعی در برخی بخش‌های خود، توانست و می‌تواند، همچون نظریهٔ ویژگی‌های ممکن فضا، در نظر گرفته شود. با وجود این، همهٔ این‌ها به‌صورت فرضیه‌ای باقی بوده، تا این که در سال ۱۹۱۵، *اینشتین* در نظریهٔ نسبیت عمومی خود، اندیشهٔ *لباچوسکی و ریمان* را تأیید کرد. نظریهٔ نسبیت عمومی تأیید کرد که در واقع، هندسهٔ فضای واقعی ما، تا اندازه‌ای با فضای اقلیدسی فرق دارد. این اختلاف، همان‌طور که *لباچوسکی* انتظار داشت، در مقیاس‌های نجومی روشن شد.

آن‌چه دربارهٔ فضا در این جا گفتیم، دست کم سه دشواری پدید می‌آورد. این دشواری‌ها، سرآخر به‌این مسأله منجر می‌شوند که، چه رابطه‌ای بین هندسهٔ انتزاعی و هندسهٔ فیزیکی، یعنی ویژگی‌های فضای واقعی، وجود دارد؟

نخستین دشواری مربوط به‌این است که، در حد امکان روشن کنیم چه ویژگی‌هایی از فضای واقعی و به چه مفهومی، ممکن است از آن چه هندسهٔ اقلیدسی می‌دهد، متفاوت

---

۱. درک انتزاعی اصل موضوع‌ها، یعنی جدا شدن اصل موضوع‌ها از سرچشمهٔ مضمونی خود، در بیش از پانصد سال پیش به‌وجود آمد، ولی این وضع، هیچ تغییری در این موضوع نداد که اصل موضوع‌های هندسهٔ اقلیدسی بیان کنندهٔ قانون‌های طبیعت‌اند. وقتی دربارهٔ اصل موضوع‌ها به‌طور کلی، و نه اصل موضوع‌های هندسه یا مکانیک، صحبت می‌کنیم، می‌خواهیم به‌طرح نخستین ساختمان قیاسی منطقی این دانش‌ها برسیم، ولی این، به‌معنای آن نیست که این دانش‌ها مبنای تجربی خود را از دست داده‌اند. وقتی می‌گوییم، دانشی، اصل موضوعی شده است، به‌این معناست که بر پایه‌ای از یک ساختمان قیاسی قرار دارد و می‌توان همهٔ نظریه‌های (قضیه‌های) دیگر را از این بنیان‌ها (اصل موضوع‌ها)، به‌یاری استدلال منطقی نتیجه گرفت.

باشد؟ چنان به‌هندسهٔ اقلیدسی عادت کرده‌ایم که تصور چیزی متفاوت با آن برایمان دشوار است و نیاز به‌روشن کردن دارد.

دشواری دوم به‌بیان«ویژگی‌های فضای واقعی» مربوط می‌شود. فضا به‌خودی‌خود، تهی و همگن به‌تصور می‌آید. به‌نظر می‌رسد، در خود مفهوم فضا، تصور مربوط به«همگنی آن وجود دارد. و فضای تهی یعنی چه؟» «تهی بودن» چه ویژگی‌هایی دارد؟ وقتی دربارهٔ «ویژگی‌های فضا» صحبت می‌کنیم به‌این پرسش‌ها نمی‌اندیشیم، ولی اندیشیدن دربارهٔ آن‌ها ضرورت دارد تا دشواری موجود به‌اندازهٔ کافی محسوس شود.

دشواری سوم به‌مفهوم راستی و حقیقت این یا آن هندسه مربوط می‌شود. مطلب در آغاز ساده به‌نظر می‌رسد: هندسه‌ای حقیقت دارد که متناظر با واقعیت باشد. البته این، درست است. ولی از سوی دیگر، به‌عنوان نمونه دیدیم، هندسهٔ درونی دایره می‌تواند همچون هندسهٔ *لباچوسکی* در نظر گرفته شود، زیرا هر حقیقت هندسی مربوط به‌درون دایره را می‌توان به‌عنوان قضیه‌ای از هندسهٔ *لباچوسکی* مطرح کرد. بنابراین، به‌نظر می‌رسد. هر حقیقت هندسی را می‌توان هم به‌عنوان قضیه‌ای از هندسهٔ اقلیدسی و هم به‌عنوان قضیه‌ای از هندسهٔ *لباچوسکی* شرح داد. یعنی، هر دو هندسه، متناظر با واقعیت‌اند. به‌این‌ترتیب، کدام‌یک از آن‌ها حقیقت است و به‌چه مفهومی؟ چرا با وجود این گمان می‌کنیم در دایره، هندسهٔ اقلیدسی برقرار است و هندسهٔ *لباچوسکی* تنها به‌کمک آن تفسیر می‌شود؟

روشن است، در این پرسش‌ها، دشواری خاصی وجود دارد که بسیاری از بزرگترین ریاضی‌دانان زمان ما هم، خود را درگیر آن کرده‌اند.

بررسی را باید از دشواری دوم آغاز کرد، زیرا بعد از درک «ویژگی‌های فضا»، می‌توان حس دشواری‌های دیگر را دنبال کرد.

۲. موضوع هندسه ــ «ویژگی‌های فضا» ــ عبارت است از ویژگی‌های جسم‌های واقعی، رابطه‌های مادی شکل‌های آن‌ها. در فضای واقعی، «مکان»، «نقطه»، «جهت» و غیر آن، به‌وسیلهٔ جسم‌های مادی معین می‌شود. مفهوم‌هایی مثل «این‌جا»، «آن‌جا»، «این‌سو»، «آن‌سو» و غیره، تنها در رابطه با این با آن موضوع مادی معنا پیدا می‌کنند. «این جا» می‌تواند به‌معنای «روی زمین»، «در این اتاق»، یا چیزی از این گونه باشد؛ «این‌جا» همیشه به‌معنای مکان است که با این یا آن اثر و نشانهٔ مادی معین می‌شود. درست به‌همین‌ترتیب، از جمله، خط راست به‌خودی خود وجود ندارد و تنها به‌عنوان نخ کشیده، یا لبهٔ خط‌کش، یا پرتوی از

نور فهمیده می‌شود. «خط راست» مفهومی انتزاعی است که ویژگی‌های مشترک این پدیده‌های مادی را بازتاب می‌دهد، درست همان‌گونه «خانه به‌معنای کلی خود» مفهومی انتزاعی است که ویژگی‌های مشترک خانه را منعکس می‌کند، «خانه»، بدون ارتباط با خانه‌های جداگانه و واقعی موجود، مفهومی ندارد.

این خصلت عینیِ ویژگی‌های فضا، موقعیت ماتریالیسم دیالکتیک را نشان می‌دهد: ***فضا، صورتی از وجود ماده است.*** شکل و صورت هر چیزی، در بستگی‌ها و رابطه‌های بخش‌های آن، معین می‌شود. ساختار فضا، عبارت است از قانون‌مندی کلی و مشترک یک رشته رابطه‌های مربوط به‌جسم‌ها و پدیده‌های مادی. این، یعنی رابطه‌های فضایی، ردیف فضایی چیزها، موقعیتی که نسبت به‌هم دارند، فاصلهٔ بین آن‌ها و غیره. ولی همان‌طور که هر شکلی از مضمون خود جدا نیست، همان‌طور که هر شکلی، ولو به‌صورت انتزاعی، درچارچوب معینی به‌مضمون خود وابسته است، همان‌طور هم فضا، را نمی‌توان جدا از ماده دانست. تصور فضا «به‌خودی خود» و فضای بدون ماده، تصوری انتزاعی است که نمی‌توان از آن سوء استفاده کرد. رابطه‌ها و شکل‌های فضایی، مثل «این جا»، «بین»، «در درون»، «خط راست»، «کره» و غیره، همه جا رابطه‌ها و شکل‌های جسم‌های مادی را بازتاب می‌دهند؛ و هندسه، این رابطه‌ها و شکل‌ها را به‌صورت انتزاعی خود بررسی می‌کند. این انتزاع از حالت‌های مشخص لازم است، زیرا در غیر این صورت، نمی‌توان وجه مشترک رابطه‌های مشخص مختلف چیزها، پدیده‌ها و روندها را بازشناخت. ولی این انتزاع را نباید مطلق کرد و مفهوم‌های انتزاعی را به‌جای واقعیت‌های عینی نشاند.

در فضایی که به‌طور مطلق تهی و محروم از هرگونه ردپای ماده باشد، نه مکان را می‌توان تشخیص داد و نه جهت را و، بنابراین، در آن‌جا، مکان و جهتی وجود ندارد. به‌نحوی که فضای تهی (به‌طور مطلق تهی) به«هیچ» تبدیل می‌شود. حتی، ضمن تصور انتزاعی فضای تهی، آوای خاموشی وجود دارد که، در این فضا، می‌توان مکان‌ها و جهت‌های مختلف را تشخیص داد. به‌زبان دیگر، در فضای انتزاعی هم، ویژگی‌های مربوط به‌مکان، جهت و فاصله، به‌همان مفهومی که در فضای واقعی وجود دارد، حفظ می‌شود و این به‌آن جهت است که، این فضای انتزاعی، بستگی تنگاتنگی با جسم‌های مادی دارد.

به‌این‌ترتیب، فضا شکلی از وجود ماده است؛ و بنابراین، «ویژگی‌های فضا»، عبارت است از ویژگی‌های ماده، ویژگی‌های رابطه‌های مشخص جسم‌های مادی، وضع استقرار متقابل آن‌ها، اندازه‌های آن‌ها و غیره.

سپس، برای این‌که قضیه‌های هندسی مفهوم فیزیکی داشته باشند، باید دانست که در این قضیه‌ها، «خط راست»، «فاصله» و دیگر مفهوم‌های هندسی به‌چه معنایی گرفته شده‌اند. در بند ۴ داریم، یک نظریهٔ هندسی، می‌تواند تعبیرهای متفاوتی داشته باشد.

بنابراین، ضمن مقایسهٔ هندسه با تجربه، باید تا جایی که ممکن است معنای هندسی مفهوم‌های فیزیکی را دریافت، زیرا هندسه، تنها وقتی می‌تواند ویژگی‌های فضای واقعی را شرح دهد که مفهوم‌های آن متناظر با مفهوم‌های فیزیکی باشند. بدون این مفهوم فیزیکی، قضیه‌های هندسه، خصلتی انتزاعی، ریاضی و صوری پیدا می‌کنند. برای پیدا کردن راه‌حلِ دشواری دوم هم، که در این جا به‌آن پرداختیم، باید به‌همین صورت عمل کرد. این دشواری از آن‌جا پدید می‌آید که، به‌جای فضای واقعی، که پیوسته به ماده است، تنها به فضای «خالص» بیندیشیم، یعنی فضا به‌معنای «چیزی» که جز در حالت انتزاعی خود وجود ندارد.

۳. اکنون به‌سادگی می‌توانیم راه حل دو دشواری دیگر را پیدا کنیم.

اول، چگونه می‌توان تصور کرد، فضای واقعی در ویژگی‌های خود، با فضای اقلیدسی فرق دارد؟ فرض کنید بخواهیم گزاره‌ای از هندسهٔ اقلیدسی را آزمایش کنیم، مثل این قضیه که مجموع زاویه‌های مثلث برابر ۱۸۰ درجه، یا طول محیط دایره برابر $۲\pi R$ است. برای آزمایش حکم اول، باید ببینیم، چگونه مثلث‌های فیزیکی معینی را در نظر بگیریم و زاویه‌های آن‌ها را چگونه معین کنیم. ضلع‌های مثلث را پرتوهای نور در خلاء می‌گیریم. در چنین حالتی، هیچ تعجبی ندارد، اگر ضمن آزمایش دقیق، معلوم شود مجموع زاویه‌های مثلث با ۱۸۰ درجه فرق دارد. به‌همین ترتیب، می‌توان اندیشید، اگر شعاع و محیط دایره‌ای را با یک مقیاس اندازه بگیریم، منجر به‌نتیجه‌ای شود که به طور کامل در دستور $l=۲\pi R$ صدق نکند. این، همان وضعی است که از جمله، در روی سطح زمین پیش می‌آید. در روی سطح زمین، طول محیط دایره متناسب با طول شعاع نیست؛ این نسبت (یعنی نسبت محیط دایره به شعاع آن) به‌کندی رشد می‌کند و وقتی به حداکثر خود می‌رسد که شعاع برابر نیمی از نصف‌النهار باشد. می‌توان به این استدلال اعتراض کرد، زیرا روی سطح زمین، نقش خط‌های راست به عهدهٔ کمان‌هایی از دایره‌های عظیمه است، به نحوی که در این جا، شعاع مفهوم دیگری پیدا می‌کند، بنابراین، نتیجه‌ای که به دست آوردیم، هندسهٔ اقلیدسی را نقض نمی‌کند با وجود این، بنابر نظریهٔ نسبیت، در نزدیکی جسمی که جرم زیادی داشته باشد، نسبت طول محیط دایره به شعاع «واقعی» آن، اندکی با ۲π فرق دارد و دستور تقریبی زیر،

برای نسبت طول استوای جسم کروی همگن به طول شعاع آن برقرار است:

$$\frac{\text{طول محیط دایره}}{\text{طول شعاع}} = ۲\pi(۱ - \frac{kM}{۳Rc^{۲}})$$

که در آن M جرم جسم (به تُن)، R شعاع جسم (به کیلومتر) و c، برابر ۳۰۰۰۰۰ کیلومتر بر ثانیه (سرعت نور) و k ثابت جاذبه است که با واحدهایی که انتخاب کرده‌ایم، برابر $۶٫۶۶ \times ۱۰^{-۱۸}$ می‌شود.

می‌بینیم، نسبت طول محیط دایره به طول شعاع آن برابر ۲π نیست و اندکی کمتر از آن است. با محاسبه ثابت می‌شود که این نسبت در سطح خورشید، به تقریب، به اندازهٔ ۰٫۰۰۰۰۰۴ و در سطح قمر شعرای یمانی که میانگین چگالی آن ۵۰۰۰۰ برابر چگالی آب است، به اندازه ۰٫۰۰۰۱۴ با ۲π اختلاف دارد.

می‌توان اعتراض کرد که، با همهٔ این‌ها، در تصورهای علنی، فضای ما همیشه اقلیدسی است. این گونه اعتراض‌ها نباید ما را پریشان کند؛ اول به این دلیل که مسألهٔ دانش این نیست که تنها تصوری عینی از پدیده‌ها به ما بدهد، بلکه در این است که موفق به درک این پدیده‌ها شویم. تصور عینی محدود است و در دایرهٔ نمونه‌های عادی که در دسترس اندام‌های حسی ماست، می‌چرخد. ما در موقعیتی نیستیم که به طور عینی بتوانیم به پرتوهای فرابنفش، یا به پراکندگی موج‌های رادیویی، یا به حرکت الکترون در درون اتم و بسیاری پدیده‌های دیگر پی‌ببریم و در نتیجه، همهٔ این‌ها از میدان آگاهی‌های ما بیرون می‌رود. ولی نقص اندام‌های حسی ما، به این معنا نیست که، این پدیده‌ها، برای ما قابل فهم نیستند. برعکس، برای نمونه، پیشرفت‌های دانش رادیو به روشنی نشان می‌دهد که به موج‌های رادیویی دست یافته‌ایم و بنابراین آن‌ها را به خوبی می‌شناسیم. مهم این که در واقع، چه چیزهایی را می‌توان تصور کرد و آیا آن چه قابل تصور نیست مربوط به عادت‌ها و ورزیدگی‌های ماست؟ تنها یک نمونه کافی است. آیا می‌توان «آنتی‌پودها» را تصور کرد دربارهٔ آن‌ها ذهن خود را آزار داد؟ تنها دانش می‌تواند در این باره پاسخ دهد و حقیقت را آشکار کند.

۴. اکنون به آخرین دشواری می‌پردازیم: کدام هندسه را باید درست و واقعی دانست؟ ضمن طرح این پرسش یادآوری کردیم که، حقیقت‌های هندسی درون دایره می‌تواند هم به عنوان قضیه‌هایی از هندسه اقلیدسی و هم به عنوان قضیه‌هایی از هندسهٔ ***لباچوسکی*** در نظر گرفته شود. بنابراین، هردوی این هندسه‌ها، پاسخ گوی واقعیت‌اند. یعنی هر دوی

آن‌هادرست‌اند. و در این موضوع، سرآخر، هیچ‌گونه شگفتی وجود ندارد. یک پدیده را همیشه می‌توان با روش‌های مختلف شرح داد. یک مقدار می‌توان با واحدهای مختلفی اندازه گرفت. یک منحنی را می‌توان به‌وسیلهٔ معادله‌های مختلف، بسته به انتخاب دستگاه محورهای مختصات، معرفی کرد. درست به همین ترتیب، مجموعهٔ مفروض و برگزیده‌ای از بستگی‌های هندسی را (مثل بستگی‌هایی که در درون دایره وجود دارد) می‌توان با روش‌های متفاوت شرح داد. ولی در این جا، مسأله را نه دربارهٔ مجموعهٔ جدا شده‌ای از حقیقت‌های هندسی، بلکه دربارهٔ بستگی‌های فضایی در تمامی کلیت آن، طرح می‌کنیم. فضا عبارت است از شکل همه جانبهٔ ماده و، بنابراین، ضمن طرح مسأله دربارهٔ ویژگی‌های فضا، هیچ بخشی از حقیقت‌های آن را، نمی‌توان به طور مصنوعی جدا کرد.

وقتی مسأله را به این صورت طرح می‌کنیم، کمیت‌ها و حقیقت‌های هندسی را نمی‌توان، جدا از پدیده‌های دیگری که به ناچار با آن‌ها بستگی دارند، جدا کرد. از جمله، طول پاره خط راست از راه کنار هم گذاشتن مقیاسی که سخت و صُلب فرض شده است، به دست می‌آید، به نحوی که اندازه‌گیری طول، به ناچار، با حرکت جسم صُلب، یعنی به مکانیک بستگی دارد. هندسه از مکانیک جدایی‌ناپذیر است. در این میان، اندازه‌گیری طول در درون دایره، ضمن تعبیر هندسهٔ *لباچوسکی*، منجر به روش به‌کلی دیگری می‌شود (همان طور که در بند ۴ دیدیم)؛ در ضمن، طول وتر برابرِ بی‌نهایت در می‌آید. روشن است، چنین تعریفی از اندازه‌گیری، متناظر با طرح نخستینی که برای اندازه‌گیری داشتیم و بر پایهٔ جابه‌جایی مکانیکی جسم‌های واقعی بوده نیست. به طور کلی، شکل‌هایی به مفهوم هندسه اقلیدسی با هم برابرند، که با جابه‌جایی یکی و قرار دادن بر دیگری و از راه حرکت مکانیکی برهم منطبق شوند. در تعبیر هندسهٔ *لباچوسکی*، مفهوم برابری به صورت دیگری تعریف می‌شود. در آن جا، نقش حرکت، به عهده تبدیل دیگری است. بنابراین، اگر حقیقت‌های هندسی درون دایره، در رابطه با مکانیک در نظر بگیریم، باید بپذیریم که در واقع، در درون دایره، هندسهٔ اقلیدسی (با دقت بالایی) بر قرار است.

هندسهٔ اقلیدسی عبارت است از هندسه‌ای که در آن، نقش حرکت به عهده حرکت مکانیکی عادی جسم‌های صُلب است. درست به همین دلیل است که انسان، در آغاز، هندسهٔ اقلیدسی را کشف کرده و نه هندسهٔ دیگری، ولی پیشرفت فیزیک به این نتیجه رسیده است که قانون‌های مکانیک نیوتونی و، همراه با آن قانون‌های هندسهٔ اقلیدسی، نسبت به قانون‌ها دقیق‌تر و کلی‌تر، تقریبی‌اند. در این دگرگونی قانون‌های هندسه، ضمن

عبور از هندسهٔ اقلیدسی به هندسهٔ ریمانی (که در نظریهٔ نسبیت تحقق یافت)، تنها مکانیک دخالت نداشت: نظریهٔ پدیده‌های الکترومغناطیسی و اُپتیک هم، اگر نگوییم بیشتر دست کم به همان اندازه نقش داشتند. هندسه، به عنوان دانش ویژگی‌های فضا، به فیزیک بستگی دارد و تنها به صورتی انتزاعی و در چارچوب معینی می‌تواند جدا از آن بررسی شود.

بستگی هندسه را با فیزیک، بستگی ویژگی‌های فضا را با ماده، ***لباچوسکی*** هم به روشنی یادآوری می‌کند. او تغییر قانون‌های هندسه را به دلیل پیدایی زمینه‌های تازه‌ای در بررسی پدیده‌های فیزیکی، پیش‌بینی کرده است. بر خلاف دیدگاه ***لباچوسکی***، ***پوانکاره*** ریاضی‌دان مشهور فرانسوی تاکید می‌کند، گویا انتخاب این یا آن هندسه تنها نتیجه‌ای از تلاش ذهنی یا «قانون‌های تفکر» در حال و هوای ذهن‌گرایی ***ماخ*** است. ***پوانکاره***، براساس این دیدگاه حتی پیش بینی می‌کند که، دانش می‌تواند از قانون انتشار نور روی خط راست صرف نظر کند، ولی هرگز هندسهٔ اقلیدسی را کنار نمی‌گذارد، زیرا هندسهٔ اقلیدسی «ساده‌ترین» هندسه‌هاست. ولی تنها سه سال بعد از مرگ پوانکاره، نظریهٔ عمومی نسبیت پایه‌گذاری شد که درست عکس دیدگاه پوانکاره را ثابت کرد. معلوم شد که از هندسهٔ اقلیدسی می‌توان صرف نظر کرد، ولی قانون انتشار نور، البته به صورت تعمیم‌یافتهٔ، به قوت خود باقی است: نور در مسیرهای منحنی‌های ژئودزیک منتشر می‌شود. و این، نخستین پیش‌بینی بزرگ مکانیکی نبود. در آستانهٔ سدهٔ بیستم، ذهن‌گرایان و کسانی که از منطق مکانیکی استفاده می‌کردند، با استدلالی شبیه ***پوانکاره*** می‌گفتند، گویا اتم‌ها تنها نهادهایی زاییدهٔ اندیشه یا چیزی از این گونه‌اند و واقعیت عینی ندارند. ولی هنوز چند سالی نگذشته بود که واقعیت اتم‌ها، به صورتی قانع کننده ثابت شد. با وجود این، درک ذهن‌گرایانهٔ هندسه، همچون دیگر دانش‌ها، هنوز در میان برخی دانشمندان معتقد به نظام سرمایه‌داری وجود دارد. این‌ها کسانی‌اند که نظام سرمایه‌داری را عینی‌ترین حقیقت می‌پندارند.

به این ترتیب، می‌توانیم به این نتیجه برسیم. مجموعه‌ای از حقیقت‌های جداگانه را می‌توان، در حالت کلی، به‌گونه‌های مختلف شرح داد؛ همهٔ این شرح‌ها درست‌اند، زیرا یک واقعیت را منعکس می‌کنند. با وجود این، حقیقت‌های هندسی را در مجموع خود، نمی‌توان جدا از دیگر پدیده‌ها بررسی کرد. تنها با این روش است که می‌توان به ویژگی‌های فضا پی‌برد، زیرا فضا، شکل همه جانبه‌ای از وجود ماده است. وقتی هندسه را در رابطه با فیزیک در نظر می‌گیریم. باید به سازگاری آن‌ها، توجه داشته باشیم؛ در این صورت است که اختلاف بین «هندسه‌های» متفاوت روشن می‌شود؛ اختلاف این هندسه‌ها، اگر جدا از

فیزیک بررسی شود، تنها در این است که یکی ساده‌تر و دیگری دشوارتر است. هندسهٔ اقلیدسی به این دلیل پدید نیامد که از دیگر هندسه‌ها ساده‌تر است، بلکه به دلیل سازگاری آن با مکانیک بود. واکنون، در رابطه با پیشرفت فیزیک در نظریهٔ نسبیت، هندسه به صورتی پیچیده‌تر، یعنی هندسهٔ ریمانی، در آمده است.

سخن کوتاه، در رابطه با ویژگی‌های فضای واقعی، حق با هندسه‌ای است که بتواند ویژگی‌های رابطه‌های فضایی را، به مفهوم عام آن، به اندازهٔ کافی دقیق بازتاب دهد و نه تنها پاسخ‌گوی حقیقت‌های خالص هندسی باشد. بلکه در ضمن پاسخ‌گویی مکانیک و به طور کلی تمامی فیزیک باشد.

۵. مختصری که تا این جا دربارهٔ نظریهٔ نسبیت گفتیم، برای مضمون آن، اساسی نبود. نظریه نسبیت، در زمینهٔ درک فضا، خیلی جلوتر از آن رفت که ***لباچوسکی*** می‌اندیشید.

اساسی‌ترین جنبهٔ نظریه نسبیت که فضا را، در رابطه‌ای جدانشدنی از زمان در نظر می‌گیرد، به نحوی که فضا و زمان، با هم، شکل یگانه‌ای از وجود ماده را تشکیل می‌دهند: صورت‌های گوناگون «فضا ـ زمان» چهار بعدی پیش‌آمدها در جهان چهار بعدی «مکان ـ زمان» رخ می‌دهند و بنابراین دارای چهار مختص هستند: سه مختص مربوط به فضا و مختص چهارم مربوط به زمان رخ داد پیش آمد. پیش‌آمدها، به این مفهوم، مجموعه‌ای چهار بعدی را تشکیل می‌دهند. و این همان مجموعهٔ چهار بعدی (از دیدگاه ساختاری آن) است که نظریهٔ نسبیت، جد از ویژگی‌های پدیده‌های جداگانه، در میدان بررسی خود قرار داده است. این، به اساس و ماهیت نظریه کار ندارد، نه توجه به نظریهٔ حرکت‌های سریع دارد، نه به گیتی‌شناسی، نه به نظریهٔ تازه فضا یا زمان، این، به‌ویژه نظریهٔ «فضا ـ زمان» به عنوان شکل یگانهٔ وجود ماده است.

البته، در مکانیک نیوتونی هم می‌توان فضا و زمان را با هم در نظر گرفت و در یک دستگاه چهار بعدی قرار داد. پیش از این هم گفته‌ایم، اندیشهٔ فضای چند بعدی، برای نخستین بار به وسیلهٔ ***لاگرانژ*** و به این دلیل پدید آمد که، در رابطه با حرکت نقطهٔ مادی، مختصاتی فضایی x، y، z، نقطه را همراه با زمان t در نظر گرفت. در این صورت، حرکت نقطه به وسیلهٔ یک منحنی در فضای چهار بعدی و با مختصات x، y، z، t نشان داده می‌شود؛ با حرکت نقطه، هر چهار مختص تغییر می‌کند: جای (x, y, z) و زمان t. ولی در این‌جا، همراه‌کردن فضا با زمان، خصلتی صوری دارد و هیچ‌گونه ضرورت درونی ارتباط بین فضا و زمان برقرار نمی‌شود. البته، در قانون حرکت هر جسم مفروض، رابطه‌ای بین موضع

فضایی آن با زمان وجود دارد. ولی این، تنها مربوط به هر حرکت مفروضی است و هیچ گونه بستگی کلی درونی بین فضا و زمان، نه در مکانیک و نه به طور کلی در فیزیک پیش از نظریهٔ نسبیت وجود ندارد. همیشه و به صورتی یک ارزشی، رابطه‌های فضایی، ردیف فضایی چیزها پدیده‌ها، از رابطه‌ها و ردیف آن‌ها در زمان تشخیص داده می‌شود. دنبالهٔ پیش‌آمدهای زمانی و تداوم فاصله‌های زمانی به صورتی مطلق در نظر گرفته می‌شوند که به هیچ چیز دیگری مربوط نیستند. سخن کوتاه، مفهوم زمان مطلق و انتزاعی، بدون هیچ انباز و شریکی، بر فیزیک حکومت می‌کرد.

کار عظیم *اینشتین*، تنها در کشف پایه‌های اساسی نظریهٔ نسبیت نبود، بلکه در این هم بود که توانست در درک فلسفی و فیزیکی مسأله‌های مربوط به فضا و زمان هم، دگرگونی جدی پدید آورد و روشن کند که، درواقع، چیزی به نام زمان مطلق وجود ندارد. بلافاصله بعد از آن که *اینشتین* در سال ۱۹۰۵ نظریهٔ خود را ساخت[1]، بین **کووسکی** نشان داد که، ماهیت این نظریه، بیش از آن چه مربوط به کنار گذاشتن مطلق بودن فضاست، به ایجاد بستگی بین فضا و زمان مربوط می‌شود که، به یاری آن، به شکل انتزاعی یگانه‌ای از فضا، به معنای وجود ماده، رسید: «فضا ــ زمان» جدا کردن فضا ــ مختصات فضایی ــ از زمان و از مختص زمانی t، تا حد معینی نسبی و مربوط به دستگاه مادی است ــ «دستگاه محاسبه‌ای» که در رابطه با آن، ردیف فضایی و زمانی پدیده‌ها معین می‌شود. پیش‌آمدهایی که نسبت به دستگاه، هم زمان‌اند، ممکن است نسبت به دستگاه دیگر، هم زمان نباشند.

البته، تعیین ردیف پدیده‌ها، بستگی کافی با دستگاه محاسبه ندارد و نمی‌تواند داشته باشد. به خودی خود معلوم است، ردیف پیش‌آمدهایی که با تأثیر متقابل مستقیم به هم بستگی دارند، در رابطهٔ با همهٔ دستگاه‌ها، به یک صورت است، به نحوی که همیشه، اثر پیش از نتیجه قرار دارد. ولی ردیف پیش‌آمدهایی که با تأثیر متقابل به هم بستگی ندارند، می‌تواند در زمان نسبی باشد. چون ردیف فضایی (به صورت خالص آن)، بستگی به پیش‌آمدهای هم زمان دارد، و هم زمانی نسبی است، جدا کردن رابطه‌های خالص فضایی از مجموعهٔ کلی رابطه‌های «فضا ـ زمان» نسبی از آب در می‌آید و بستگی به دستگاه محاسبه دارد. فضا به مفهوم مطلق خود به معنای «بریدن» مجموعهٔ فضا ــ زمان چهاربعدی و (در

---

۱. نظریه‌ای را که *اینشتین* در سال ۱۹۰۵ آورد، نظریهٔ خصوصی نسبیت نام گرفت و نظریهٔ «عمومی» او در سال ۱۹۱۵ تدوین شد.

رابطه با دستگاه مفروض) در میان پیش‌آمدهای هم زمان است.

ما نمی‌توانیم در این جا به شرح پایه‌های نظریهٔ نسبیت بپردازیم، همین قدر تلاش کردیم، به‌کوتاه‌ترین صورتی که ممکن بود، این مطلب را روشن کنیم که چگونه این نظریه می‌تواند طبیعی‌ترین روش درک هندسهٔ انتزاعی باشد. و این درک، به‌کلی غیر از آن چیزی است که خود *اینشتین* آغاز کرده بود.

جهان را می‌توان همچون مجموعه‌ای از پیش‌آمدهای گوناگون بررسی کرد. در ضمن، منظور از پیش‌آمد، پدیده‌ای نیست که در فضا پراکنده شود و در طول زمان حمل کند، بلکه پدیده‌ای لحظه‌ای و نقطه‌ای، همچون جرقه‌ای لحظه‌ای از یک لامپ نقطه‌ای است. به زبان هندسه، پیش‌آمد عبارت است از نقطه‌ای در جهان چهاربعدی.

«فضا ـ زمان» شکلی از وجود ماده و شکلی از گوناگونی جهان است. ساختار «فضا ـ زمان»، یعنی «هندسهٔ» آن، چیزی جز نوعی از ساختار جهان نیست، یعنی بنابر بررسی ما، «هندسهٔ» مجموعهٔ پیش‌آمدهاست. این ساختار، با برخی تأثیرهای متقابل و رابطه‌های مادی پیش‌آمدها، به صورت کلی خود، تعریف می‌شود، به‌ویژه:

اول، همان‌طور که پیش از این هم دربارهٔ رابطه‌های فضایی گفتیم، این‌ها باید رابطه‌ها و تأثیرهای متقابل مادی باشند. همین وضع دربارهٔ رابطهٔ پدیده‌ها در زمان هم درست است. رابطه‌های فضایی و زمانی، به خودی و به صورت خالص، تنها مفهوم‌هایی انتزاعی‌اند.

دوم، رابطه‌های بین پیش‌آمدها، که ساختار «فضا ـ زمان» را معین می‌کنند، بنابر خصلت همه جانبهٔ «فضا ـ زمان» باید خصلتی عام داشته باشند.

چنین رابطهٔ مادی عام بین پیش‌آمدها، عبارت است از یک بستگی همه‌جانبه که هم سببی است و هم نتیجه‌ای. هر پیش‌آمدی، به نحوی، مستقیم یا غیرمستقیم، بر برخی پیش‌آمدهای دیگر تأثیر می‌گذارد و، به نوبهٔ خود، تأثیر پیش‌آمدهای دیگر را تحمل می‌کند. همین رابطهٔ مربوط به تأثیر یک پیش‌آمد بر پیش‌آمدهای دیگر است که ساختار «فضا ـ زمان» را معین می‌کند.

بنابراین، نظریهٔ نسبیت، امکان تعریف زیر را به ما می‌دهد. «فضا ـ زمان» عبارت است از مجموعهٔ همهٔ پیش‌آمدها که از همهٔ ویژگی‌ها و رابطه‌های خود، به‌جز رابطهٔ کلی تأثیر برخی پیش‌آمدها بر برخی دیگر، جدا شده باشد، خود این رابطه هم باید به مفهوم کلی در نظر گرفته شود که از همهٔ صورت‌های مشخص خود جدا شده باشد.

در نظریهٔ نسبیت خصوصی، «فضا ـ زمان» تا بیشترین حد ممکن، همگن در نظر گرفته

می‌شود. و این، به معنای آن است که پیش‌آمدهای گوناگون را می‌توان تبدیل کرد، به نحوی که رابطهٔ تأثیر بین پیش‌آمدها به هم نخورد و، در ضمن، گروه این تبدیل‌ها، به مفهوم معینی، کوچک نیست. دو زوج پیش آمد A، Bو A'، B'را در نظر می‌گیریم، با این شرط که A بر B تأثیر نکند و همینطور Bهم بر Aتأثیری نداشته باشد؛ به همین ترتیب برای دو پیش‌آمد A' ، B'. در این صورت، چنان رابطه‌ای بین پیش‌آمدها وجود دارد که، برای آن، Aمتناظر A' و B متناظر B'است؛ در ضمن برای هر زوج پیش‌آمدی، رابطهٔ تأثیر (یا عدم تأثیر) خراب نمی‌شود، یعنی اگر Xبر Y تأثیر می‌گذارد، X'هم بر Y' تأثیر بگذارد، و اگر Xبر Y تأثیر نمی‌گذارد، X'هم بر Y' تأثیر نگذارد.

به این ترتیب، از دیدگاه نظریهٔ نسبیت خصوصی، «فضا ـ زمان» نوعی فضای چهار بعدی است که هندسهٔ آن، با گروهی از تبدیل‌ها تعریف می‌شود. این تبدیل‌ها هم، همان تبدیل‌های مشهور *لورنس* هستند. قانون‌های هندسه و فیزیک، ضمن این تبدیل‌ها، تغییر نمی‌کنند. این دیدگاه نسبت به هندسه، شبیه دیدگاهی است که **کلاین** در برنامهٔ ارلانگن ارائه داده بود و در بند ۶ دربارهٔ آن صحبت کردیم.

۶. نظریهٔ نسبیت عمومی جلوتر می‌رود؛ در این نظریه، همگنی «فضا ـ زمان» کنار گذاشته می‌شود. در این نظریه، «فضا ـ زمان» تنها در حوزه‌های به اندازهٔ کافی کوچک و آن هم به تقریب، همگن به حساب می‌آید، ولی در کل خود، فضایی ناهمگن است. بنابر نظریهٔ *اینشتین*، ناهمگنی «فضا ـ زمان» با پراکندگی و حرکت ماده مشخص می‌شود، ساختار «فضا-زمان»، به‌نوبهٔ خود، قانون‌های حرکت جسم‌ها را معین می‌کند که آن را می‌توان در پدیده‌های مربوط به‌جاذبهٔ عمومی کشف کرد. نظریهٔ نسبیت عمومی، شبیه نظریهٔ جاذبه، ساختار «فضا ـ زمان» را به‌یاری حرکت ماده توضیح می‌دهد.

تصور دربارهٔ «فضا ـ زمان»، که تنها با تقریب معینی در حوزه‌های کوچک همگن است، شبیه تصور دربارهٔ فضای ریمانی است که تنها در محدودهٔ «بی‌نهایت» کوچک، اقلیدسی است از دیدگاه ریاضی، «فضا ـ زمان» در نظریه عمومی نسبیت، همچون فضای ریمانی، و البته با مفهومی تغییر یافته، تعبیر می‌شود.

به‌ویژه، در فضای چهار بعدی ریمانی در نزدیکی هر نقطه، می‌توان مختصات را به‌نحوی در نظر گرفت که مربع عنصر خطی با این دستور بیان شود:

$$ds^۲ = dx_۱^۲ + dx_۲^۲ + dx_۳^۲ + dx_۴^۲$$

در «فضا ـ زمان» می‌توان، در همسایگی هر پیش‌آمد، مختصات x ، y ، z ، t را طوری وارد کرد که عنصر خطی با دستور

$$ds^{2} = dx^{2} + dy^{2} + dz^{2} - c^{2}dt^{2}$$

بیان شود که در آن، c عبارت است از سرعت نور، و باتوجه به‌انتخاب واحد اندازه‌گیری، می‌توان آن را واحد به‌حساب آورد. در این جا، x ، y ، z، مختصات فضایی و t زمان است. وجود علامت منفی برای dt^{2}، بیان صوری این حقیقت است که مختصات زمانی، به‌طور ریشه‌ای، با مختصات فضایی (یعنی با فضا) فرق دارد.

در نظریهٔ جاذبه، مهم‌ترین نقش به‌عهدهٔ مفهوم انحنای «فضا ـ زمان» است. معادله‌های اصلی این نظریه، که به‌وسیلهٔ *اینشتین* داده شد، معرف بستگی مقدارهای مربوط به‌انحنای فضا ـ زمان، با مقدارهای مربوط به‌پراکندگی و حرکت ماده است. این معادله‌ها، در عین حال، معادله‌های میدان جاذبه‌اند و، همان‌طور که *اینشتین* با همکاری *و.آ. فوک* ثابت کرد، قانون‌های حرکت جسم در میدان جاذبه رامشخص می‌کنند.

ساختار «فضا ـ زمان» در نظریهٔ نسبیت عمومی مرکب و پیچیده است و در آن، نمی‌توان، حتی درحدی که در نظریهٔ نسبیت خاص ممکن است، فضا را از زمان جدا کرد. ولی با تقریب معینی و با پیش‌فرض‌های معینی، می‌توان این جدایی را انجام داد. فضا در حوزه‌هایی که در مقایسه با مقیاس‌های کیهانی کوچک باشد، با دقتی کافی، اقلیدسی است، ولی در حوزه‌های بزرگ، از هندسهٔ اقلیدسی منحرف می‌شود. این، انحراف، مربوط به‌پراکندگی و جرم ماده است و خود را ، ولو به‌میزان خیلی کوچک، در نزدیکی ستاره‌های با جرم بالا نشان می‌دهد؛ و در مقیاس‌های کیهانی به‌طور کامل و ملموس با این انحراف برخورد می‌کنیم. در ردیف فرضیه‌هایی که برای ساختار کیهان در مجموع خود، وجود دارد، فرض می‌کنند که جرم به‌تقریب و به‌طور متوسط، به‌صورتی یکنواخت پراکنده شده است. *فریریمان*، ریاضی‌دان و فیزیک‌دان شوروی، نظریه‌ای دارد که، بنابر آن، هندسهٔ فضا در مجموع خود، با هندسهٔ *لباچوسکی* سازگار است.

هندسهٔ انتزاعی، در نظریهٔ نسبیت کاربرد خود را پیدا کرد؛ به‌جز این، اندیشه‌های مربوط به‌فضای انتزاعی، به‌عنوان وسیله‌ای برای عمیق‌ترکردن این نظریهٔ و تنظیم دقیق‌تر آن، مورد استفاده قرار گرفت. نیرو و امکان هندسهٔ انتزاعی در بررسی واقعیت‌هاست و اندیشه انتزاعی توانست در این‌جا، پیروزی خود را جشن بگیرد. هندسهٔ انتزاعی، که از بررسی

تجربی رابطه‌ها و شکل‌های جسم‌هایی فضایی برخاسته است، امروز برای بررسی فضای واقع، و به‌عنوان روش حاضر و آماده‌ای از ریاضیات، به‌کار می‌رود. مسیر هر دانشی همین‌گونه است: از تجربهٔ مستقیم آغاز می‌کند، به‌سمت تعمیم نظری و انتزاع پیش می‌رود و سپس دوباره، به عنوان ابزاری برای شناخت دقیق‌تر ماهیت پدیده‌ها، به‌تجربه باز می‌گردد، پدیده‌های آشنا را دقیق‌تر می‌شناساند، پدیده‌های تازه‌ای را پیش‌بینی می‌کند، به‌فعالیت‌های عملی انسان یاری می‌رساند و از این راه، روش درستِ پیشرفت آیندهٔ خود را پیدا می‌کند.

بخش هجدهم

توپولوژی

پ. س. آلکساندرف

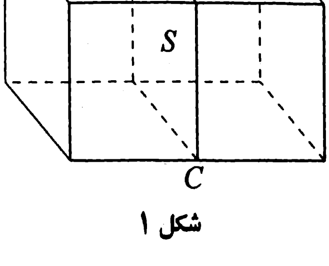
۱. موضوع توپولوژی

سایش و تماس (یا مجاورت) که ویژه جسم‌هاست. ما وقتی می‌توانیم جسمی را هندسی بنامیم که، جدا از دیگر خاصیت‌ها، اعم از پایدار یا تصادفی، فقط این ویژگی جسم را درنظر بگیریم».

نیکلای ایونوویچ لباچوسکی، بخش نخست کتاب «مقدمات تازهٔ هندسهٔ» خود را، با این واژه‌ها آغاز می‌کند.

معنای این واژه‌ها را روی شکل روشن می‌کنیم (شکل ۱). *لباچوسکی* ادامه می‌دهد: «دو جسم A و B، ضمن تماس باهم، یک شکل هندسی C را تشکیل می‌دهند ... برعکس، هر جسم C، به‌وسیلهٔ بُرش دل‌خواه S، به‌دو بخش A و B تقسیم می‌شود».

این مفهوم سایش و تماس، که به‌معنای همسایگی و «بی‌نهایت نزدیکی» است، و هم‌چنین مفهوم دوگان آنها، یعنی بُرش جسم، مفاهیمی هستند که *لباچوسکی* در اساس تمامی هندسه قرار داده است، در واقع، پایه‌های عمدهٔ توپولوژی را در تمامی حجم آن،

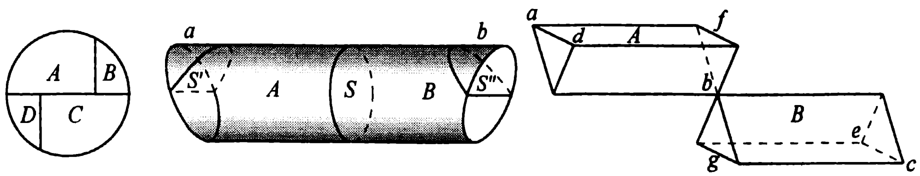


شکل ۱

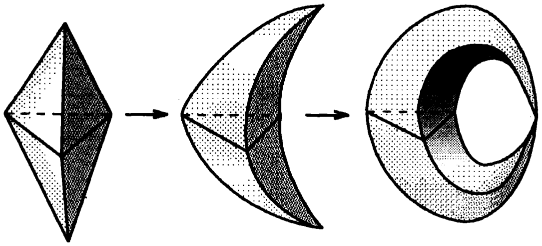
به‌مفهومی که امروز از این شاخهٔ ریاضیات می‌شناسیم، مشخص می‌کند. بنابراین، تفسیرهایی که امروز از کار این هندسه‌دان بزرگ می‌شود، درست است که: «*لباچوسکی*، نخستین کسی در تاریخ دانش‌های ریاضی است که در ساختار هندسه، از ویژگی‌های توپولوژیک جسم‌ها استفاده کرده است ... *لباچوسکی* مفهوم‌های سطح، خط و نقطه را با اصطلاح‌های بُرش و سایش جسم‌ها تعریف می‌کند». تصوری را که *لباچوسکی* دربارهٔ گونه‌های مختلف مضمون مشخص هندسی داشته است و می‌تواند مفهوم‌های سایش و برش حبه‌ها را منعکس کند، می‌توان از شکل ۲ که از همان کتاب *لباچوسکی* برداشته شده است، دریافت.

هر تبدیلی از شکل‌های هندسی که ضمن آن، رابطهٔ سایش (یا مجاورت) بین بخش‌های مختلف شکل آن ویران نشود، **تبدیل پیوسته** نامیده می‌شود؛ اگر نه تنها سایش ویران نشود، بلکه سایش تازه‌ای هم به‌وجود نیاید، آن وقت با **تبدیل توپولوژیک** سروکار داریم. به‌این ترتیب، ضمن تبدیل توپولوژیک یک شکل، بخش‌هایی از شکل که در سایش با یکدیگرند، در سایش باهم باقی می‌مانند و بخش‌هایی که در سایش با یکدیگر نیستند، نمی‌توانند با یکدیگر سایش پیدا کنند؛ به‌زبان کوتاه‌تر، ضمن تبدیل توپولوژیک، نه‌پارگی (ناپیوستگی) به‌وجود می‌آید و نه بخش‌هایی از شکل روی هم قرار می‌گیرند. به‌ویژه، دو نقطهٔ مختلف نمی‌توانند برهم قرار گیرند و به یک نقطه منجر شوند (اگر دو نقطهٔ مختلف، به‌یک نقطه منجر شوند، به‌معنای آن است که سایش تازه‌ای پدید آمده است: شکل ۳). بنابراین، تبدیل توپولوژیک یک شکل هندسی، به‌عنوان مجموعه‌ای از نقطه‌های تشکیل دهندهٔ آن، تبدیلی است نه تنها پیوسته، بلکه در ضمن در تناظر یک به‌یک (یعنی یک ارزشی): هر دو نقطهٔ مختلف شکل، ضمن تبدیل به‌دو نقطهٔ مختلف منجر می‌شوند. در نتیجه، تبدیل توپولوژیک، عبارت است از تبدیلی که در هر دو جهت یک ارزشی و پیوسته‌اند.

از دیدگاه عینی، تبدیل توپولوژیک یک شکل (خط، سطح و غیره) را می‌توان به‌این



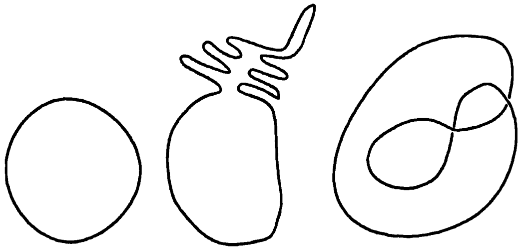
شکل ۲



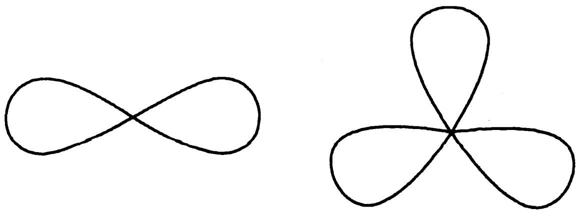
شکل ۳

صورت تصور کرد:

فرض می‌کنیم، شکلی که از ماده‌ای نرم و کشسان، مثل لاستیک، ساخته شده است، در اختیار داشته باشیم. در این صورت، می‌توان شکل را به‌صورتی پیوسته تغییر داد و از حالت نخستین خود دگرگون کرد: در برخی بخش‌ها «کشیده» و در برخی بخش‌های دیگر «به‌هم فشرده» و، به‌طور کلی، در شکل و اندازه‌ها، دگرگون شود. فرض کنیم منحنی بسته‌ای، مثل دایره، از لاستیک داشته باشیم، می‌توانیم آن را از دو طرف بکشیم و به‌یک بیضی تبدیل کنیم، می‌توانیم یک چند ضعلی منتظم یا نا منتظم از آن بسازیم و می‌توانیم به‌صورت منحنی بستهٔ عجیب و غریبی، که دو نمونهٔ آن در شکل ۴ نشان داده شده است، درآوریم. ولی حق نداریم آن را به‌صورت عدد هشت لاتین درآوریم (چرا که، برای این منظور باید دو نقطهٔ آن را برهم منطبق کنیم: شکل ۵) و یا از آن، یک پاره‌خط راست درست کنیم (چرا که، برای این منظور، یا باید نیمی از محیط دایره را بر نیم دیگر آن قرار داد و یا محیط دایره را، در یکی از نقطه‌های آن، پاره کرد). دایره، منحنی بستهٔ ساده‌ای است که تنها از یک «حلقه» تشکیل شده است، در حالی که عدد هشت لاتین (یا نماد بی‌نهایت: ∞) شامل دو حلقه و یا منحنی «سه برگی» (شکل ۵) شامل سه حلقه است. این ویژگی دایره که منحنی بستهٔ ساده‌ای است،



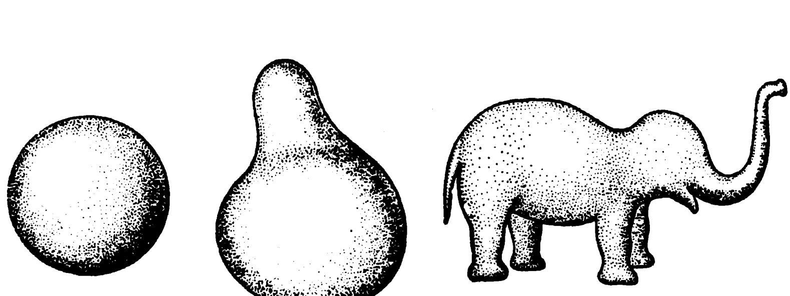
شکل ۴



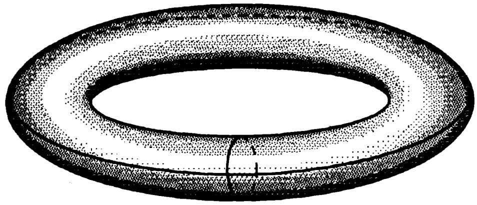
شکل ۵

چنان ویژگی است که، ضمن هر تبدیل توپولوژیک، تغییر نمی‌کند و، به‌همین جهت، به‌آن ویژگی توپولوژیک گویند.

اگر سطحی کره‌ای را در نظر بگیریم، که در شکل یک توپ لاستیکی با جدار نازک باشد، آن وقت می‌توانیم به‌یاری تبدیل توپولوژیک به‌شکل‌های گوناگونی در آوریم (شکل ۶). ولی به یاری تبدیل توپولوژیک، نمی‌توانیم سطح کره‌ای را به‌مربع یا یک سطح حلقه‌ای (مثل نان حلقه‌ای، یا لاستیکی برای نجات از غرق شدن به‌کار می‌رود)، یعنی به‌شکل چنبره (شکل ۷) درآورد. در واقع، سطح کره دو ویژگی دارد که، ضمن تبدیل توپولوژیک تغییر نمی‌کنند. ویژگی اول به‌بسته بودن سطح مربوط می‌شود: دارای مرزهایی نیست (در حالی که مربع، مرز دارد)؛ ویژگی دوم، مربوط به‌این است که، هر منحنی بستهٔ روی آن، به‌بیان *لباچوسکی* مقطعی از آن است، یعنی اگر سطح کرهٔ لاستیکی را روی منحنی بسته‌ای واقع بر این سطح ببریم، سطح به‌دو بخش جدا از هم تبدیل می‌شود. چنبره، این ویژگی را ندارد. اگر سطح چنبره را روی نصف‌النهار برش دهیم، به‌بخش‌های جداگانه تقسیم نمی‌شود (شکل ۷)، به‌لوله‌ای منجر می‌شود که با تبدیل توپولوژیک، به‌سادگی و با راست کردن آن،



شکل ۶

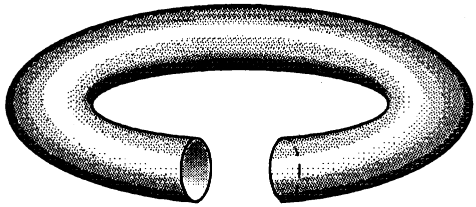


شکل ۷

به‌استوانه تبدیل می‌شود (شکل ۸). به‌این ترتیب، برخلاف کره، هر منحنی بسته روی سطح چنبره، مقطعی از آن نیست (یعنی آن را به‌دو بخش جدا از هم منجر نمی‌کند). بنابراین سطح کروی را نمی‌توان با تبدیل توپولوژیک، به چنبره تبدیل کرد. می‌گویند، کره و چنبره، از نظر توپولوژی سطح‌هایی متفاوت‌اند، یا متعلق به‌گونه‌های توپولوژیک متفاوتی هستند و یا، سرانجام، این سطح‌ها، **همسانریخت** (یا همسان یا همانریخت) نیستند. برعکس، کره و بیضوی و، به‌طور کلی، هر سطح محدود و محدب، متعلق به‌یک گونه توپولوژیک‌اند، یعنی با هم همسانریخت‌اند. و این به‌معنای آن است که آن‌ها را با تبدیل توپولوژیک، می‌توان به‌هم تبدیل کرد.

۲. سطح‌ها

همان‌طور که گفتیم، هر ویژگی از یک شکل هندسی، که ضمن هر تبدیل توپولوژیک آن تغییر نکند، ویژگی توپولوژیک‌شکل نامیده می‌شود. کار توپولوژی، بررسی ویژگی‌های توپولوژیک شکل‌هاست؛ به‌جز این، در توپولوژی، هرگونه تبدیل پیوستهٔ شکل‌های هندسی هم بررسی می‌شود.

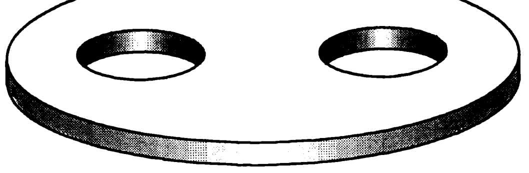


شکل ۸

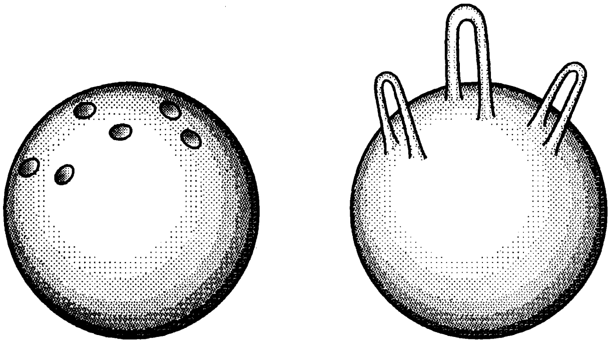
مثال‌هایی از ویژگی‌های توپولوژیک را دیدیم. به‌این ترتیب، بعضی ویژگی‌های توپولوژیک چنین‌اند: ویژگی بسته بودن منحنی یا سطح، ویژگی ساده بودن یک منحنی بسته (یعنی منحنی‌ای که یک حلقه تشکیل بدهد)، ویژگی سطح بسته، که هر منحنی بستهٔ روی آن، مقطعی از سطح باشد (سطح کره دارای این ویژگی است، ولی سطح چنبره این ویژگی را ندارد) و غیره.

بیشترین تعداد منحنی‌های بسته‌ای را که می‌توان روی سطح مفروض رسم کرد، به‌نحوی که این منحنی‌ها مقطع تشکیل ندهند (یعنی به شرطی که روی همهٔ این منحنی‌ها برش انجام دهیم سطح به بخش‌های جداگانه برش نخورد) **مرتبهٔ همبندی** سطح گویند. این عدد، به‌احتمالی، می‌تواند مهم‌ترین آگاهی را دربارهٔ ساختمان توپولوژیک سطح به‌ما بدهد. دیدیم، مرتبهٔ همبندی برای سطح کره، برابر صفر است (نمی‌شود منحنی بسته‌ای روی آن رسم کرد که مقطعی از سطح نباشد). روی چنبره، دو منحنی بسته می‌توان یافت که در مجموع خود (یعنی با برش روی هردو) مقطعی به‌وجود نیاید: یکی از این منحنی‌ها را می‌توان هر نصف‌النهاری از آن، و دیگری را موازی چنبره گرفت (شکل ۷). ولی روی چنبره نمی‌توان سه منحنی بسته رسم کرد که با برش روی هر سه منحنی، مقطعی پدید نیاید؛ لذا مرتبهٔ همبندی چنبره برابر است با ۲. مرتبه همبندی سطحی شبیه شکل ۹، برابر است با ۴ و غیره. به‌طور کلی، سطح کروی را در نظر می‌گیریم و روی آن به‌تعداد $2p$ سوراخ گرد ایجاد می‌کنیم (در شکل ۱۰، حالت $p=3$ نشان داده شده است). این سوراخ‌ها را به p زوج تقسیم و در هر زوج، لوله‌های استوانه‌ای وار («دستگیره») قرار می‌دهیم. شکلی کروی با p «دستگیره» یا آن طور که می‌گویند، سطح نرمال با گونای p به‌دست می‌آید. مرتبهٔ همبندی این سطح برابر است $2p$.

همهٔ این گونه سطح‌ها، به‌بیان ***لباچوسکی***، «مقطع‌هایی» از فضا هستند: هر کدام از آن‌ها، فضا را به‌دو حوزه تقسیم می‌کنند، حوزه درونی و حوزهٔ بیرونی؛ در ضمن، مرزهای این



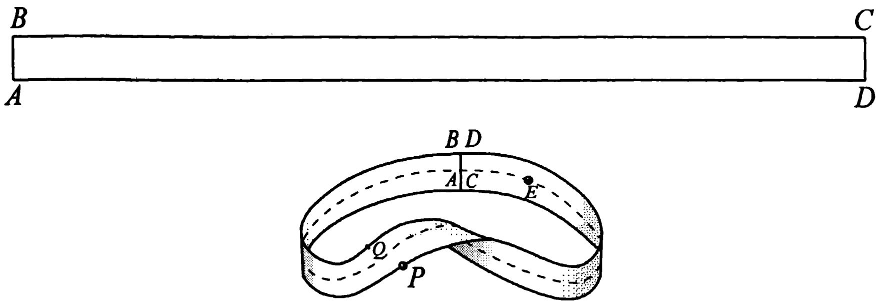
شکل ۹



شکل ۱۰

حوزه‌ها برهم منطبق‌اند. این موقعیت، معنای دیگری دارد که برای ما جالب‌تر است؛ هریک از این سطح‌ها دو رویه دارند: رویهٔ بیرونی و رویهٔ درونی (یک رویه را می‌توان با یک رنگ و رویهٔ دیگر را با رنگ دوم رنگ‌آمیزی کرد).

ولی، در کنار این سطح‌ها، سطح‌های به‌اصطلاح یک رویه هم وجود دارند، سطح‌هایی که، در آن‌ها دو رویهٔ جداگانه وجود ندارد. مشهورترین و در ضمن ساده‌ترینِ این سطح‌های یک رویه، «**نوار موبیوس**» است. **نوار موبیوس** را می‌توان به‌این ترتیب به‌دست آورد: نوار کاغذی مستطیلی شکل $ABCD$ را انتخاب می‌کنیم، سپس ضلع‌های روبه‌رو و کوچکتر AB و CD از مستطیل $ABCD$ را طوری روی هم می‌گذاریم که رأس A بر رأس C و رأس B بر رأس D قرار گیرد. سطحی به‌دست می‌آید که در شکل ۱۱ نشان داده شده است و «**نوار موبیوس**» نامیده می‌شود. خیلی ساده می‌توان تحقیق کرد که، این سطح، دو رویه ندارد که بتوان هر یک از آن‌ها را به‌رنگی غیر از رنگ رویهٔ دیگر درآورد. کنارهٔ این سطح، یک منحنی



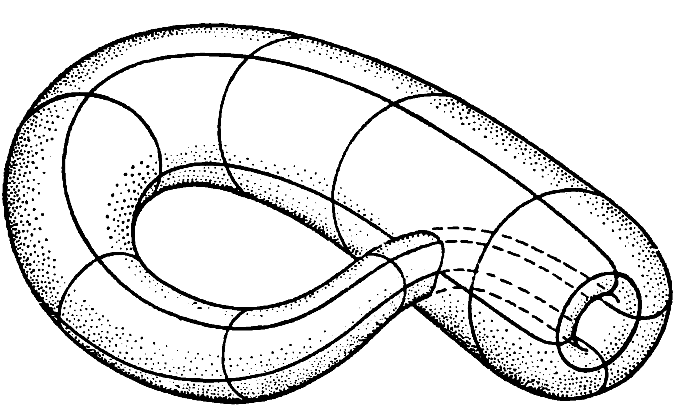
شکل ۱۱

بسته تشکیل می‌دهد که اگر از نقطه‌ای واقع بر آن، مثل نقطهٔ P حرکت کنیم، از همهٔ نقطه‌های کناره می‌گذرد و دوباره به P می‌رسد. همچنین اگر از نقطهٔ میانهٔ نوار موبیوس، مثل E و روی سطح حرکت کنیم، از تمامی سطح عبور می‌کنیم و دوباره به E می‌رسیم. نوار موبیوس، سطحی یک رویه است و یک لبه دارد.

این پرسش پیش می‌آید: آیا سطح بستهٔ یک رویه وجود دارد، یعنی سطح یک رویه‌ای که مرز نداشته باشد؟ چنین سطح‌هایی وجود دارند، ولی این گونه سطح‌ها، که نمی‌توان آن‌ها را در فضای سه بعدی قرارداد، همیشه با خودشان برخورد دارند. نمونهٔ مشخص سطح بستهٔ یک رویه در شکل ۱۲ نشان داده شده است که به‌آن «چنبرهٔ یک رویه» یا **بطری کلاین** می‌گویند. اگر از برخورد سطح با خودش واهمه ندارید، می‌توانید، با تصور ذهنی، دو نمونه ***نوار موبیوس*** در نظر بگیرید و کناره یا مرز یک نوار را روی مرز نوار دوم قرار دهید (گفتیم مرز ***نوار موبیوس***، تنها از یک منحنی بسته تشکیل شده است)، آن وقت **بطری کلاین** به‌دست می‌آید.

اکنون می‌توانیم قضیهٔ اصلی توپولوژی سطح‌ها را، برای سطح‌های دو رویه بیاوریم: هر سطح بستهٔ دو رویه، همسانریخت (هومئومورف *homeomorphic*) یک سطح نرمال با گونای p، یعنی «کره‌ای با p دستگیره»، است. دو سطح بستهٔ دو رویه همسانریخت هستند، وقتی و تنها وقتی که گوناهای آن‌ها برابر p باشد (یعنی از مرتبهٔ همبندی $2p$)، یعنی وقتی که همسانریخت سطح کره‌ای با p دستگیره باشند.

برای سطح‌های یک رویه هم، شبیه صورت‌های نرمال سطح‌های دو رویه از خانواده p،



شکل ۱۲

«صورت نرمال» وجود دارد، ولی آن‌ها را دشوارتر می‌توان تصور کرد. برای این منظور باید کره‌ای را در نظر گرفت که دارای p سوراخ گرد باشد و برای هر سوراخ یک ***نوار موبیوس*** انتخاب کرد، به‌نحوی که مرز نوار روی مرز سوراخ متناظر قرار گیرد. دشواری، ضمن تلاش برای تصور نحوه قرار گرفتن لبه نوار بر مرز سوراخ، از این جا ناشی می‌شود که، در عمل، نمی‌توان آن را انجام داد: ضمن تلاش برای انجام این عمل، برخورد سطح با خودش پیش می‌آید که برای هر سطح بستهٔ یک رویه در فضا، ناچاری است.

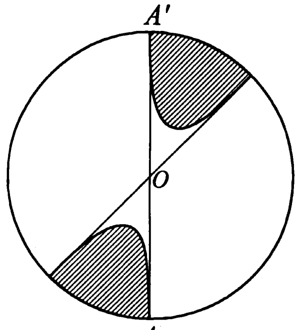
نباید گمان کرد که، سطح‌های بستهٔ یک رویه، به‌حوزهٔ معماها و سرگرمی‌های ریاضی مربوط می‌شود و هیچ ربطی به‌مسأله‌های جدی دانش ندارد. برای این که در این باره قانع شویم، کافی است به‌یاد آوریم، یکی از موفقیت‌های اساسی تفکر هندسی، پیدایی هندسهٔ تصویری است که عنصرهای آن، در برنامهٔ دورهٔ هنر دانشگاه‌ها و مرکزهای تربیت معلم گنجانده شده است. سرچشمهٔ تاریخی هندسهٔ تصویری را باید در نظریهٔ پرسپکتیو و تصویر مرکزی دانست، که خود این هم، در دوران نوزایی (رنسانس، ***لئوناردو داوینچی***) در بستگی با نیازهای معماری، نقاشی و تصویر برداری صنعتی به‌وجود آمد. نخستین قضیه‌های هندسهٔ تصویری، در سده‌های شانزدهم و هفدهم کشف شد. هندسهٔ تصویری که در رابطهٔ کامل با نیازهای عملی پدید آمد، در مسیر پیشرفت خود، به‌یکی از معتبرترین اندیشه‌ها و نظریه‌های مربوط به‌تعمیم هندسه تبدیل شد. به‌برکت هندسهٔ تصویری، از جمله، برای نخستین بار، مفهوم هندسهٔ نااقلیدسی ***لباچوسکی***، به‌طور کامل فهم شد (برای نمونه بند ۶ از بخش هفدهم را ببینید).

گذر از صفحهٔ عادی، آن‌طور که در هندسهٔ مقدماتی آموزش داده می‌شود، به‌هندسهٔ تصویری، به‌معنای تکمیل کردن صفحه با عنصرهای انتزاعی تازه، یعنی نقطه‌های «ناسره» یا نقطه‌های بی‌نهایت دور بوده. تنها پس از چنین تکمیلی بود که تصویر یک صفحهٔ به‌صفحهٔ دیگر، تبدیلی یک ارزشی و در تناظر یک به‌یک از صفحه‌ای بر صفحهٔ دیگر شد. تکمیل صفحه با نقطه‌های ناسره، که در هندسهٔ تحلیلی به‌معنای گذر از مختصات دکارتی معمولی به‌مختصات همگن است، به‌این ترتیب انجام گرفت. هر خط راست به‌وسیله یک نقطه (و تنها یک نقطه) ناسره («بی‌نهایت دور») تکمیل می‌شود؛ در ضمن دو خط راست، وقتی و تنها وقتی نقطهٔ مشترک ناسره‌ای دارند که موازی باشند. وقتی نقطهٔ یگانهٔ بی‌نهایت دور را به‌خط راست اضافه کنیم، به‌صورت خطی بسته در می‌آید؛ در ضمن مجموعهٔ نقطه‌های بی‌نهایت دورِ همهٔ خط‌های راست ممکن، بنابر تعریف، خط راست ناسره یا خط راست

بی‌نهایت دور را می‌سازند.

از آن جا که خط‌های راست موازی دارای نقطهٔ مشترک بی‌نهایت دورند، بنابراین، برای این‌که تمامی روند تکمیل صفحه را با نقطه‌های ناسره پیش خود مجسم کنیم، کافی است خط‌های راستی را در نظر بگیریم که از یک نقطهٔ صفحه، و در شکل از مبدأ مختصات O گذشته‌اند (شکل ۱۳). نقطه‌های ناسره این خط‌های راست، شامل همهٔ نقطه‌های ناسرهٔ صفحهٔ تصویری‌اند (زیرا نقطهٔ ناسره هر خط راست صفحه، همان نقطهٔ ناسرهٔ خط راستی است که از نقطهٔ O موازی آن رسم شود). به‌این‌ترتیب «مدل» صفحهٔ تصویری بدست می‌آید: دایره‌ای با شعاع «بی‌نهایت بزرگ» به‌مرکز نقطهٔ O است و، در ضمن، هر نقطهٔ دو انتهای قطر، یعنی A و A' از محیط این دایره، باید به‌عنوان یک نقطهٔ یگانهٔ «بی‌نهایت دور» خط راست AA' در نظر گرفته شود. خود محیط این دایره، در این ضمن، به‌خط راست بی‌نهایت دور منجر می‌شود؛ این موضوع را باید به‌طور جدی به‌یاد داشت که دو انتهای هر قطر این دایره، باید همچون نقطه‌های منطبق برهم (یعنی یک نقطهٔ یگانه) تصور شود. از این‌جا، بلافاصله‌نتیجه می‌شودکه، صفحهٔ تصویری، یک‌سطح بسته‌است وهیچ مرزی ندارد.

اگر روی صفحهٔ تصویری منحنی درجه دومی در نظر بگیریم که در این جا آن را به‌صورت یک هذلولی رسم کرده‌ایم (شکل ۱۳ را ببینید)، روشن می‌شود که این هذلولی، روی صفحهٔ تصویری، یک منحنی بسته است (تنها روی دو شاخهٔ خط راست بی‌نهایت دور، از هم جدا شده است). باتوجه به‌این که نقطه‌های دو انتهای قطر محیط دایرهٔ اصلی ما به‌هم چسبیده‌اند، بدون دشواری قانع می‌شویم، بخش هاشور خوردهٔ درون هذلولی، در



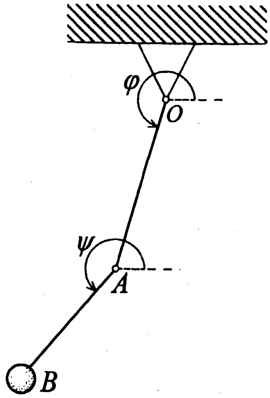
شکل ۱۳

شکل ۱۳، همسانریخت با درون دایرهٔ عادی، و بخش تکمیلی هاشور نخوردهٔ صفحهٔ تصویری، همسانریخت ***نوار موبیوس*** است. بنابراین، ***صفحهٔ تصویری*** از دیدگاه توپولوژی، عبارت است از برهم قرار دادن سطح دایره (و در حالت ما، درون هذلولی) با ***نوار موبیوس*** در مرز آن‌ها. از این جا نتیجه می‌شود که صفحهٔ تصویری، یعنی موضوع اصلی مطالعهٔ هندسهٔ تصویری روی صفحه، عبارت است از سطح بسته یک رویه.

نمونهٔ صفحهٔ تصویری، به‌جز ارزش بی‌اندازهٔ هندسی خود، از این جهت هم جالب است که به‌یاری آن می‌توان به‌روشنی یکی از ویژگی‌های اندیشهٔ امروزی را دربارهٔ هندسه تشخیص داد. اندیشه‌ای که بر پایهٔ کشف ***لباچوسکی*** تنظیم شد. اندیشهٔ هندسی، بنابر ویژگی و خصلت شکل هندسی، همیشه انتزاعی بوده است. در زمان ما، هندسهٔ پلهٔ دیگری از انتزاع را بالا رفته است. این انتزاع تازه، عبارت است از تکمیل صفحهٔ معمولی به‌وسیلهٔ عنصرهای تازهٔ انتزاعی، یعنی نقطه‌های ناسره. البته، این عنصرهای انتزاعی هم، بازتابی از واقعیت‌اند (هر «نقطهٔ ناسره»، چیزی جز معرف دسته‌ای از خط‌های راست انتزاعی نیست)، ولی در بررسی‌های ما همچون عنصرهای انتزاعی هندسی ظاهر می‌شوند که تنها به‌صورتی نه چندان کامل، می‌توان آن‌ها را به‌عنوان نتیجهٔ برهم قرارگرفتن دو انتهای هر قطر دایره تصور کرد (چیزی که از نظر فیزیکی قابل تحقق نیست). این گونه ساختمان‌های انتزاعی، در تمامی توپولوژی، به‌ویژه برای عبور از سطح به خمینه‌های (یا چندلایه‌های) سه‌بعدی و یا با بُعد بیشتر، ارزش فراوانی دارد.

۳. خمینه‌ها (یا چندلایه‌ها، manifolds)

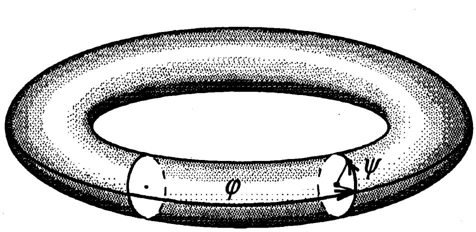
حالت ساده‌ای را بررسی می‌کنیم که به‌آن پاندول دوگانه سطحی هم می‌گویند (شکل ۱۴) و عبارت است از دو میلهٔ OA و AB، که با لولا به‌نقطهٔ A وصل شده‌اند؛ نقطهٔ O بی‌حرکت است، میلهٔ OA در صفحهٔ ثابتی آزادانه دور نقطهٔ O می‌چرخد و میلهٔ AB در همان صفحه، آزادانه دور نقطهٔ A دوران می‌کند. هر موقعیت ممکن این دستگاه، به‌طور کامل و با مفروض بودن زاویهٔ φ و ψ مشخص می‌شود. این‌ها زاویه‌هایی هستند که OA و AB با جهت ثابتی واقع در صفحه، از جمله با جهت مثبت محور طول، پدید می‌آورند. هر یک از این دو زاویه، از ۰ تا 2π تغییر می‌کند و می‌توانیم آن‌ها را به عنوان «مختصات جغرافیایی» واقع بر چنبره



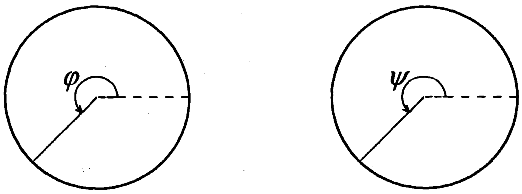
شکل ۱۴

به‌حساب آوریم که به‌ترتیب، از «استوای» چنبره و یکی از «نصف‌النهارهای» آن در نظر گرفته شده است (شکل ۱۵).

بنابراین، می‌توانیم بگوییم، خمینهٔ همهٔ حالت‌های ممکن در دستگاه مکانیکی ما، خمینه‌ای دو بعدی است، یعنی چنبره. با تغییر هر یک از زاویه‌های φ و ψ، متناظر با نقطه‌ای از محیط دایره، که روی آن نقطهٔ مبدأ و جهت محاسبهٔ کمان داده شده باشد، (شکل ۱۶)، می‌توانیم بگوییم، هر حالت ممکن دستگاه مکانیکی ما، وقتی به‌طور کامل مشخص می‌شود که به‌وسیلهٔ نقطه‌ای روی محیط هر یک از دو دایره داده شده باشد (روی یکی از آن‌ها عرض φ و روی دیگری طول ψ). به‌زبان دیگر، هر نقطه از صفحه ـ شبیه هندسهٔ تحلیلی ـ با دو عدد مشخص می‌شود (مختصات آن)، به‌نحوی که در این‌جا می‌توانیم هر نقطه‌ای از چنبره (یعنی هر موقعیتی از پاندول) را با دو مختص جغرافیایی آن، یعنی با دو نقطه مشخص کنیم که یکی بر محیط یک دایره و دیگری بر محیط دایرهٔ دیگر قرار دارد. این موقعیت به‌این معناست که همهٔ حالت‌های ممکن پاندول دوگانه سطحی ما، یعنی چنبره، عبارت است از



شکل ۱۵



شکل ۱۶

حاصل‌ضرب توپولوژیک دو دایره.

اکنون دستگاه خود را به‌صورت دیگری در نظر می‌گیریم. فرض می‌کنیم، مثل قبل، دستگاه شامل دو میلهٔ OA و AB باشد؛ در ضمن، میلهٔ OA بتواند آزادانه در صفحهٔ معینی دور نقطهٔ O بچرخد؛ میلهٔ AB در نقطهٔ A به‌میلهٔ OA با یک لولای کره‌ای مربوط باشد، به‌نحوی که در هر وضع مفروض نقطهٔ A، میلهٔ AB بتواند آزادانه دور این نقطه، در فضا دوران کند و در جهت هر نیم‌خط راستی که از A آغاز شده است، قرار گیرد. اکنون موقعیت دستگاه، با سه پارامتر مشخص می‌شود یکی از این پارامترها، مثل حالت قبل، زاویهٔ φ است که میلهٔ OA با جهت مثبت محور طول می‌سازد و، پارامتر دیگر، جهت میلهٔ AB را در فضا معین می‌کنند. جهت اخیر (یعنی جهت میلهٔ AB) را می‌توان از جمله به‌یاری نقطهٔ B' از سطح کرهٔ واحد به‌مرکز مبدأ مختصات O تعیین کرد که در آن جا شعاع OB' موازی میلهٔ AB کره را قطع می‌کند، یا این که دو مختص جغرافیایی نقطهٔ B' را روی کره، به‌دست داد. به‌این‌ترتیب، همهٔ حالت‌های گوناگون این دستگاه لولایی، یک خمینهٔ سه‌بعدی است و به‌سادگی می‌توان فهمید که آن را همچون حاصل‌ضرب توپولوژیک دایره و کره می‌توان تفسیر کرد.این، خمینه‌ای بسته است، یعنی مرزی ندارد و بنابراین نمی‌توان آن را به‌صورت شکلی واقع در فضای سه‌بعدی نشان داد. برای این‌که بتوانیم این خمینه را دست‌کم تا اندازه‌ای عینی به‌تصور آوریم، بخشی از فضا را که بین دو کرهٔ هم‌مرکز واقع است، در نظر می‌گیریم. هر شعاعی که از مرکز مشترک دو کره خارج شود، سطح دو کره را در دو نقطه سوراخ می‌کند. اگر فرض کنیم، هر دو نقطه از این نقطه‌ها، یکی شوند (به‌صورت یک نقطه به‌هم بچسبند)، آن وقت خمینهٔ سه‌بعدی به دست می‌آید که حاصل‌ضرب توپولوژیک کره در دایره است.

دستگاه لولایی خود را می‌توانیم پیچیده‌تر کنیم، یعنی فرض کنیم، به‌جز میلهٔ AB، بلکه میلهٔ OA هم بتواند آزادانه دور نقطهٔ O و در فضا بچرخد. مجموعهٔ همهٔ موقعیت‌های ممکن

این دستگاه، خمینه‌ای بسته و چهاربعدی است، یعنی حاصل‌ضرب دو کره.

به‌این ترتیب، می‌بینیم، ساده‌ترین دستگاههای مکانیکی، منجر به خمینه‌های توپولوژیک می‌شود که در ضمن، سه بعد یا بعدهای بیشتری دارند. برای بررسی مفصل‌تر مسأله‌های مکانیک، خمینه‌ها (و همچنین فضاهای چندبعدی) اهمیت زیاد دارند و ***فضاهای فازی*** دستگاه‌های دینامیکی نامیده می‌شوند و به‌یاری آن‌ها، نه تنها ترکیب‌بندی ممکنی که دستگاه مکانیکی مفروض می‌تواند داشته باشد، بلکه سرعت‌های نقطه‌های مختلف تشکیل‌دهندهٔ آن هم معین می‌شود. یکی از ساده‌ترین مثال‌ها را می‌آوریم. نقطه‌ای را در نظر می‌گیریم که بتواند روی محیط دایره با سرعتی دل‌خواه حرکت کند. هر حالت این دستگاه با دو داده معین می‌شود: جای نقطه در روی محیط دایره و سرعت آن در لحظهٔ مفروض. روشن است، حالت خمینه‌ای (فضای فازی) دستگاه مکانیکی مفروض، عبارت است از استوانهٔ بی‌پایان (حاصل‌ضرب محیط دایره در خط راست)

بعد فضای فازی، همراه با افزوده شدن درجه‌های آزادی دستگاه مفروض، بالا می‌رود. بسیاری از ویژگی‌های دینامیکی این یا آن دستگاه مکانیکی، بیان خود را در ویژگی‌های توپولوژیک فضای فازی آن پیدا می‌کنند. به‌عنوان مثال، هر حرکت تناوبی دستگاه مفروض، متناظر است با منحنی بسته‌ای در فضای فازی آن.

بررسی فضاهای فازی دستگاه‌های دینامیکی، در بسیاری از مسأله‌های گوناگون مکانیک، فیزیک و اخترشناسی (مکانیک آسمانی، گیتی‌شناسی)، توجه ریاضی‌دانان را به‌توپولوژی خمینه‌های چندبعدی جلب کرده است. به‌ویژه، ***پوانکاره***، ریاضی‌دان بزرگ فرانسوی در دههٔ نود سدهٔ نوزدهم، در رابطه با این مسأله‌ها، به‌طور منظم ساختن خمینه‌های توپولوژیک را در بررسی‌های خود قرار داد و برای این منظور از روش ترکیباتی (یا به‌قول برخی از ریاضی‌دانان امروز ایرانی، روش ترکیبیاتی) استفاده کرد که تا زمان ما، یکی از روش‌های اساسی توپولوژی است.

۴. روش ترکیبی

از نظر تاریخی، نخستین قضیهٔ مربوط به توپولوژی، قضیه یا دستور ***اولر*** است (که به‌ظاهر، ***دکارت*** هم آن را می‌شناخت). این قضیه را می‌آوریم. سطح یک چندوجهی کوژ را درنظر

می‌گیریم. α_0 را تعداد رأس‌های آن، α_1 را تعداد یال‌های آن و α_2 را تعداد وجه‌های آن فرض می‌کنیم. در این صورت، این رابطه، که **دستور *اولر*** نامیده می‌شود، برقرار است:

$$\alpha_0 - \alpha_1 + \alpha_2 = 2 \qquad (1)$$

این قضیه هندسی به‌این مناسبت به توپولوژی مربوط می‌شود که این دستور، برای هر چندوجهی کوژ با هر تبدیل دل‌خواه توپولوژیک درست است. ضمن چنین تبدیل‌هایی، در حالت کلی، یال‌ها از صورت خط راست بیرون می‌روند، وجه‌ها به‌صورتی غیرمسطح درمی‌آیند و رویهٔ چندوجهی به‌سطحی خمیده تبدیل می‌شود؛ باوجود این، برابری (۱) برای تعداد رأس‌ها، تعداد یال‌ها (که اکنون به‌صورت خمیده درآمده‌اند) و تعداد وجه‌ها به‌قوت خود باقی می‌ماند. مهم‌ترین حالت، که همهٔ وجه‌ها مثلثی‌شکل باشند که در این صورت، با به‌اصطلاح ***مثلث‌بندی*** سرو کار داریم (تقسیم سطح به‌مثلث‌های با ضلع‌های راست یا منحنی). اثبات دستور برای هر چندوجهی دل‌خواه را به‌سادگی می‌توان به‌این حالت منجر کرد: کافی است هر وجه را به‌مثلث‌هایی تبدیل کنیم (که از جمله می‌توان از هر رأس وجه، همهٔ قطرهای آن را رسم کرد). بنابراین، می‌توان خود را به‌بررسی سطح مثلث‌بندی شده محدود می‌کنیم. روش ترکیبی در توپولوژی سطح، عبارت است از بررسی این سطح به‌صورت مثلث‌بندی آن؛ در ضمن، تنها به‌چنان ویژگی‌هایی از سطح مثلث‌بندی شده علاقه‌مندیم که ربطی به‌شیوهٔ تصادفی نوع مثلث‌بندی نداشته باشندو برای همه گونه‌های مثلث بندی سطح مفروض مشترک باشند و، در نتیجه، ویژگی‌های خود سطح را بیان کنند.

دستور *اولر*، ما را به‌یکی از این ویژگی‌ها می‌رساند و، به‌همین دلیل، اندکی مفصل‌تر دربارهٔ آن صحبت می‌کنیم. سمت چپ دستور اولر، یعنی $\alpha_0 - \alpha_1 + \alpha_2$ (که در آن، α_0 تعداد رأس‌ها، α_1 تعداد یال‌ها α_2 تعداد مثلث‌ها در مثلث‌بندی است)، **مشخصه (*یا مفسر*) *اولر*** این مثلث‌بندی نامیده می‌شود. قضیهٔ اولر براین تأکید دارد که برای مثلث‌بندی هر سطح، که همسانریخت کره باشد، مشخصه *اولر* برابر است با دو. معلوم می‌شود برای هر سطحی (و نه تنها سطحی که همسانریخت کره باشد)، همهٔ مثلث‌بندی‌های این سطح دارای مشخصه *اولر* برابر هستند.

به‌سادگی می‌توان، مشخصه *اولر* را برای سطح‌های مختلف، پیدا کرد. قبل از همه، برای سطح استوانه‌ای، این مشخصه برابر صفر است. در واقع، اگر از مثلث‌بندی کره، دو مثلث

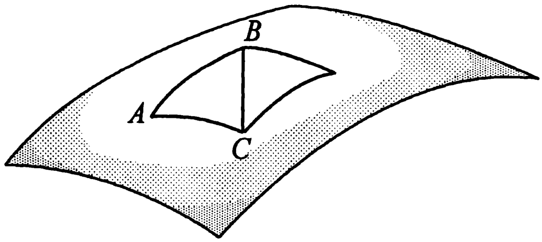
غیر مجاور را کنار بگذاریم، ولی مرزهای این مثلث‌ها را نگه داریم، به مثلث‌بندی سطحی می‌رسیم که با سطح جانبی استوانه همسانریخت است. در ضمن، تعداد راس‌ها و تعداد یال‌ها تغییر نمی‌کند، در حالی که تعداد مثلث‌ها، ۲ واحد کمتر می‌شود و بنابراین مشخصه *اولر* مثلث‌بندی جدید، برابر صفر خواهد شد. اکنون سطحی را انتخاب می‌کنیم که با کنار گذاشتن تعداد $2p$ از مثلث‌های مثلث‌بندی شده کره به‌دست آمده است، به‌نحوی که هیچ دو مثلثی که کنار گذاشته‌ایم، مجاور نباشند (یعنی نه رأس‌های مشترک داشته باشند و نه یال‌های مشترک)[1]. در این صورت، مشخصه *اولر*، به‌تعداد $2p$ کاهش می‌یابد. به‌سادگی می‌توان در این باره قانع شد که، اگر به‌هر زوج سوراخی را که در سطح کره به‌وجود آمده است، با لوله‌ای استوانه‌ای بچسبانیم، مشخصه *اولر* تغییری نمی‌کند. این، از آن جا ناشی می‌شود که، مشخصه *اولر* لوله‌هایی که وصل کرده‌ایم، همان‌طور که دیدیم برابر صفر است و در کناره‌های این لوله‌ها، تعداد رأس‌ها با تعداد یال‌ها برابر است. به‌این ترتیب، سطح دو رویهٔ بستهٔ با گونای p، دارای مفسر *اولر* برابر $2p - 2$ است (این حقیقت را، برای نخستین بار، ده ژُن کی‌یر، دریاسالار فرانسوی ثابت کرد).

ویژگی مهم دیگری از مثلث‌بندی را می‌آوریم که با به‌اصطلاح شرط ناوردایی توپولوژیک سازگار است (یعنی به‌شرطی که برای یک مثلث‌بندی سطح مفروض برقرار باشد، برای همهٔ دیگر مثلث‌بندی‌های همان سطح هم برقرار است). این ویژگی، **جهت‌پذیری** است. پیش از آن که این ویژگی را تنظیم کنیم، یادآوری می‌کنیم که هر مثلث را می‌توان جهت‌دار کرد که آن را، توجیه مثلث هم می‌گویند)، که به‌معنای دادن جهتی روی مرز مثلث که، بنابر آن، حرکتی روی محیط مثلث را مشخص کند و روشن باشد، وقتی در یک رأس قرار گرفته‌ایم، به‌سمت کدام رأس حرکت کنیم. هر یک از دو حالت ممکن در این باره، تعقیب رأس‌ها را، ضمن حرکت روی محیط معین می‌کند[2]. اکنون فرض می‌کنیم دو مثلث مجاور روی سطحی مفروض باشند، به‌نحوی که در یک ضلع مشترک باشند و، به‌جز این ضلع، نقطهٔ مشترک دیگری نداشته باشند (شکل ۱۷). دو مثلث جهت‌دار از این گونه

---

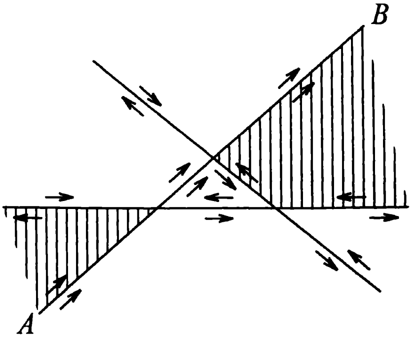
۱. برای این‌که چنین عملی ممکن باشد، تنها باید مثلث‌بندی خود را به‌اندازهٔ کافی «خُرد» بگیریم. این کار هم همیشه با تقسیم مثلث‌های هر مثلث‌بندی ممکن است.

۲. در ضمن به‌سادگی معلوم می‌شود که، دو ردیف از رأس‌ها، وقتی و تنها وقتی، یک جهت حرکت را ضمن دور زدن مثلث، معین می‌کنند که با جایگشتی «زوج»، یکی به‌دیگری منجر شود. به‌این ترتیب (ABC)، (BCA) و (CAB) یک جهت را، و (BAC)، (ACB) و (CBA) جهت دیگری را در مثلث ABC معین می‌کنند (دربارهٔ جایگشت‌های فرد و زوج *permutation*، به‌عنوان نمونه، بند ۳ از بخش بیستم را ببینید).



شکل ۱۷

(یعنی مثلث‌های مجاور) را هم‌آهنگ یا سازگار گویند، وقتی که روی ضلع مشترک مثلث‌ها، دو جهت مخالف هم پدید آید (روی صفحه یا هر سطح دو رویه، این موضوع به‌معنای آن است که ـ اگر به‌دو مثلث از یک سمت سطح نگاه کنیم ـ جهت هر دو مثلث یکی باشد، یعنی یا هر دو در جهت حرکت عقربه‌های ساعت و یا هر دو برخلاف حرکت عقربه‌ها باشند). مثلث‌بندی یک سطح مفروض بسته را جهت‌پذیر گوییم، وقتی جهت همهٔ مثلث‌های داخل این مثلث‌بندی را بتوان طوری انتخاب کرد که هر دو مثلث مجاور (یعنی دو مثلثی که در یک ضلع مشترک‌اند و نقطهٔ مشترک دیگری ندارند)، جهت‌هایی هم‌آهنگ داشته باشند. این حکم درست است: هر مثلث‌بندی سطح دو رویه جهت‌پذیر است، در حالی که هرگونه سطح یک رویه‌ای، جهت‌پذیر نیست. به‌همین مناسبت، سطح دو رویه را سطح جهت‌پذیر و سطح یک رویه را سطح جهت‌ناپذیر هم می‌گویند. اگر خواننده مثلث‌بندی دل‌خواهی از نوار موبیوس را در نظر بگیرد، بی‌دشواری، قانع می‌شود که جهت‌پذیر نیست. برای این که ساده‌ترین مثلث‌بندی صفحهٔ تصویری را به‌دست آوریم، باید روی آن، سه خط راست دل‌خواهی که از یک نقطه نمی‌گذرند، رسم کرد (شکل ۱۸). این خط‌های راست، از صفحه تصویری چهار مثلث جدا می‌کنند که یکی از آن‌ها در بخش محدودی از صفحه قرار می‌گیرد و سه تای دیگر، هر کدام شامل دو بخش‌اند و خط در بی‌نهایت آنها را دو بخش می‌کند. در شکل ۱۸، یکی از این سه مثلث که به بی‌نهایت دور می‌رود، هاشور خورده است. روی همین شکل دیده می‌شود که هرگونه تلاشی، برای جهت‌دار کردن هم‌آهنگ این مثلث‌ها، با عدم موفقیت روبه‌رو می‌شود. در حالت خاصی که روی شکل ۱۸، جهت‌ها را معین کرده‌ایم، به‌جای این که مجموع جبری مرزها برابر صفر شود (چیزی که در جهت‌گذاری هم‌آهنگ به‌دست می‌آید)، دو برابر خط راست AB را به‌دست می‌آوریم.



شکل ۱۸

مشخصه *اولر* و ویژگی جهت‌پذیری یا جهت‌ناپذیری سطح‌های بسته، دستگاه کامل ناورداهای توپولوژیک را، برای سطح‌های بسته به ما می‌دهند. مفهوم این بیان این است که دو سطح، وقتی و تنها وقتی همسانریخت (هومِئومورف) هستند که مثلث‌بندی این سطح‌ها: ۱) دارای یک مشخصه *اولر* باشند، ۲) هر دو جهت‌پذیر یا هر دو جهت‌ناپذیر باشند.

روش ترکیبی، به‌جز بررسی سطح‌ها (خمینه‌های دو بعدی)، برای بررسی خمینه‌های چند بعدی هم به‌کار می‌رود. ولی، برای نمونه، در حالت خمینه‌های سه بعدی، نقش مثلث‌بندی عادی، به‌عهدهٔ چهاروجهی‌ها خواهد بود که در این صورت، آن را مثلث‌بندی سه‌بعدی یا تقسیم خمینه به‌وسیلهٔ سادک‌ها گویند. مشخصه *اولر* در مثلث‌بندی سه‌بعدی، با عدد $\alpha_۰ - \alpha_۱ + \alpha_۲ - \alpha_۳$ به‌دست می‌آید که در آن α_i $(i = ۰ , ۱ , ۲ , ۳)$ به‌معنای تعداد عنصرهای iبعدی این مثلث‌بندی است ($\alpha_۰$ تعداد راس‌ها، $\alpha_۱$ تعداد یال‌ها، $\alpha_۲$ تعداد وجه‌های دوبعدی، $\alpha_۳$ تعداد چهاروجهی‌ها). وقتی تعداد بعدها از ۳ بیشتر باشد $(n>۳)$، خمینه به‌سادک‌های nبعدی، یعنی ساده‌ترین چندوجهی‌های کوژ nبعدی، تقسیم می‌شود، شبیه مثلث‌ها $(n=۲)$ و چهاروجهی‌ها $(n=۳)$. سادک‌هایی که تقسیم‌های خمینهٔ nبعدی را تشکیل داده‌اند، و وجه‌های آن‌ها، مثلث‌بندی nبعدی این خمینه را می‌سازند. در این جا هم می‌توان از مشخصه *اولر* صحبت کرد که برابر است با مجموع $\sum_{i=۰}^{n} (-۱)^i \alpha_i$، که در آن، α_i عبارت است از تعداد عنصرهای iبعدی موجود در مثلث‌بندی و i می‌تواند برابر عددهای از ۰ تا n باشد. در ضمن، مثل قبل، مشخصه *اولر* برای همهٔ مثلث‌بندی‌های خمینهٔ n بعدی مفروض (و همهٔ خمینه‌های همسانریخت با آن)، تنها یک مقدار را می‌دهد، یعنی یک ناوردای توپولوژیک است: ولی دربارهٔ دستگاه ناورداها (به‌این مفهوم که برای سطح،

مشخصه *اولر* و جهت‌پذیری انجام می‌دهد) برای خمینه‌های سه‌بعدی هم، با سطح آگاهی‌های امروزی، نمی‌توانیم چیزی بگوییم.

اهمیت روش ترکیبی در توپولوژی امروزی، بسیار زیاد است. این روش به ما امکان می‌دهد، از برخی شیوه‌های جبری، برای حل مسأله‌های توپولوژی استفاده کنیم. این استفادهٔ از مفهوم‌های جبری را در بالا هم، به‌صورت سادهٔ خود دیدیم، آن جا که از مجموع جبری مرزهای مثلث‌های جهت‌دار در مثلث‌بندی صفحهٔ تصویری یاد کردیم. در واقع، وقتی مثلث جهت‌دار باشد، یعنی جهتی برای دور زدن آن انتخاب کرده باشیم، آن وقت به‌ناچار بحث مجموعهٔ ضلع‌های آن پیش می‌آید و برای هر ضلع هم جهت معینی پیدا می‌شود، همان جهتی که دور زدن مثلث را میسر می‌کند.

اکنون همهٔ مثلث‌های T_i^2 $(i = 1, 2, \dots, \alpha_2)$ را، که در مثلث‌بندی مفروضی از یک سطح داخل شده‌اند، در نظر می‌گیریم. برای هر یک از این مثلث‌ها، می‌توان دو جهت را در نظر گرفت؛ مثلث T_i^2 را با یکی از دو جهت ممکن با t_i^2 و همان مثلث را با جهت دیگر (مخالف جهت قبلی) با $-t_i^2$ نشان می‌دهیم. به‌همین‌ترتیب، هر یک از عنصرهای یک‌بعدی (یال) T_k^1 $(k = 1, 2, \dots, \alpha_1)$ را که در مثلث‌بندی مفروض وارد شده است، می‌توان جهت‌دار کرد، یعنی به‌آن یکی از دو جهت ممکن را داد. پاره‌خط راست T_k^1 را در یکی از این جهت‌ها با t_k^1 و در جهت دیگر با $-t_k^1$ نشان می‌دهیم. در این صورت، اگر ضلع‌های مثلث T_i^2، پاره‌خط‌های راست T_1^1، T_2^1، T_3^1 باشند، آن وقت مرز مثلث جهت‌دار t_i^2، مجموعهٔ همین ضلع‌هاست، البته باتوجه به‌جهتی که دارند، یعنی مرز تشکیل شده است از پاره‌خط‌های راست جهت‌دار $\varepsilon_1 t_1^1$، $\varepsilon_2 t_2^1$، $\varepsilon_3 t_3^1$؛ در این‌جا $\varepsilon_i = 1$، به‌شرطی که این جهت برای یال T_i^1، منطبق برجهت خاص t_i^1 باشد، و در حالت عکس $\varepsilon_i = -1$. مرز t_i^2 را با Δt_i^2 نشان می‌دهند. همان‌طور که می‌بینیم $\Delta t_i^2 = \sum \varepsilon_k t_k^1$؛ در ضمن، این مجموع را می‌توان برای همهٔ یال‌های مثلث‌بندی در نظر گرفت، با این شرط که ضریب ε_k را برای پاره‌خط‌های راستی که در مرز مثلث بندی t_i^2 شرکت ندارند، برابر صفر بگیریم.

طبیعی به‌نظر می‌رسد، که به‌طورکلی مجموع $x^1 = \sum a_k t_k^1$ را که برای همهٔ یال‌های مثلث‌بندی گسترش داده شده است، بررسی کنیم[1]. مفهوم هندسی چنین مجموع‌هایی خیلی ساده است: هر جملهٔ مجموع، پاره‌خط راستی است که در مثلث‌بندی ما شرکت دارد

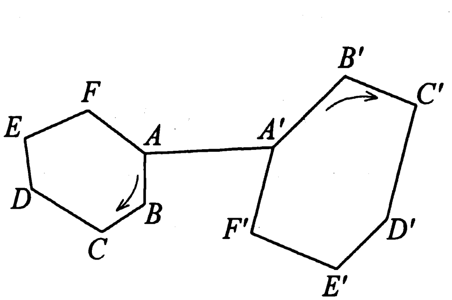
---

۱. ضریب‌های a_k را عددهایی درست به‌حساب می‌آوریم.

که با جهت معین و با ضریب معینی در نظر گرفته شده است («درجهٔ تعدد» یا «چندگانگی» معین). این مجموع جبری، بیان مسیری از پاره‌خط‌های راست است که در آن، هر پاره‌خط راست به‌تعداد ضریب مربوط به‌آن، شرکت کرده است. اگر در شکل، ابتدا از چندضلعی $ABCDEF$ (شکل ۱۹) در جهتی که نشان داده شده است، عبور کنیم، سپس از راه پاره‌خط راست AA'، $A'B'C'D'E'F'$ را در جهت پیکان دور بزنیم، بعد دوباره از راه خط راست $A'A$ به چندضلعی $ABCDEF$ برگردیم و در همان جهت قبلی حرکت کنیم، آن وقت مجموعی به‌دست می‌آوریم که در آن پاره‌خط‌های راست AB، BC، CD، DE، EF، FA ضریبی برابر ۲، پاره‌خط‌های راست $A'B'$، $B'C'$، $C'D'$، $D'E'$، $E'F'$، $F'A'$ ضریبی برابر ۱ دارند، ولی پاره‌خط راست AA' وارد مجموع نشده است (AA' ضریبی برابر صفر دارد، زیرا از آن دوبار و در دو جهت مخالف هم عبور کرده‌ایم).

مجموع‌های به‌صورت $x^1 = \sum a_k t_k^1$ را **زنجیرهٔ یک‌بعدی** مثلث‌بندی مفروض می‌نامند. از دیدگاه جبری، این مجموع‌ها، عبارت‌اند از صورت‌های خطی (چندجمله‌ای همگن درجهٔ اول)؛ آن‌ها را می‌توان باهم جمع و یا از هم کم و یا در عدد درست دل‌خواهی ضرب کرد (بنا بر قانون‌های معمول جبر). بین زنجیره‌های یک‌بعدی، به‌ویژه **دورهای یک‌بعدی** اهمیت دارند. از نظر هندسی، این زنجیره‌ها متناظر با مسیرهای بسته‌اند (آن طور که برای نمونه، در شکل ۱۹ داشتیم).

برای تعریف جبری دور، قرار می‌گذاریم از دو رأس پاره‌خط جهت‌دار $\overrightarrow{AB}$، راس پایانی B که در مرز پاره‌خط راست $\overrightarrow{AB}$ قرار دارد، با علامت مثبت باشد (با ضریب ۱+)، و راس آغازین A، با علامت منفی (با ضریب ۱−). در این صورت مرز پاره‌خط راست $\overrightarrow{AB}$ را می‌توان به‌صورت $\Delta(\overrightarrow{AB}) = B - A$ نوشت.



شکل ۱۹

با پذیرش این شرط، بلافاصله روشن می‌شود که، مجموع پاره‌خط‌های راست مرزی که $\overrightarrow{AB}$، $\overrightarrow{BC}$، $\overrightarrow{CD}$، ...، $\overrightarrow{FA}$ که مسیر بسته‌ای را (به معنای عادی آن) تشکیل می‌دهند، برابر صفر است. و این به معنای آن‌است که زنجیرهٔ یک‌بعدی $z^1 = \sum_k a_k t_k^1$ یک دور است اگر مجموع مرزهای جمله‌های آن، یعنی مجموع $\sum_k a_k \Delta t_k^1$ برابر صفر باشد. به سادگی می‌توان ثابت کرد که، مجموع دو دور، خود یک دور است. از ضرب یک دور در یک عدد درست، دوباره یک دور به‌دست می‌آید. این وضع اجازه می‌دهد دربارهٔ ترکیب‌های خطی دورهای z_1^1، z_2^1، ...، z_s^1، یعنی دربارهٔ دورهای به‌صورت $z = \sum_{r=1}^{s} c_r z_r^1$ صحبت کنیم که در آن c_rها، عددهایی درست‌اند.

شبیه مفهوم زنجیرهٔ یک‌بعدی مثلث‌بندی مفروض، می‌توان دربارهٔ زنجیره‌های دوبعدی این مثلث‌بندی، یعنی دربارهٔ عبارت‌هایی به‌صورت $x^2 = \sum_i a_i t_i^2$ صحبت کرد که در آن، t_i^2 مثلث‌های جهت‌دار مثلث‌بندیِ مفروض‌اند. چون مرز هر مثلث جهت‌دار، یک دور یک‌بعدی است، بنابراین زنجیرهٔ $\sum_i a_i \Delta t_i^2$ هم یک دور است. این دور را مرز Δx^2 از زنجیرهٔ $x^2 = \sum_i a_i t_i^2$ به‌حساب می‌آوریم.

مفهوم **مرز زنجیره**، اجازه می‌دهد مفهوم مانستگی (یا ***همتایی***، هومولوژی= *homology*) را تنظیم کنیم: دور یک‌بعدی z^1 از مثلث‌بندی مفروض، در این مثلث‌بندی، مانسته با (یا **همتای**) **صفر** نامیده می‌شود، به‌شرطی که مرز یک زنجیرهٔ دوبعدی این مثلث‌بندی باشد. در هر مثلث‌بندی از سطح بستهٔ کوژ و به‌طور کلی، روی هر سطح همسانریخت کره، هر دور یک‌بعدی، مانسته با صفر است؛ معنای هندسی این وضع روشن است: هر چندضلعی بسته روی یک سطح کوژ، مرز تکه‌ای از این سطح است. در چنبره این طور نیست: نصف‌النهار چنبره، همچنین استوای آن، مرز تکه‌ای از سطح چنبره نیست. اگر یک مثلث‌بندی روی چنبره در نظر بگیریم، در آن دوری شبیه نصف‌النهار و استوای چنبره به‌دست می‌آید؛ در ضمن این دور مانسته با صفر نیست.

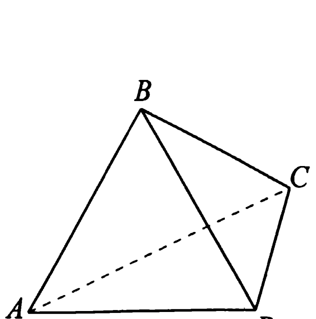
مثلث‌بندی صفحهٔ تصویری، که پیش از این هم از آن یاد کردیم، پدیدهٔ به‌کلی تازه‌ای در برابر ما می‌گذارد. اگر خط راستی، مثل خط راست AB را، (شکل ۱۸ را ببینید)، به‌عنوان دور این مثلث‌بندی در نظر بگیریم، آن وقت این دور مانسته با صفر نمی‌شود. ولی اگر همین خط راست را، با ضریب ۲ بگیریم، آن وقت مانسته با صفر می‌شود. بنابراین، استفاده از ضریب‌هایی غیر از ۱± برای تعیین زنجیره‌ها، که در آغاز صوری و غیر لازم به‌نظر

می‌رسد، امکان می‌دهد به‌ویژگی‌های هندسی مهمی از سطح‌ها و به‌طور کلی خمینه‌ها پی ببریم. در این حالت، این به‌اصطلاح ویژگی تاب (*torsion*) است که به معنی وجود زنجیره‌هایی است که در خمینهٔ مفروض مانسته با صفر نیست (هیچ تکه‌ای از سطح را محدود نمی‌کند)، ولی می‌تواند مانسته با صفر شود، به‌شرطی که ضریب‌های درست عددی به‌آن‌ها بدهیم.

در رابطه با آن چه گفتیم، مفهوم بسیار مهم ***استقلال مانستگیی دورها*** را مطرح می‌کنیم. دورهای z_1 ، .... ، z_s را در مثلث‌بندی مفروض، مستقلِ مانسته گویند، وقتی که هیچ‌یک از ترکیب‌های خطی $\Sigma c_i z_i$، که در آن دست‌کم یکی از ضریب‌های c_i مخالف صفر است، در این مثلث‌بندی مانسته با صفر نباشد. به‌عنوان نمونه‌ای از دورهای مستقل مانسته، روی چنبره، می‌توان یک استوا و یک نصف‌النهار چنبره را، به‌عنوان دورهایی از مثلث‌بندی چنبره، در نظر گرفت.

با مفهوم‌های اساسی تمامی توپولوژی ترکیبی ـ مفهوم‌های مرز، دور، مانستگی ـ برای خمینه‌های یک‌بعدی آشنا شدیم، ولی این مفهوم‌ها برای هر تعداد بعد هم وجود دارند. از جمله زنجیرهٔ بعدی $z^2 = \Sigma a_i t_i^2$، وقتی دور نامیده می‌شود که مرز آن $\Delta z^2 = \Sigma a_i \Delta t_i^2$ برابر صفر باشد. زنجیره سه‌بعدی، عبارتی به‌صورت $x^3 = \Sigma a_h t_h^3$ است که در آن، t_h^3 سادک سه‌بعدی جهت‌دار است (چهاروجهی).

مثل حالت مثلث، جهت‌دار کردن سادک سه‌بعدی (چهاروجهی) هم، با تعریفی برای ردیف راس‌های آن داده می‌شود؛ در ضمن، دو ردیفی از رأس‌ها، که با تبدیلی زوج به‌یکدیگر منجر شوند، یک و تنها یک ردیف را مشخص می‌کنند. مرز سادک سه‌بعدی جهت‌دار $t^3 = (ABCD)$، عبارت است از یک زنجیره (دور) دوبعدی (شکل ۲۰)



شکل ۲۰

$$\Delta t^3 = (BCD) - (ACD) + (ABD) - (ABC)$$

مرز زنجیره سه‌بعدی، به‌عنوان مجموع مرزهای سادک‌های آن، با همان ضریب‌هایی که این سادک‌ها در زنجیرهٔ مفروض وارد شده‌اند، معین می‌شود. به‌سادگی می‌توان تحقیق کرد که مرز هر زنجیرهٔ سه‌بعدی را، یک دور دوبعدی تشکیل می‌دهد (کافی است این حکم را برای مرز یک سادک سه‌بعدی ثابت کنیم). دور دوبعدی را در خمینهٔ مفروض، مانسته با صفر گوییم، وقتی که مرز یک زنجیرهٔ سه‌بعدی از این خمینه باشد. و به‌همین ترتیب برای دورهای بابعدهای بیشتر. یادآوری می‌کنیم، ازتعریفی‌که برای‌مثلث‌بندی‌های‌جهت‌پذیر وجهت‌ناپذیر آوریم[1]، بلافاصله نتیجه می‌شود که، در هر مثلث‌بندی جهت‌پذیر، دورهایی پیدا می‌شود (در حالت سطح ـ دوبعدی) که مخالف صفرند، ولی در مثلث‌بندی‌های جهت‌ناپذیر، چنین دورهایی وجود ندارند؛ این نتیجه‌گیری را می‌توان بلافاصله برای هر بعد تعمیم داد.

این مفهوم‌ها، امکان می‌دهند، مرتبه‌های همبندی یک‌بعدی، دوبعدی و...، در خمینه‌های مفروض با هر بعد معین کنیم. بیشترین تعداد دورهایی که در مثلث‌بندی خمینه مستقل مانسته‌اند، **مرتبهٔ همبندی** یا **عدد بتی** (متناظر با آن بعد) نامیده می‌شود.

عدد بتی یک‌بعدی، برای سطح بستهٔ جهت‌دار از خانوادهٔ p، برابر است با $2p$ (یعنی مرتبهٔ همبندی سطح، آن‌طور که در بند ۲ تعریف شد). عدد یک‌بعدی بتی، برای صفحهٔ تصویری برابر صفر است (در آن‌جا، هر دوری که به بی‌نهایت نرفته باشد، بخشی از صفحه را محدود می‌کند، یعنی مانسته با صفر است؛ و دوری که به بی‌نهایت رفته باشد، مثل خط راست تصویری، اگر آن را دوبار به‌حساب آوریم، مانسته با صفر است). **عدد دوبعدی بتی**، برای هر سطح جهت‌ناپذیر، برابر صفر است (روی چنین سطحی، یک دور دوبعدی مخالف صفر هم پیدا نمی‌شود).

عدد دوبعدی بتی، برای هر سطح جهت‌پذیر برابر است با ۱. در واقع، اگر همهٔ مثلث‌های یک مثلث‌بندی را، که روی سطح جهت‌پذیر قرار دارد، جهت‌دار کنیم، دور به دست می‌آید (به اصطلاح، دور اصلی سطح) به‌سادگی روشن می‌شود، هر دور دوبعدی، از ضرب دور اصلی در عددی درست به دست می‌آید. این نتیجه‌گیری‌ها را می‌توان برای خمینه‌های n بعدی تعمیم داد.و یادآوری می‌کنیم، عدد، بتی «صفر بعدی» در خمینه‌های

---

۱. تعریفی که در این جا برای مثلث‌بندی‌های سطح آوردیم، برای مثلث‌بندی‌های خمینه‌های با هر تعداد بعد هم، درست است.

همبند (یعنی وقتی که به‌قطعه‌های جداگانه تقسیم نشده باشد) را برابر ۱ به‌حساب می‌آورند.

عدد بتی برای بعدهای مختلف، به‌مشخصه *اولر* خمینه و با استفاده از دستور جالبی که به‌وسیلهٔ **پوانکاره** با تعمیم قضیهٔ *اولر* داده شد، مربوط می‌شود. این دستور که به‌دستور *اولر*- **پوانکاره** مشهور است، ساده و به‌این صورت است:

$$\sum_{r=0}^{n}(-1)^r \alpha_r = \sum_{r=0}^{n}(-1)^r p_r$$

در این‌جا، در سمت چپ برابری، مشخصه *اولر* از یک مثلث‌بندی خمینه مفروض قرار دارد و عدد p_r در سمت راست برابری، عدد **بتی** برای بعدهای مختلف r از این خمینه است. در حالت خاص برای سطح جهت‌پذیر، همان‌طور که پیش از این هم دیدیم، داریم: $p_0 = p_2 = 1$ ، $p_1 = 2p$ (p، گونای سطح است). و این قضیه *اولر* را برای سطح جهت‌پذیر به ما می‌دهد:

$$\alpha_0 - \alpha_1 + \alpha_2 = 2 - 2p$$

۵. میدان‌های برداری

ساده‌ترین معادلهٔ دیفرانسیلی

$$\frac{dy}{dx} = F(x, y) \qquad (2)$$

که در حوزهٔ سطح مفروض G داده شده است، در نظر می‌گیریم. مفهوم هندسی این معادله چنین است: در هر نقطهٔ (x, y) از حوزهٔ G، جهتی معین شده است که ضریب زاویهٔ آن (شیب آن) برابر $F(x, y)$ است، که در آن $F(x, y)$، تابعی پیوسته از نقطهٔ (x, y) است. می‌گویند در حوزهٔ G، میدان پیوستهٔ جهت‌ها داده شده است؛ می‌توانیم با انتخاب بردار به‌طول واحد در هر یک از جهت‌های مفروض، به‌سادگی آن را به‌میدان برداری پیوسته تبدیل کنیم. مسألهٔ انتگرال‌گیری معادلهٔ دیفرانسیلی (۲) به‌این معناست که حوزهٔ سطح مفروض را (در صورت امکان)، به‌منحنی‌هایی که دو به‌دو غیرمتقاطع‌اند («منحنی‌های

انتگرالی» معادله)، به‌نحوی تجزیه کنیم که در هر نقطهٔ حوزه، جهتی برآن داده شود که جهت مماس بر تنها منحنی انتگرالی باشد که از این نقطه می‌گذرد.

برای نمونه، این معادله را در نظر می‌گیریم:

$$\frac{dy}{dx}=\frac{y}{x}$$

در هر نقطهٔ $M(x\ ,\ y)$ از صفحه، جهتی نسبت داده می‌شود که همان جهت نیم‌خط راست $\overrightarrow{OM}$ است (O را مبدأ مختصات گرفته‌ایم). منحنی‌های انتگرالی، خط‌های راستی هستند که از نقطهٔ O می‌گذرند. از هر نقطهٔ صفحه، به‌جز نقطهٔ O، تنها یک منحنی انتگرالی می‌گذرد. نقطهٔ O، نقطهٔ تکین برای معادلهٔ مفروض است (به اصطلاح، «نقطهٔ گرهی») که همهٔ منحنی‌های انتگرالی از آن می‌گذرند.

اگر این معادلهٔ دیفرانسیلی را در نظر بگیریم:

$$\frac{dy}{dx}=-\frac{x}{y}$$

آن وقت می‌بینیم که، برای هر نقطهٔ $M(x\ ,\ y)$، به‌جز نقطهٔ O، جهتی نسبت داده می‌شود که بر $\overrightarrow{OM}$ عمود است. منحنی‌های انتگرالی در این حالت، محیط‌های دایره‌هایی به‌مرکز نقطهٔ O را تشکیل می‌دهند؛ خود نقطهٔ O هم نقطهٔ تکینی از معادلهٔ دیفرانسیلی است، ولی نقطهٔ تکینی، که به‌کلی از گونهٔ دیگری است. نقطهٔ O، یک «گره» نیست و به اصطلاح یک «مرکز» است. نقطه‌های تکین، از گونه‌های دیگری هم وجود دارند (بند ۹، بخش پنجم از جلد دوم را ببینید) که نمونه‌هایی از آن‌ها را در شکل‌های ۲۱ و ۲۲ نشان داده‌ایم. برای معادلهٔ دیفرانسیلی $\frac{dy}{dx}=\frac{y}{x}$، منحنی‌های انتگرالی بسته وجود ندارد. برعکس، معادلهٔ دیفرانسیلی $\frac{dy}{dx}=-\frac{x}{y}$، تنها دارای منحنی‌های انتگرالی بسته است. ممکن است منحنی‌های انتگرالی به‌صورت حلزونی، دور نقطهٔ تکین بپیچند، که در این صورت، نقطهٔ تکین را **کانون** می‌گویند.

در بین حالت‌های مختلفی که وجود دارد، حالتی بسیار مهم است که دور حدی نامیده می‌شود و آن عبارت است از منحنی انتگرالی بسته‌ای که دیگر منحنی‌های انتگرالی، به‌صورت حلزونی، دور آن می‌پیچند. حالت‌های بسیار دیگری هم از وضع استقرار منحنی‌های انتگرالی نسبت به‌هم و یا وضع استقرار آن‌ها نسبت به‌نقطهٔ تکین وجود دارد.

همهٔ مسأله‌های مربوط به‌شکل و وضع استقرار منحنی‌های انتگرالی یک معادلهٔ دیفرانسیلی، هم‌چنین تعداد، خصلت و استقرار نقطه‌های تکین نسبت به‌هم، به نظریهٔ کیفی معادله‌های دیفرانسیلی، مربوط می‌شود. همان‌طور که از نام آن پیدا است، نظریهٔ کیفی معادله‌های دیفرانسیلی، از انتگرال‌گیری مستقیم معادلهٔ دیفرانسیلی «به‌صورت محدود»، یا روش‌های تقریبی و انتگرال‌گیری عددی صرف‌نظر می‌کند. موضوع اساسی نظریهٔ کیفی، در واقع، توپولوژی میدان جهت‌ها و دستگاه‌های منحنی‌های انتگرالی معادلهٔ دیفرانسیلی مفروض است.

دیدگاه کیفی به معادلهٔ دیفرانسیلی به معنای پرداختن به‌مسأله‌هایی از نوع وجود منحنی‌های انتگرالی بسته است؛ به ویژه همهٔ مسأله‌های مربوط به تعداد و پیدایی دورهای حدی است که، در درجهٔ نخست، از جانب مسأله‌های مکانیک، فیزیک و صنعت مطرح می‌شود. این مسأله‌ها، برای نخستین بار، ضمن بررسی‌های ***پوانکاره*** در مکانیک آسمانی و کیهان‌زایی پدید آمد و، همان‌طور که دیدیم، بهانه‌ای برای جست‌وجوهای توپولوژیک هندسه‌دان فرانسوی شد. بعدها، مسأله‌های مربوط به نظریهٔ توپولوژیک معادله‌های دیفرانسیلی در مرکز توجه پژوهشگران مشهور شوروی قرار گرفت که روی نظریهٔ نوسان و رادیوتکنیک کار می‌کردند، منظورم مکتب ***لئونیدایساآنکوویچ ماندل شتام*** و ***آلکساندر آلکساندرویچ آندرونوف*** است که یکی از مرکزهای عمدهٔ کار روی نظریهٔ کیفی توپولوژیک معادله‌های دیفرانسیلی بود. مرکز دیگر کار در زمینهٔ نظریهٔ کیفی معادله‌های دیفرانسیلی و یا اندازهٔ زیادی با روش‌های توپولوژیک را ***و.و ستهپانوف*** و ***و.و.نرتس کوف*** در دانشگاه مسکو تشکیل دادند. در کارهای ریاضی‌دانان لنین‌گراد، سِوردلوسک و تازان هم که روی نظریهٔ کیفی معادله‌های دیفرانسیلی بررسی می‌کردند، نقش اصلی به‌عهدهٔ روش‌های توپولوژیک بود.

نظریهٔ معادله‌های دیفرانسیلی منجر به‌مطالعهٔ برخی میدان‌ها، نه تنها روی صفحه، بلکه در ضمن روی خمینه‌های چندبعدی شد؛ ساده‌ترین دستگاه‌های معادله‌های دیفرانسیلی، تعبیر هندسی خود را به‌عنوان میدان جهت‌ها در فضاهای اقلیدسی چندبعدی پیدا کردند. انتگرال‌های اولیه معادله، به‌معنای آن است که از بین همهٔ منحنی‌های انتگرالی، آن‌هایی را جدا کنیم که روی خمینه‌ای قرار دارند که با انتگرال‌های اولیه معین می‌شوند. هر دستگاه دینامیکی (به‌مفهوم کلاسیک این واژه)، در حالت کلی، منجر به‌خمینهٔ چندبعدی حالت‌های ممکن آن می‌شود (بند ۳ را ببینید) و به‌دستگاهی از معادله‌های دیفرانسیلی می‌رسد که

منحنی‌های انتگرالی آن، ضمن پر کردن فضای فازی مفروض، حرکت‌های ممکن دستگاه مفروض را معرفی می‌کنند. هر یک از این حرکت‌ها، به‌وسیلهٔ این یا آن شرط‌های اولیه معین می‌شود. بنابراین، موضوع اصلی بررسی در این حالت، عبارت است از میدان جهت‌ها و دستگاه مسیرها در خمینهٔ مفروض. آن چه در سال‌های اخیر در این زمینه انجام گرفته است، موجب پیشرفت نظریهٔ کیفی معادله‌های دیفرانسیلی و چشم‌انداز گسترده آن و بنابراین، پیشرفت توپولوژی، به‌عنوان اساس این نظریه، شده است. به‌ویژه، زمینه‌های مربوط به‌مکانیک، فیزیک و اخترشناسی، توپولوژی را بی‌اندازه رشد داده است. آن‌هایی که بخواهند، مسأله‌های توپولوژیک در نظریهٔ معادله‌های دیفرانسیلی و جنبه‌های فیزیکی و تکنیکی آن آشنا شوند، می‌توانند از جمله، از کتاب مشهور «نظریهٔ نوسان‌ها» نوشتهٔ *آندرونوف* و هایکین استفاده کنند.

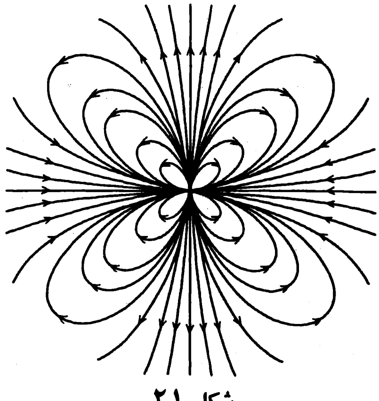
برای نمونه، مسألهٔ مشخصی را می‌آوریم که در نظریهٔ میدان‌های برداری درخمینه‌های چندبعدی حل شده است و مربوط به‌مسأله جبری تعداد نقطه‌های تکین این میدان برداری است.

خمینه‌ای هموار را در نظر می‌گیریم. برای سادگی‌کار، سطح بستهٔ هموار را فرض می‌کنیم. فرض می‌کنیم، در هر نقطهٔ این خمینه، بردار مماس برآن را داده باشند که هم طول و هم جهت آن، به‌طور پیوسته، بستگی به‌نقطه دارد. نقطه‌های تکین این میدان برداری، نقطه‌هایی از خمینه‌اند که در آن‌ها، بردار مربوط ، برابر صفر باشد، یعنی نتوان برای آن جهتی معین کرد. فرض می‌کنیم، هر یک از این نقطه‌ها، نقطه‌ای تنها باشد. در حالت خمینهٔ بسته، این مطلب به‌معنای آن است که تنها تعداد محدودی نقطهٔ تکین وجود دارد (در غیر این صورت، بخشی از این نقطه‌ها در نزدیکی نقطهٔ حدی متراکم می‌شود، که به‌دلیل پیوستگی میدان، خود یک نقطهٔ تکین است و، بنابراین نمی‌تواند تنها باشد).

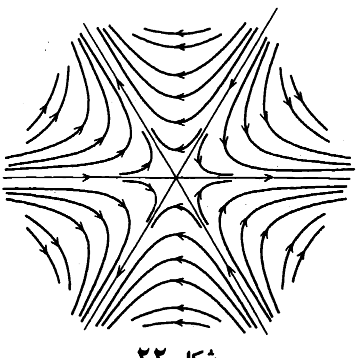
برای نقطهٔ تکین تنها، می‌توان زیرنشان (اندیس یا شاخص) آن را تعریف کرد، که مفهوم عادی آن، شبیه مفهوم تکرار ریشه‌ها در معادله‌های جبری است. برای تعیین زیرنشان، نقطهٔ تکین مفروض را به وسیلهٔ منحنی بسته C (که روی صفحه، می‌تواند محیط یک دایره باشد)، برای «تنهاکردن آن»، احاطه می‌کنیم؛ این منحنی باید چنان باشد که از هیچ نقطهٔ تکینی نگذرد و در درون آن، تنها همان یک نقطه تکین (که آن را بررسی می‌کنیم) قرار داشته باشد. در همهٔ نقطه‌های منحنی C، جهت میدان، به‌صورتی یک‌ارزشی معین است. برای سادگی کار، بخش درونی منحنی C را که در همسایگی نقطهٔ تکین قرار دارد، تخت (یا مسطح)

می‌گیریم (در حالت کلی، می‌توان این همسایگی را همراه با میدان روی آن، روی صفحه نگاشت). با دور زدن منحنی در جهت مثبت، زاویه‌ای که جهت میدان با جهت ثابتی می‌سازد، وقتی به مقدار آغازین خود برمی‌گردد، در مسیر خود با جمله‌ای به صورت $2k\pi$ رشد می‌کند، که در آن، k عددی درست و مشخص است. همین عدد را «زیرنشان» **نقطهٔ تکین** و همچنین **عدد چرخش** میدان دور منحنی C می‌نامند. یادآوری می‌کنیم، این عدد به انتخاب نوع منحنی بسته‌ای که نقطهٔ تکین مفروض را جدا کرده است، بستگی ندارد. روی شکل‌های ۲۱ و ۲۲، نقطه‌های تکین متناظر با «زیرنشان‌های» ۳+ و ۲− نشان داده شده است. به‌همین‌ترتیب، البته به‌صورتی پیچیده‌تر، می‌توان زیرنشان نقطه‌های تکین را در میدان برداری (جهت‌ها)، برای خمینه‌های nبعدی، به‌ازای $n>2$، به‌دست آورد. معلوم شد، این قضیه که در سال ۱۹۲۶ میلادی و به‌وسیلهٔ **هوپف** [*Hopf*] ریاضی‌دان آلمانی ثابت شد، درست است: اگر برای خمینهٔ مفروض با میدان برداری پیوسته، تعداد محدودی نقطه‌های تکین وجود داشته باشد، آن وقت مجموع زیرنشان‌ها، یا آن‌طورکه می‌گویند، **عدد جبری** این نقطه‌های تکین، بستگی به‌میدان ندارد و همیشه برابر است با **مشخصه *اولر*** خمینه.

از قضیهٔ **هوپف** نتیجه می‌شود: میدان‌های برداری بدون تکینی، تنها برای خمینه‌هایی ممکن است که مشخصهٔ ***اولر*** آن‌ها برابر صفر باشد؛ روی خمینه‌های اخیر هم، همیشه می‌توان میدان برداری بدون تکینی ساخت. بنابراین، از بین همهٔ سطح‌های بسته، تنها روی چنبره و روی به اصطلاح چنبره یک رویه (**بطری کلاین**)، می‌توان میدان برداری بدون تکینی ساخت.



شکل ۲۱



شکل ۲۲

نظریهٔ نگاشت‌های پیوستهٔ خمینه برخودش و نتیجهٔ ناشی از آن، یعنی وجود نقطه‌های بی‌حرکت (یا ثابت) ضمن این نگاشت، با نظریهٔ میدان‌های بُرداری، رابطهٔ تنگاتنگی دارد. نقطهٔ xرا، نقطهٔ بی‌حرکت در نگاشت مفروض f گویند، وقتی تصویر آن، ضمن این نگاشت بر خودش منطبق شود، یعنی وقتی

$$f(x)=x$$

برای این که خصلت این رابطه را روشن کنیم، ساده‌ترین حالت را در نظر می‌گیریم: حالت نگاشت پیوستهٔ f از قرص بستهٔ (یا سطح دایرهٔ) K برخودش. اگر هر نقطهٔ x از قرص بستهٔ K را به‌تصویر خودش $f(x)$ بپیوندیم، بردار $u_x=\overrightarrow{x\ ,\ f(x)}$ به‌دست می‌آید. این بردار، وقتی و تنها وقتی برابر صفر می‌شود که داشته باشیم $f(x)=x$، یعنی وقتی که x در نقطهٔ بی‌حرکت این نگاشت باشد. ثابت می‌کنیم، به‌واقع، چنین نقطه‌ای وجود دارد. برای این منظور، فرض می‌کنیم چنین نباشد و عدد چرخش میدان برداری مفروض را در طول محیط دایرهٔ C از قرص K معین می‌کنیم.

ضمن تغییر پیوستهٔ میدان، روشن است که عدد چرخش آن در طول محیط دایرهٔ C، تنها می‌تواند پیوسته تغییر کند. اما، از آنجا که عدد چرخش عددی درست است، باید مقدار ثابتی باشد. از این‌جا نتیجه می‌شود، عدد چرخش میدان در طول محیط دایرهٔ C برابر است با ۱. در واقع، از آن‌جا که هر نقطه‌ای که متعلق به قرص K باشد، به توی همین دایره نگاشته می‌شود، برای نقطهٔ x که روی محیط دایرهٔ C قرار دارد، بردار u_x (که بنابر فرض مخالف صفر است)، جهتی به درون دایره پیدا می‌کند و، بنابراین، با شعاع Ox، که آن را به‌عنوان برداری در جهت به سمت مرکز O در نظر می‌گیریم، زاویه‌ای حاده می‌سازد.

به‌تغییر جهت‌های همهٔ بردارهای u_x، برای نقطه‌های واقع بر محیط دایرهٔ C، به‌صورتی پیوسته، توجه می‌کنیم. این تغییر به‌این معناست که همهٔ این بردارها را به اندازهٔ زاویه‌های حادهٔ متناظر آن‌ها طوری تغییر بدهیم که به سمت مرکز O جهت بگیرند. همان‌طور که هم‌اکنون گفتیم، عدد چرخش میدان در طول محیط دایرهٔ C تغییر نمی‌کند. ولی نتیجهٔ تبدیل ما به‌این‌جا رسید که میدان اولیهٔ روی C به‌میدان بردارهای شعاعی تبدیل شد که به روشنی عدد چرخش ۱ را دارد. به این ترتیب، میدان اولیه هم در طول محیط دایرهٔ C، دارای عدد چرخش ۱ است.

باتوجه به پیوسته بودن میدان برداری اولیه، عدد چرخش آن در طول محیط‌های دو

دایره، که هر دو به مرکز O باشند و طول شعاع‌های آن‌ها اندکی متفاوت باشد، یک عدد را به ما می‌دهد[1]. به‌این‌ترتیب، عدد چرخش میدان در طول محیط هر دایره‌ای به مرکز O و واقع در قرص K، یک مقدار، یعنی همه‌جا عدد ۱ را به ما می‌دهد. چون بنابر فرض، بردار u_x معین و مخالف صفر است (در تمامی نقطه‌های قرص)، از جمله در نقطهٔ O هم باید معین و مخالف صفر باشد، ولی عدد چرخش میدان در طول محیط دایره‌ای به‌مرکز O و شعاع بسیار کوچک، ناگزیر برابر صفر می‌شود. به‌تناقض می‌رسیم و، بنابراین، ثابت می‌شود، ضمن نگاشت قرص بر خودش، دست‌کم یک نقطهٔ بی‌حرکت وجود دارد. این قضیه، حالت خاصی از قضیهٔ بسیار مهم *براوور* است که می‌گوید: در هر نگاشت پیوستهٔ یک کرهٔ n بعدی برخودش، دست‌کم یک نقطهٔ بی‌حرکت وجود دارد.

مسألهٔ وجود نقطه‌های بی‌حرکت در نگاشت‌های این یا آن گونه، بخش مهمی از توپولوژی خمینه‌ها را تشکیل می‌دهد.

۶. پیشرفت توپولوژی

توپولوژی سطح‌های بسته، تنها زمینه‌ای از توپولوژی است که تا پایان سدهٔ نوزدهم، کم و بیش روی آن کار شده بود. پدید آمدن این نظریه، به پیشرفت نظریهٔ تابع‌های با متغیر مختلط در طول سدهٔ نوزدهم مربوط می‌شود. این نظریه، که یکی از پدیده‌های ارزشمند در تاریخ ریاضیات سدهٔ نوزدهم است، با روش‌های گوناگونی ساخته شد. یکی از پربارترین این روش‌ها (پربارترین، به‌معنای درک ماهیت پدیده‌هایی که در معرض مطالعه بود)، روش هندسی *ریمان* است. روش *ریمان* به‌طور قانع کننده‌ای نشان داد که در نظریهٔ کلی تابع‌های با متغیر مختلط، نمی‌توان تنها به‌تابع‌های یک‌ارزشی قناعت کرد و، سپس، آن‌چه که امروز سطح‌های ریمانی می‌نامیم، ساخت. این سطح‌ها، در ساده‌ترین حالت تابع‌های جبری با متغیر مختلط، همیشه سطح‌های بستهٔ جهت‌پذیرند. بررسی ویژگی‌های توپولوژیک این گونه سطح‌ها، به‌مفهوم معینی، هم‌ارز با بررسی تابع جبری

---

۱. با توجه به‌این که فرض کردیم، در نگاشت f، نقطهٔ بی‌حرکت وجود ندارد، میدان مورد بررسی ما، همه جا معین و مخالف صفر است؛ و این اجازه می‌دهد دربارهٔ عدد چرخش آن در طول هر منحنی در داخل K صحبت کنیم.

مفروض است. پیشرفت بعدی اندیشهٔ ***ریمان***، به‌وسیلهٔ ***پوانکاره***، ***کلاین*** و ادامه دهندگان راه آن‌ها انجام گرفت که منجر به‌پیوندی عمیق و نامنتظر بین نظریهٔ تابع‌ها، توپولوژی سطح‌های بسته و هندسهٔ نااقلیدسی و به‌ویژه، نظریهٔ گروه‌های حرکت روی صفحهٔ لباچوسکی شد. به‌این‌ترتیب، برای نخستین بار، توپولوژی با شاخه‌های گوناگونی از ریاضیات پیوند خورد.

ضمن پیشرفت بعدی، روشن شد که تنها توپولوژی سطح کافی نیست و لازم است مسأله‌های مشخصی از توپولوژی n بعدی را حل کرد. نخستین این مسأله‌ها، مسألهٔ مربوط به ناوردایی و تغییرناپذیری توپولوژیک تعداد بعدهای فضا بود. مفهوم مسأله این بود که ثابت شود، از نظر توپولوژی نمی‌توان فضای اقلیدسی n بعدی را روی فضای m بعدی (برای $m \neq n$) نگاشت. این مسأله دشوار در سال ۱۹۱۱ به‌وسیلهٔ برا اوور حل شد[1]. در رابطه با حل این مسأله، روش‌های توپولوژیک تازه‌ای کشف شد که خیلی زود منجر به ساختن نظریهٔ نگاشت‌های پیوسته خمینه‌های چندبعدی و نظریهٔ میدان‌های برداری روی آن‌ها شد. در همهٔ این بررسی‌ها، نخستین مفهوم‌های اساسی توپولوژی «نظری ـ مجموعه‌ای» (یا توپولوژی عمومی) ضرورت پیدا کرد که بر پایهٔ نظریهٔ کلی مجموعه‌ها (که ***کانتور*** در ربع آخر سدهٔ نوزدهم ساخته بود) استوار بود.

در توپولوژی « نظری ـ مجموعه‌ای» خود موضوع بررسی، یعنی ردهٔ شکل‌های هندسی قابل بررسی، گسترش بسیار یافت و، اگر نه همهٔ مجموعه‌های فضاهای اقلیدسی به‌طور کلی، دست‌کم همهٔ مجموعه‌های بسته را در برگرفت. در حرکت تند جهت‌گیری «نظری ـ مجموعه‌ای» توپولوژی، دانشمندانی از کشورهای مختلف شرکت داشتند که به‌ویژه باید به‌مکتب توپولوژی لهستان اشاره کرد.

مسیر تازه‌ای در پیشرفت توپولوژی «نظری ـ مجموعه‌ای» را باید به‌طور اساسی در کارهای ریاضی‌دانان شوروی جست و جو کرد. این کارها با تلاش ***پاول ساموئیلوویچ اوریزُن*** (۱۸۹۸-۱۹۲۴) ریاضی‌دان جوانمرگ شوروی آغاز می‌شود. او نظریهٔ عمومی بعدها را آورد و براساس کلی‌ترین مفهوم‌های مجموعه‌های نقطه‌ای، معیاری برای تعداد بعدها به‌دست آورد. کار ***اوریزن*** بسیار ثمربخش بود و دیدگاه‌های تازه برای بررسی کلی‌ترین فرم‌های

---

۱. در واقع، برای پیشرفت نظریهٔ تابع‌های با متغیر مختلط، باید در آغاز مسأله دشوارتری حل می‌شد: تصویر توپولوژیک یک حوزهٔ n بعدی واقع در فضای n بعدی، به‌نوبهٔ خود، همیشه حوزه است. این مسأله را هم برا اوور حل کرد.

هندسی را برآورد[1]. اندیشه‌های *اوریزن*، که در نظریهٔ بعدهای او تکامل یافت، زمینه‌ای برای پدید آمدن دیدگاه‌های جالب *ل.آ.لوسترنیک* (با همراهی *ل.گ شنی ریمان*) دربارهٔ حساب وردشی شد.

در کارهای این ریاضی‌دانان، در کنار نتیجه‌گیری‌های دیگر، حل کامل و نهائی مسألهٔ مشهور *پوانکاره* دربارهٔ وجود سه منحنی بستهٔ ژئودزیک بدون نقطه‌های چندگانه روی هر سطح همسانریخت با کره، به‌دست آمد.

از طرف دیگر، نظریهٔ بعدها، *پ.س.آلکساندرف* را به‌روش‌های جبری توپولوژی ترکیبی در نظریهٔ مجموعه‌ها رساند که به‌نوبهٔ خود زمینه‌ای برای سمت‌گیری تازه‌ای در بررسی‌های مربوط به‌توپولوژی شد و در آن، ریاضی‌دانان شوروی، نقش جدی و درجه اول داشتند[2].

آن چه به‌ویژه به توپولوژی ترکیبی مربوط می‌شود، بعد از کارهای *پوانکاره* و *برا اوور*، که حدود سال ۱۹۱۵ میلادی انجام گرفت، نوبت به پژوهش‌های ریاضی‌دانان امریکایی می‌رسد: *ویلن*، *بیرک هویف*، *آلکساندر ولف‌شتس*، که توانستند نتیجه‌های بسیار با ارزشی به‌دست آورند. از جمله، *آلکساندر* ناوردایی توپولوژیک عددهای بتی را ثابت کرد، همچنین باید از قضیهٔ اصلی دوگانی او یاد کرد که آغازی برای بررسی‌های بعدی *پون‌تریاگین* شد؛ لف شتس دستور مشهور مربوط به‌عدد جبری نقطه‌های بی‌حرکت را در هر نگاشت پیوستهٔ خمینه‌ها داد و پایهٔ نظریهٔ کلی جبری نگاشت‌های پیوسته را ریخت که سپس به‌وسیله *هویف* تکامل یافت؛ *بیرک هویف* که روی نظریهٔ دستگاه‌های مکانیکی کار می‌کرد، در زمینهٔ فضاهای توپولوژیک و فضاهای متری هم کار کرد و غیره. پیشرفت

---

۱. تعریف استقرایی بعد را برای مجموعه‌ها، که به‌وسیلهٔ *اوریزن* پیشنهاد شد، می‌توان به‌عنوان تکامل اندیشهٔ *لباچوسکی* دانست که برش‌ها را پایه‌ای برای عمل‌های هندسی قرار داده بود. این تعریف، با تقریبی نه چندان دقیق، چنین است: مجموعه را، بعد صفر می‌گیریم، وقتی که بتوان آن را به‌صورت مجموعی از بخش‌های به‌اندازهٔ کافی کوچک در نظر گرفت، به‌نحوی که بین هر دو بخش آن «سایش» وجود نداشته باشد. مجموعه را n بعدی گویند، وقتی که زیر مجموعه‌های $(n-1)$ بعدی آن را بتوان طوری به‌بخش‌های به‌اندازهٔ کافی کوچک تقسیم کرد که دو به‌دو سایشی، باهم نداشته باشند و اگر این کار عملی نباشد، این عمل را با زیر مجموعه‌های با بعد کمتر از $(n-1)$ انجام داد (برای تعریف دقیق سایش، آن‌طور که در توپولوژی امروز فهمیده می‌شود، بند ۷ را ببینید).

۲. در این جا باید به نظریهٔ همتایی بعدها اشاره کرد که *پ.س.آلکساندروف* طرح کرد و براساس آن، *ل.س.یون تریاگین* و سپس در ادامه کار او *م.ف بوگ شتاین*، *و.گ.بولتیانسکی* و به‌ویژه *ک.آ.سیت نیکوف* توانستند نظریهٔ همتایی بعدها را تکامل دهند. دربارهٔ همزادی یا دوگانی (duolitg) *یون تریاگین* صحبت خواهیم کرد.

عمیق بعدی توپولوژی خمینه‌ها و نگاشت‌های پیوستهٔ آن‌ها، به‌کارهای **هوپف** مربوط می‌شود که، در کنار بسیاری از نتیجه‌گیری‌های دیگر خود، وجود بی‌نهایت نگاشت پیوستهٔ کرهٔ سه بعد روی کره دو بعدی را ثابت کرد، به‌نحوی که همهٔ آن‌ها باهم اختلاف دارند. اختلاف به‌این مفهوم که هیچ دو نگاشتی نمی‌توانند با تغییر پیوسته به‌یکدیگر منجر شوند. به‌این‌ترتیب، هوپف را باید بنیان‌گذار سمت‌گیری تازه‌ای به‌نام توپولوژی هموتوپی (هم‌مکانی = *homotopy*) دانست. نمایندگان مکتب تازهٔ توپولوژی در فرانسه (له ره، سی‌یرو و غیره) را باید ادامه‌دهندگان بعدی جنبه‌های مختلف توپولوژی و به‌ویژه توپولوژی ترکیبی دانست.

بررسی‌های عمیق ***اوریزن***، همان‌طور که گفتیم، سرآغازی برای فعالیت‌های ریاضی‌دانان شوروی در زمینهٔ توپولوژی بود. این بررسی‌ها به‌توپولوژی نظری ـ مجموعه‌ای مربوط می‌شد، ولی از اواخر سال‌های بیست سدهٔ بیستم، ریاضی‌دانان شوروی به‌سمت توپولوژی ترکیبی هم کشیده شدند. این بررسی‌ها، شکلی خود ویژه داشت و از روش‌های ترکیبی در مطالعهٔ مجموعه‌ها بسته، یعنی موضوع‌هایی که طبیعتی کلی داشتند، استفاده می‌کردند. در این زمینه، ***ل.س.پون تریاگین***، یکی از مهم‌ترین کشف‌های هندسی را تنظیم و ثابت کرد: قانون کلی دوگانی، که توانست رابطهٔ عمیق و از جهتی نهایی بین ساختمان توپولوژیک مجموعهٔ بسته مفروض واقع در فضای n بعدی اقلیدسی را با بخش‌های متمم فضا، پیدا کند. ***پون تریاگین*** در رابطه با قانون دوگانی، نظریهٔ کلی مشخصه‌های گروه‌های جابه‌جایی را ساخت که او را به‌بررسی بعدی او در زمینهٔ گروه‌های توپولوژیک و کلاسیک پیوسته لی، در حالت کلی، راهنمایی کرد. به‌دنبال آن، ***پون تریاگین*** و شاگردان او، یک رشته بررسی‌های مهم در توپولوژی خمینه‌ها و نگاشت‌های پیوستهٔ آن‌ها انجام دادند (***و.گ. بولتیانسکی، م.م.پوستن‌کوف*** و غیره). روش‌های تازه‌ای در این بررسی‌ها به‌کار گرفته شد: روش همانستگی ∇، که ***کولموگوروف*** و، بدون ارتباط با او، ***آلکساندر***، در توپولوژی ترکیبی وارد کرده بودند. این روش، که جای نخست را در تمامی توپولوژی هموتوپی گرفت، امکان ادامهٔ همه‌جانبهٔ نظریهٔ دوگانی ***پون تریاگین*** را فراهم کرد که منجر به قضیه‌های دوگانی ***کولموگوروف*** (و ***آلکساندر***)، ***پ.س.آلکساندرف*** و ***ک.آ. سیت نیکوف*** شد که از جملهٔ مهم‌ترین نتیجه‌گیری‌های توپولوژی امروزی است. این روش، کاربردی اساسی در کارهای ***ل.آ.لوسترنیک***، در حساب وردشی هم پیدا کرد.

۷. فضاهای متریک و توپولوژیک

در آغاز این بخش، دربارهٔ سایش، مجاورت یا تماس (بخش‌های مختلف شکل مفروض)، به‌عنوان مفهوم اصلی توپولوژی صحبت کردیم، و تبدیل‌های پیوسته را به‌عنوان تبدیل‌هایی که این رابطه را حفظ می‌کنند، تعریف کردیم. با وجود این، در آن‌جا تعریف دقیق این مفهوم اصلی را نیاوردیم. این تعریف را به‌صورتی کلی، تنها بر پایهٔ مفهوم‌های نظریهٔ مجموعه‌ها می‌توان داد. در این بند می‌خواهیم در این باره صحبت کنیم و پیش از آن، مقدمه‌ای دربارهٔ مفهوم فضای توپولوژیک می‌آوریم.

در نظریهٔ مجموعه‌ها می‌توان به مفهوم‌هایی از شکل‌های هندسی رسید که در ریاضیات «سنتی» قابل دسترسی نیستند. موضوع بررسی‌های هندسی، و در حالت خاص، بررسی‌های توپولوژیک، شامل هرگونه مجموعه‌های نقطه‌ای می‌شود، یعنی مجموعه‌ای که عضوهای آن، نقطه‌هایی از فضای n بعدی اقلیدسی باشند. برای نقطه‌های فضای n بعدی، مفهوم ***فاصله*** تعریف شده است: فاصلهٔ بین نقطه‌های

$$A(x_۱ , x_۲ , \dots , x_n) , \; B(y_۱ , y_۲ , \dots , y_n)$$

بنابه تعریف، برابر است با عدد نامنفی

$$\rho(A , B) = \sqrt{(x_۱ - y_۱)^۲ + (x_۲ - y_۲)^۲ + \dots + (x_n - y_n)^۲}$$

مفهوم فاصله، این امکان را پدید می‌آورد که سایش یا مجاورت را، در آغاز بین مجموعه و نقطه، و سپس بین دو مجموعه، تعریف کنیم. می‌گوییم، نقطهٔ A، نقطهٔ سایش (یا نقطهٔ حدی) مجموعهٔ M است (یا نقطهٔ A به مجموعهٔ M چسبیده است) وقتی که در مجموعهٔ M نقطه‌ای وجود داشته باشد که فاصلهٔ آن از نقطهٔ A، کمتر از هر عدد مثبت مفروضی باشد که از پیش انتخاب کرده‌ایم. روشن است، هر نقطه از مجموعهٔ مفروض، نقطهٔ سایش آن است، ولی نقطه‌ای هم که عضو مجموعهٔ مفروض نباشد، ممکن است نقطهٔ سایش آن باشد. برای نمونه، بازهٔ باز (۱ , ۰) را روی محور عددی در نظر می‌گیریم، یعنی مجموعهٔ همهٔ نقطه‌های واقع در بین (۱ , ۰)، نقطه‌های ۰ و ۱ متعلق به‌این مجموعه نیستند [در بازهٔ (۱ , ۰) قرار ندارند]، ولی نقطه‌های سایش‌اند، زیرا در بازهٔ (۱، ۰) نقطه‌هایی وجود دارند که هر چه بخواهیم به نقطهٔ ۰ نزدیک‌اند و، به‌همین ترتیب، نقطه‌هایی که به‌دل‌خواه نزدیک به ۱ هستند.

مجموعه را بسته می‌نامیم، وقتی که شامل همهٔ نقطه‌های سایش خود باشد. پاره‌خط راست [۱ , ۰] از خط راست عددی، یعنی مجموعهٔ همهٔ نقطه‌های x که در نابرابری ۱ ≥ x ≥ ۰ صدق کنند، بسته است.

مجموعه‌های بسته روی صفحه، و از آن بیشتر، مجموعه‌های بسته در فضای سه‌بعدی یا با بعدهای بیشتر، ممکن است ساختار پیچیده‌ای داشته باشند؛ این مجموعه‌ها هم، به‌طور عمده، موضوع مطالعهٔ توپولوژی «نظری ـ مجموعه‌ای» فضای n بعدی‌اند. سپس یادآوری می‌کنیم، دو مجموعهٔ P و Q را در سایش با یکدیگر گوییم، وقتی که دست‌کم یکی از آن‌ها شامل نقطه‌ای در سایش با مجموعهٔ دیگر باشد. با توجه به‌این تعریف، دو مجموعهٔ متقاطع (یعنی مجموعه‌هایی که نقطه‌های مشترک دارند) همیشه در سایش با یکدیگرند، از آن چه گفتیم می‌توان نتیجه گرفت که، دو مجموعهٔ بسته، تنها وقتی سایش دارند که دست کم یک نقطهٔ مشترک داشته باشند؛ ولی در مثل، پاره‌خط راست [۱ , ۰] و بازهٔ (۲ , ۱) که نقطهٔ مشترکی ندارند، در سایش‌اند، زیرا نقطهٔ ۱، که به‌پاره‌خط راست [۱ , ۰] تعلق دارد، در ضمن نقطهٔ سایش بازهٔ (۲ , ۱) است. اکنون می‌توان گفت: مجموعهٔ R به‌وسیلهٔ مجموعهٔ S که در آن واقع است تقسیم شده است (یا «برش» یافته است)، یا S «مقطعی» (یا بخشی) از مجموعهٔ R است، وقتی که مجموعهٔ $R-S$، که شامل همهٔ نقطه‌هایی از مجموعهٔ R است که به S تعلق ندارند، بتواند به‌عنوان مجموع دو مجموعه‌ای تصور شود که باهم سایش ندارند.

به‌این ترتیب، اندیشه‌های ***لباچوسکی*** دربارهٔ سایش‌ها و مقطع‌های مجموعه‌ها، در توپولوژی جدید، دقت و بیان کلی خود را پیدا کرد. بر پایهٔ همین اندیشه‌ها بود که ***اوریزن*** تعریف بعد را برای هر مجموعه‌ای ارائه داد، تعریفی که، پس از آن، درون‌مایه‌ای دقیق‌تر و کلی‌تر پیدا کرد. تنظیم دقیق‌تر این تعریف، به‌نگاشت‌ها یا تبدیل‌های پیوسته مربوط می‌شود: نگاشت f از مجموعهٔ X روی مجموعهٔ Y را پیوسته گویند، وقتی که ضمن این نگاشت مفهوم سایش به‌قوت خود باقی بماند، یعنی از این فرض که نقطهٔ A از مجموعهٔ X، برای زیر مجموعهٔ P از مجموعهٔ Y، نقطهٔ سایش است، نتیجه شود که تصویر نقطهٔ A یعنی $f(A)$ هم برای تصویر مجموعهٔ P یعنی $f(P)$، نقطهٔ سایش باشد. سرانجام نگاشت یک‌ارزشی متقابل (یعنی با متناظر یک به‌یک) از مجموعهٔ X روی مجموعهٔ Y، وقتی ***توپولوژیک*** نامیده می‌شود که خود این نگاشت پیوسته باشد و، در ضمن برعکس، نگاشت وارون مجموعهٔ Y روی مجموعهٔ X هم پیوسته باشد. این تعریف‌ها، پایه‌ای دقیق برای طرح موضوع‌هایی شد که در این بخش آورده‌ایم.

ولی توپولوژی «نظری ـ مجموعه‌ای» در همین جا متوقف نشد که تنها به‌بررسی همهٔ مجموعه‌های نقطه‌ای، به‌عنوان شکل‌های هندسی بپردازد. به‌طور طبیعی، این مسأله مطرح شد که مفهوم فاصله، نه‌تنها برای نقطه‌های یک فضای اقلیدسی، بلکه برای موضوع‌های دیگری هم، که هیچ رابطی به‌هندسه ندارند، به‌کار رود.

برای نمونه، مجموعهٔ همهٔ تابع‌های پیوسته‌ای را در نظر می‌گیریم که روی پاره‌خط راست [۱ , ۰] معین باشند. فاصلهٔ $\rho(f, g)$ بین دو تابع f و g را، به‌عنوان حداکثر عبارت $|f(x) - g(x)|$، وقتی x همهٔ پاره‌خط راست [۱ , ۰] را می‌پیماید، تعریف می‌کنیم. این «فاصله» دارای همهٔ ویژگی‌های اساسی فاصلهٔ بین دو نقطه از فضا است $\rho(f, g)$ بین دو تابع f و g، وقتی و تنها وقتی برابر صفر است که این تابع‌ها برهم منطبق باشند. از طرف دیگر برای هر دو تابع f و g داریم: $\rho(f, g) = \rho(g, f)$؛ در ضمن، این تعریف فاصله، با به‌اصطلاح «نابرابری مثلثی» سازگار است، یعنی برای هر سه تابع $f_۱$، $f_۲$ و $f_۳$ داریم:

$$\rho(f_۱ , f_۲) + \rho(f_۲ , f_۳) \geq \rho(f_۱ , f_۳)$$

در واقع، این تعریف از فاصله، مجموعهٔ تابع‌های ما را، در فضای متری قرار می‌دهد (که به‌طور معمول با C نشان می‌دهند). وقتی از فضای متری صحبت می‌کنیم، مجموعه‌ای از موضوع‌های دل‌خواه را در نظر داریم که آن‌ها را ***نقطه‌های*** فضای متری می‌نامیم، با این تنها شرط که فاصلهٔ بین هر دو نقطهٔ آن معین و نامنفی و، در ضمن، با نابرابری مثلثی سازگار باشد.

وقتی یک فضای متری داده شده باشد، می‌توان دربارهٔ نقطه‌های سایش زیرمجموعه‌های آن و بنابراین، دربارهٔ سایش زیر مجموعه‌های آن باهم و به‌طور کلی دربارهٔ مفهوم‌های توپولوژیک صحبت کرد (مثل مجموعه‌های بسته، نگاشت‌های پیوسته و دیگر مفهوم‌هایی در این فضا وجود دارد). از این راه، میدان گسترده و بی‌اندازه ثمربخش برای به‌کار گرفتن اندیشه‌ها توپولوژیک و به‌طور کلی اندیشه‌های هندسی در این گروه از موضوع‌های ریاضی گشوده می‌شود، موضوع‌هایی که دربارهٔ آن، به‌ظاهر، صحبتی از این یا آن هندسه نمی‌توان داشت. مثالی می‌آوریم.

دوباره معادلهٔ دیفرانسیلی (۲) را انتخاب می‌کنیم.

$$\frac{dy}{dx} = F(x , y)$$

اگر $y=\varphi(x)$ را، برای $\circ \leq x \leq ۱$، جوابی از این معادله بگیریم که با فرض $y=\circ$ به‌ازای $x=\circ$ به‌دست آمده است، آن وقت روشن است که تابع $\varphi(x)$ در این معادلهٔ انتگرالی صدق می‌کند:

$$\varphi(x)=\int_{\circ}^{x}F(x\,,\,\varphi(x))dx \qquad (۳)$$

اکنون انتگرال $G(f)=\int_{\circ}^{x}F(x\,,\,f(x))dx$ را در نظر می‌گیریم که در آن $\circ \leq x \leq ۱$ و $f(x)$، تابعی پیوسته روی پاره‌خط راست $[\circ\,,\,۱]$ باشد. این انتگرال، تابعی پیوسته مثل $g(x)$ است که در ضمن، روی پاره‌خط راست $[\circ\,,\,۱]$ معین است. به‌این‌ترتیب، بیان $G(f)=\int_{\circ}^{x}F(x\,,\,f(x))dx$، هر تابع f را متناظر با تابع $g=G(f)$ می‌کند؛ به‌زبان دیگر، به‌نگاشت G می‌رسیم که فضای پیوسته و متری C را به‌توی خودش می‌نگارد. ببینیم تابع $\varphi(x)$ را (که تعداد آن‌ها ممکن است بی‌نهایت باشد)، که جواب معادلهٔ (۲) یا هم‌ارز آن، معادلهٔ (۳) است، چگونه مشخص می‌شود؟ روشن است که، ضمن این نگاشت، این تابع منجر به‌خودش می‌شود، یعنی نقطه‌ای بی‌حرکت در نگاشت G است: در واقع، وجود چنین نقطهٔ بی‌حرکتی در نگاشت G نتیجه‌ای از قضیهٔ کلی مربوط به‌نقطه‌های بی‌حرکت نگاشت‌های فضاهای متری است که در سال ۱۹۲۶ به‌وسیلهٔ ***پ.س.آلکساندروف*** و ***و.وسنریک*** ثابت شد. در زمان ما، بررسی فضاهای مختلف متری، که نقطه‌های آن‌ها را گونه‌ای از تابع‌ها (همچنین فضاهای تابعی) تشکیل می‌دهند، همه‌جا به‌عنوان ابزاری از آنالیزها کاربرد دارد؛ مطالعهٔ فضاهای تابعی تا اندازه‌ای با روش‌های توپولوژیک، و به‌طور گسترده‌ای با روش‌های جبری (به‌مفهوم وسیع آن)، درون‌مایهٔ آنالیز تابعی را تشکیل می‌دهد (بخش نوزدهم را ببینید).

آنالیز تابعی، همان‌طور که پیش از این هم گفته‌ایم، جای بسیار نمایانی در دانش ریاضی امروز دارد، و این به‌دلیل بستگی‌های چندجانبه‌ای است که با دیگر شاخه‌های ریاضیات و به دلیل اهمیتی است که برای دانش‌های طبیعی و در درجهٔ اول فیزیک نظری دارد. بررسی ویژگی‌های توپولوژیک فضای تابعی رابطهٔ نزدیکی با حساب وردش‌ها و نظریهٔ معادله‌های با مشتق‌های جزئی دارد (بررسی‌های ***ل.آ.لوسترنیک***، ***مورس***، ***لِره***، ***پووادر***، ***م.آ.کراستورسلسکی*** و دیگران). مسأله‌های مربوط به وجود نقطه‌های بی‌حرکت نگاشت‌های پیوستهٔ فضاهای تابعی، جای نمایانی در این بررسی‌ها داشته است.

توپولوژی فضاهای تابعی و به‌طور کلی فضاهای متری، هنوز حرف آخر در تعمیم

نظریه‌های توپولوژیک نیست. موضوع این است که در فضاهای متری، مفهوم اصلی توپولوژی، یعنی سایش، از فاصلهٔ بین نقطه‌ها نتیجه می‌شود که، به‌خودی خود، مفهومی توپولوژیک نیست. بنابراین، این مسأله در برابر ما قرار دارد که تعریف اصل موضوعی سایش را، به‌طور مستقیم ارائه دهیم. به‌این‌ترتیب است که به‌مفهوم فضای توپولوژیک، کلی‌ترین مفهوم توپولوژی امروزی می‌رسیم.

فضای توپولوژیک عبارت است از مجموعهٔ موضوع‌هایی با طبیعتی دل‌خواه (به‌نام نقطه‌های فضای توپولوژیک) که در آن، برای هر زیرمجموعه، با این یا آن وسیله، نقطه‌های ***سایش*** مفروض باشد. در ضمن باید برخی شرط‌های طبیعی را، که اصل موضوع‌های فضای توپولوژیک می‌نامیم، در نظر داشته باشیم (از جمله، هر نقطهٔ از مجموعهٔ مفروض، نقطهٔ سایش آن است، نقطهٔ سایش اجتماع دو مجموعهٔ دست‌کم نقطهٔ سایش یکی از آن‌ها است و غیره). نظریهٔ فضاهای توپولوژیک، شاخهٔ کارآمد و عمیقی از ریاضیات است و در تکامل آن، ریاضی‌دانان شوروی نقشی جدی داشته‌اند: ***پ. س. وریزن***، ***پ. س. آلکساندرف***، ***آ. ن. یکخوفوف*** و دیگران. از نتیجه‌های تازه در فضاهای توپولوژیک، باید از کار **یو.م. سمیرنوف** نام برد که شرط لازم و کافی برای متری‌پذیر بودن فضای توپولوژیک را پیدا کرد، یعنی شرطی که، به‌ازای آن، بتوان بین نقطه‌های فضاطوری فاصله را معین کرد که «توپولوژی» فضا بتواند به‌عنوان نتیجه‌ای از مفهوم فاصله به دست آید. به زبان دیگر، نقطه‌های سایش همهٔ مجموعه‌های ممکن که از فضای متری به‌دست آمده‌اند، همان‌هایی باشند که از همان آغاز در فضای توپولوژیک مفروض مشخص بودند.

بخش نوزدهم

آنالیز تابعی

ای. م. هلفوند

پیدایش و پیشرفت آنالیز تابعی در سدهٔ بیستم، به دو دلیل بود. از طرفی لازم بود، موضوع‌ها و حقیقت‌های فراوانی که در شاخه‌های گوناگون ریاضیات در جریان سدهٔ نوزدهم روی‌هم جمع شده بود و اغلب بی‌ارتباط باهم به‌نظر می‌آمدند، یک‌جا و با دیدی یگانه بررسی و قابل فهم شود. در ضمن، مفهوم‌های اصلی آنالیز تابعی، در سمت‌های مختلف و به‌بهانه‌های مختلف پدید آمده و شکل گرفته بود. بسیاری از مفهوم‌های آنالیز تابعی، به‌صورتی طبیعی و در جریان پیشرفت حساب وردشی، در مسأله‌های مربوط به‌نوسان‌ها (ضمن عبور از نوسان دستگاه با تعداد محدودی درجهٔ آزادی، به‌نوسان محیط‌های پیوسته و کامل)، در نظریهٔ معادله‌های انتگرالی، در نظریهٔ معادله‌های دیفرانسیلی، چه عادی و چه با مشتق‌های جزئی (در مسأله‌های مرزی، در مسأله‌های مربوط به‌ویژه مقدارها و غیره)، ضمن پیشرفت نظریهٔ تابع‌های حقیقی، در محاسبه‌های اپراسیونی (عمل‌گرها)، ضمن بررسی مسأله‌های مربوط به‌تقریب تابع‌ها و غیره، پدید آمد. آنالیز تابعی موجب شد بسیاری از نتیجه‌گیری‌های پراکنده در این رشته‌ها، به‌صورتی واحد و یگانه درک شود و اغلب به‌نتیجه‌های تازه‌ای برسد. مفهوم‌هایی که به‌این‌ترتیب فراهم شده بود، در میانهٔ سدهٔ بیستم در بخش تازه‌ای از فیزیک نظری، یعنی در مکانیک کوانتایی، کاربرد پیدا کرد.

از طرف دیگر، بررسی مسأله‌هایی از ریاضیات که به‌مکانیک کوانتایی مربوط می‌شد؛ موجب تحولی قطعی در پیشرفت خود آنالیز تابعی شد و سمت‌گیری تکاملی آن را شکل داد.

آنالیز تابعی، هنوز با پایان خود فاصله دارد، در این تردیدی نیست که در تکامل آنالیز تابعی، مسأله‌ها و نیازهای فیزیک امروزی دارای همان اهمیتی است که زمانی و در سدهٔ هجدهم، مکانیک سنتی برای پدید آمدن وتکامل محاسبه‌های دیفرانسیلی وانتگرالی داشت.

بدون این که بخواهیم در این بخش به همهٔ جنبه‌ها بپردازیم، تلاش می‌کنیم اصلی‌ترین

مسأله‌های آنالیز تابعی را بیاوریم. بسیاری از موضوع‌های مهم، در بیرون از محدودهٔ این بخش‌اند. می‌خواهیم دست‌کم در محدودهٔ برخی از مسأله‌های انتخابی، خواننده را با بعضی مفهوم‌های اساسی آنالیز تابعی آشنا کنیم و تا حد امکان، بستگی‌هایی را، که از آن‌ها یاد کردیم، روشن کنیم. این موضوع‌ها، در واقع، همان‌هایی هستند که هیلبرت، یکی از پایه‌گذاران آنالیز تابعی، در آغاز سدهٔ بیستم به آن‌ها پرداخته بود. از آن زمان به بعد، سرعت پیشرفت آنالیز تابعی شتاب گرفت و به‌تقریب در همهٔ شاخه‌های ریاضیات به‌صورت گسترده‌ای کاربرد پیدا کرد: در معادله‌های دیفرانسیلی با مشتق‌های جزئی، در نظریهٔ احتمال، در مکانیک کوانتایی، در نظریهٔ کوانتایی میدان و غیره. باید اعتراف کنم که این جنبه‌های آنالیز تابعی در این جا مطرح نشده است. در واقع، برای هر کدام از این جنبه‌ها باید مقالهٔ جداگانه‌ای تهیه کرد، و به‌همین مناسبت، در این جا خود را به یکی از مسأله‌های قدیمی‌تر، یعنی ویژه‌تابع‌ها، محدود کرده‌ایم.

۱. فضای nبعدی

از این‌جا به بعد، از مفهوم‌های اساسی مربوط به فضای nبعدی استفاده کرده‌ایم. گرچه در بخش‌های مربوط به جبر خطی و فضاهای انتزاعی، از این مفهوم‌ها یاد شده است، بی‌فایده نیست که در این‌جا هم، آن‌ها را به‌کوتاهی و به ترتیبی که بعد از این با آن‌ها روبه‌رو می‌شویم، بیاوریم. برای مطالعهٔ این بند، کافی است با پایه‌های اصلی هندسهٔ تحلیلی آشنا باشیم.

می‌دانیم در هندسهٔ تحلیلی، نقطه‌ای را که در فضای سه‌بعدی واقع باشد، به‌وسیلهٔ سه عدد $(f_۱, f_۲, f_۳)$، یعنی مختصات آن، مشخص می‌کنند. فاصلهٔ این نقطه تا مبدأ مختصات برابراست با $\sqrt{f_۱^۲+f_۲^۲+f_۳^۲}$ ،کسینوس زاویهٔ بین دوبردار غیرصفرکه از مبدأ مختصات به دو نقطهٔ مختلف $A(f_۱, f_۲, f_۳)$ و $B(g_۱, g_۲, g_۳)$رفته‌اند، با این دستور به‌دست می‌آید:

$$\cos\varphi = \frac{|f_۱g_۱+f_۲g_۲+f_۳g_۳|}{\sqrt{f_۱^۲+f_۲^۲+f_۳^۲}\cdot\sqrt{g_۱^۲+g_۲^۲+g_۳^۲}}$$

از مثلثات می‌دانیم $|\cos\varphi| \leq ۱$. بنابراین، این نابرابری برقرار است:

$$\frac{|f_1g_1+f_2g_2+f_3g_3|}{\sqrt{f_1^2+f_2^2+f_3^2}\cdot\sqrt{g_1^2+g_2^2+g_3^2}}\leq 1$$

یعنی، همیشه داریم:

$$(f_1g_1+f_2g_2+f_3g_3)^2\leq(f^2{}_1+f^2{}_2+f^2{}_3)(g^2{}_1+g^2{}_2+g^2{}_3) \qquad (1)$$

این نابرابری، که خصلت جبری دارد، برای هر شش عدد دل‌خواه (f_1 , f_2 , f_3) و (g_1 , g_2 , g_3) برقرار است، زیرا هر شش عدد دل‌خواه را می‌توان به‌عنوان مختصات دو نقطه از فضا در نظر گرفت. در ضمن، نابرابری، (۱) از راه ملاحظه‌های خالص هندسی به‌دست می‌آید و بستگی کامل با هندسه دارد که امکان عینی بودن آن را فراهم می‌آورد.

وقتی رابطه‌های هندسی را به صورت تحلیلی تنظیم می‌کنیم، اغلب معلوم می‌شود که آن‌ها را می‌توان، با تغییر سه عدد به n عدد هم به‌کار برد. از جمله، نابرابری (۱) را می‌توان برای $2n$ عدد

$$(f_1 , f_2 , \ldots , f_n) \quad \text{و} \quad (g_1 , g_2 , \ldots , g_n)$$

هم نوشت. و این، به‌معنای آن است که نابرابری (۱)، برای $2n$ عدد هم درست است، یعنی داریم:

$$(f_1g_1+f_2g_2+\cdots+f_ng_n)^2\leq(f_1^2+f_2^2+\cdots+f_n^2)(g_1^2+g_2^2+\cdots+f_n^2) \qquad (1')$$

درستی این نابرابری را، که حالت خاص آن نابرابری (۱) است، می‌توان با روش خالص تحلیلی ثابت کرد[1]. به‌همین‌ترتیب، بسیاری از رابطه‌های دیگری که در هندسهٔ تحلیلی بین سه عدد به‌دست می‌آید، می‌توان برای n عدد تعمیم داد. چنین بستگی بین رابطه‌های عددی (رابطه‌های کمیتی)، که نمونهٔ آن را دربارهٔ نابرابری‌های (۱) و ('۱) دیدیم، می‌تواند به‌یاری هندسه بیشتر نمایان شود، به‌شرطی که مفهوم فضای n بعدی را در بحث خود وارد کنیم. این مفهوم در بخش شانزدهم داده شده است و در این‌جا، باکوتاهی از آن یاد می‌کنیم.

مجموعهٔ n عدد $(f_1 , f_2 , \ldots , f_n)$ را **نقطه** یا **بردار** در فضای n بعدی می‌گویند (از این به‌بعد، بیشتر از نام بردار استفاده می‌کنیم). بردار $(f_1 , f_2 , \ldots , f_n)$ را هم، به‌صورتی کوتاه و با یک حرف f نشان می‌دهیم.

---

۱. بخش شانزدهم، بند ۲، عنوان «فضای اقلیدسی n بعدی» را ببینید.

همان‌طور که در فضای سه بعدی برای جمع بردارها، مختصات آن‌ها را با هم جمع می‌کنیم، در این جا هم، مجموع بردارهای

$$f=(f_۱ , f_۲ , ... , f_n) \text{ و } g=(g_۱ , g_۲ , ... , g_n)$$

به‌عنوان بردار $(f_۱+g_۱ , f_۲+g_۲ , ... , f_n+g_n)$ شناخته و به‌صورت $f+g$ نشان داده می‌شود.

حاصل‌ضرب بردار $f=(f_۱ , f_۲ , ... , f_n)$ در عدد λ عبارت است از بردار $\lambda f=(\lambda f_۱ , \lambda f_۲ , ... , \lambda f_n)$.

طول بردار $f=(f_۱ , f_۲ , ... , f_n)$، شبیه طول بردار در فضای سه‌بعدی، به‌عنوان $\sqrt{f_۱^۲+f_۲^۲+ ... +f_n^۲}$ معین می‌شود.

زاویهٔ φ بین دو بردار

$$f=(f_۱ , f_۲ , ... , f_n) , \; g=(g_۱ , g_۲ , ... , g_n)$$

در فضای n بعدی، به‌وسیلهٔ کسینوس خود، شبیه آنچه برای کسینوس زاویهٔ بین دو بردار در فضای سه‌بعدی داشتیم، با دستور بیان می‌شود[1]:

$$\cos\varphi=\frac{f_۱g_۱+f_۲g_۲+ ... +f_ng_n}{\sqrt{f_۱^۲+f_۲^۲+ ... +f_n^۲}\cdot\sqrt{g_۱^۲+g_۲^۲+ ... +g_n^۲}} \qquad (۲)$$

حاصل‌ضرب عددی (یا حاصل‌ضرب اسکالر) دو بردار، به‌حاصل‌ضرب طول‌های این بردارها در کسینوس زاویهٔ بین آن‌ها گفته می‌شود. بنابراین، اگر

$$f=(f_۱ , f_۲ , ... , f_n) , g=(g_۱ , g_۲ , ... , g_n)$$

از آن جا که طول این بردارها، به‌ترتیب برابر

$$\sqrt{f_۱^۲+f_۲^۲+ ... +f_n^۲} , \; \sqrt{g_۱^۲+g_۲^۲+ ... +g_n^۲}$$

است، حاصل‌ضرب اسکالر آن‌ها، که به‌صورت (f , g) نشان می‌دهیم، چنین می‌شود.

---

۱. درستی نابرابری $|\cos\varphi| \leq ۱$ نتیجه‌ای است از نابرابری $(۱')$.

$$(f,\ g) = f_1 g_1 + f_2 g_2 + \cdots + f_n g_n \qquad (3)$$

به‌ویژه شرط این که دو بردار برهم عمود باشند، برابری $\cos\varphi = 0$، یعنی $(f,\ g) = 0$ است.

به‌یاری دستور (۳)، می‌توان قانع شد که، حاصل‌ضرب اسکالر در فضای n بعدی، این ویژگی‌ها را دارد:

۱. $(f,\ g) = (g,\ f)$:

۲. $(\lambda f,\ g) = \lambda(f,\ g)$:

۳. $(f,\ g_1 + g_2) = (f,\ g_1) + (f,\ g_2)$:

۴. $(f,\ f) \geq 0$، در ضمن علامت برابری تنها برای وقتی است که داشته باشیم $f = 0$ یعنی وقتی که $f_1 = f_2 = \cdots = f_n = 0$.

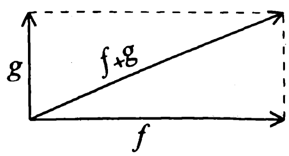
حاصل‌ضرب اسکالر بردار f درخودش، یعنی $(f,\ f)$، برابر است با توان دوم طول بردار f.

حاصل‌ضرب اسکالر، وسیلهٔ بسیار ساده‌ای برای بررسی فضای n بعدی است. در این‌جا، به‌بررسی هندسهٔ فضای n بعدی نمی‌پردازیم و تنها یک مثال می‌آوریم.

قضیه **فیثاغورس** را در فضای n بعدی در نظر می‌گیریم: توان دوم طول وتر برابر است با مجموع توان‌های دوم ضلع‌های دیگر. برای این منظور، اثباتی از قضیه در روی صفحه را انتخاب می‌کنیم که به سادگی به فضای n بعدی تعمیم داده شود.

f و g را دو بردار عمود برهم و واقع در صفحه می‌گیریم و مثلث راست‌گوشه‌ای را که روی بردارهای f و g ساخته شده است، انتخاب می‌کنیم (شکل ۱). وتر این مثلث برابر است با طول بردار $f+g$. قضیه **فیثاغورس** را، با نمادهای برداری می‌نویسیم. چون توان دوم طول بردار برابر است با حاصل‌ضرب اسکالر بردار در خودش، بنابراین با اصطلاح‌های حاصل‌ضرب اسکالر، قضیهٔ **فیثاغورس** به‌این صورت درمی‌آید:

$$(f+g,\ f+g) = (f,\ f) + (g,\ g)$$



شکل ۱

اثبات به‌سادگی و بلافاصله، از ویژگی‌های حاصل‌ضرب اسکالر نتیجه می‌شود. در واقع

$$(f+g\,,\,f+g)=(f\,,\,f)+(f\,,\,g)+(g\,,\,f)+(g\,,\,g)$$

دو جملهٔ میانی در سمت راست این برابری، به‌دلیل این که دو بردار f و g برهم عمودند، برابر صفرند.

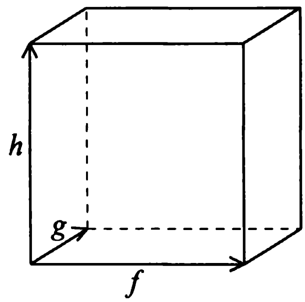
در اثباتی که آوردیم، تنها از تعریف طول بردار، عمود بودن بردارها برهم و ویژگی‌های حاصل‌ضرب اسکالر استفاده کردیم. بنابراین، اگر f و g را دو بردار عمود برهم در فضای n بعدی فرض کنیم هیچ‌چیز تغییر نمی‌کند. به‌این‌ترتیب، قضیهٔ **فیثاغورس** برای مثلث راست‌گوشه در فضای n بعدی هم درست است.

اگر سه بردار دو به‌دو عمود برهم f ، g و h را در فضای n بعدی داده باشند، آن وقت مجموع این بردارها، یعنی $f+g+h$، عبارت است از قطر مکعب مستطیلی که روی این بردارها ساخته شده است (شکل ۲) و در نتیجه، به‌این برابری می‌رسیم:

$$(f+g+h\,,\,f+g+h)=(f\,,\,f)+(g\,,\,g)+(h\,,\,h)$$

و به‌این معناست که توان دوم طول قطر مکعب مستطیل برابر است با مجموع توان‌های دوم طول‌های آن. اثبات این حکم را، که شبیه اثبات قضیه **فیثاغورس** به‌انجام می‌رسد، به‌عهدهٔ خواننده می‌گذاریم. درست به‌همین‌ترتیب، اگر k بردار دو به‌دو عمود برهم را در فضای n بعدی را در نظر بگیریم و آن‌ها را f^1 ، f^2 ، ... ، f^k بنامیم، به‌همان سادگی حالت‌های دوبعدی و سه‌بعدی، به‌این برابری می‌رسیم:

$$(f^1+f^2+\cdots+f^k\,,\,f^1+f^2+\cdots+f^k)=(f^1\,,\,f^1)+(f^2\,,\,f^2)+\cdots+(f^k\,,\,f^k) \qquad (۴)$$



شکل ۲

که به‌معنای آن است که توان دوم طول قطر «مکعب مستطیل k بعدی» در فضای n بعدی برابر است با مجموع توان‌های دوم طول‌های یال‌های آن.

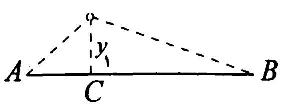
۲. فضای هیلبرت (فضای بی‌نهایت بعدی)

رابطه با فضای n بعدی. وارد کردن مفهوم فضای n بعدی، سودمندی خود را برای بررسی یک رشته از دشواری‌های ریاضیات و فیزیک نشان داده است. در ضمن این مفهوم، به تکامل مفهوم فضا و کاربرد آن در شاخه‌های گوناگون ریاضیات، تکان شدیدی داده است. در تکامل جبر خطی و هندسهٔ فضای n بعدی، مسأله‌های مربوط به‌نوسان‌های کوچک دستگاه‌های کشسان، نقش زیادی داشته است.

مثالی عادی و سنتی از این مسأله را مطرح می‌کنیم (شکل ۳). AB را نخ نرم کشسانی می‌گیریم که بین نقطه‌های A و B کشیده شده باشد، فرض می‌کنیم در نقطه‌ای از آن مثل C وزنه‌ای محکم شده باشد. اگر وزنه را از حالت تعادل خارج کنیم، با بسامدی (فرکانسی) مثل ω آغاز به نوسان می‌کند که اگر نیروی کشش نخ γ، جرم m و جای وزنه را بدانیم، می‌توانیم آن را محاسبه کنیم. در ضمن، موقعیت دستگاه در هر لحظهٔ زمانی، به وسیلهٔ یک عدد یعنی انحراف y_1 جرم m از حالت تعادل نخ داده می‌شود.

اکنون، n وزنه در نقطه‌های C_1 ، C_2 ، ... ، C_n از نخ AB قرار می‌دهیم. در ضمن، نخ را بی‌وزن به‌حساب می‌آوریم و این، به‌معنای آن است که جرم نخ آن قدر کم است که در مقایسه با جرم وزنه‌ها، می‌توان از آن صرف‌نظر کرد. موقعیت چنین دستگاهی با n عدد y_1 ، y_2 ، ... ، y_n داده می‌شود (n عددی محدود است) که انحراف وزنه‌ها را از حالت تعادل مشخص می‌کنند. مجموعهٔ عددهای y_1 ، y_2 ، ... ، y_n را می‌توان به‌صورت بردار

$$(y_1 , y_2 , \dots , y_n)$$



شکل ۳

نشان داد که برداری در فضای n بعدی است (و این نمایش برداری، از بسیاری جهت‌ها سودمند است).

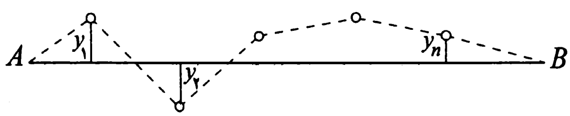
بررسی نوسان‌های کوچک، که در چنین طرحی پیش می‌آید، بستگی تنگاتنگی با مفهوم‌های اساسی فضای n بعدی دارد. از جمله، تعیین بسامد یا فرکانس نوسان‌های چنین دستگاهی را می‌توان منجر به‌مسألهٔ مربوط به پیدا کردن محور یک بیضوی (الیپسوئید) در فضای n بعدی کرد.

اکنون نوسان‌های کوچک سیمی را بررسی می‌کنیم که بین دو نقطهٔ A و B کشیده شده باشد. در ضمن، سیم ایده‌آلی را در نظر داریم، یعنی نخ نرم و کشسانی که جرم آن در طول نخ به‌صورتی پیوسته تقسیم شده باشد. به‌ویژه، وقتی از سیم همگن صحبت می‌کنیم. منظورمان سیمی است که تراکم و چگالی آن ثابت باشد.

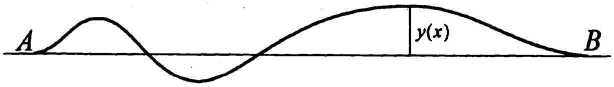
از آن‌جا که جرم در طول سیم به‌صورتی پیوسته پراکنده است، بنابراین موقعیت سیم را نمی‌توان با تعداد محدودی عدد y_1 ، y_2 ، ... ، y_n داد، بلکه باید انحراف $y(x)$ را برای هر نقطهٔ x از سیم معین کرد. به‌این‌ترتیب، موقعیت سیم در هر لحظهٔ زمانی به‌وسیلهٔ تابعی مثل $y(x)$ داده می‌شود.

موقعیت نخی که با n وزنه در نقطه‌های با طول‌های x_1 ، x_2 ، ... ، x_n داده شده است، به‌صورت نموداری، به‌وسیلهٔ خط شکسته‌ای با $n+1$ پاره‌خط راست داده می‌شود (شکل ۴). اگر تعداد وزنه‌ها بیشتر شود، متناظر با آن، تعداد پاره‌خط‌های خط شکسته هم زیادتر می‌شود. اگر تعداد وزنه‌ها، به‌طور نامحدود، زیاد کنیم، فاصلهٔ بین وزنه‌ها به‌سمت صفر میل می‌کند و در نتیجه، در حد، موقعیت پیوستهٔ جرم در طول نخ، یعنی سیم ایده‌آل، به‌دست می‌آید. در ضمن، خط شکسته‌ای که موقعیت نخ را با وزنه‌ها نشان می‌داد، به‌صورت یک منحنی درمی‌آید که معرف موقعیت سیم است (شکل ۵).

به‌این‌ترتیب می‌بینیم بین مسأله‌های مربوط به‌نوسان‌های نخی که همراه با وزنه‌هاست، و نوسان‌های سیم، رابطه‌های جدی وجود دارد. در حالت اول، موقعیت دستگاه، به‌وسیلهٔ



شکل ۴



شکل ۵

یک نقطه یا یک بردار در فضای n بعدی داده می‌شود. بنابراین طبیعی است، تابع $f(x)$، که موقعیت سیم در حال نوسان را در مسألهٔ دوم مشخص می‌کند، بازهم بوسیلهٔ بردار یا نقطه‌ای در یک فضای بی‌نهایت بعدی در نظر گرفته می‌شود. مسأله‌های مشابه دیگری هم وجود دارد، که به‌همین اندیشه می‌رسد، یعنی فضایی که نقطه‌های (یا بردارهای) آن تابع $f(x)$ هستند که در فاصلهٔ معینی داده شده‌اند[1].

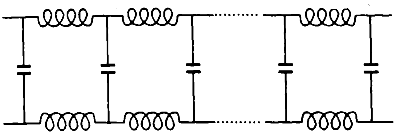
نمونه مسأله‌ای را که در این‌جا دربارهٔ نوسان‌ها بررسی کردیم، و دوباره در بند ۴ به‌آن برخواهیم گشت، تا حدی روشن می‌کند که، چگونه باید با مفهوم‌های اساسی فضای بی‌نهایت بعدی برخورد کرد.

فضای هیلبرت. در این جا، به بررسی مفهوم یکی از فضاهای بی‌نهایت بعدی که از نظر کاربردی مهم و بیش از نمونه‌های دیگر معمول شده است، یعنی مفهوم فضای هیلبرت می‌پردازیم.

بردار فضای n بعدی، به‌عنوان مجموعه‌ای از n عدد f_i تعریف می‌شود که در آن، i از ۱ تا

---

۱. به‌عنوان مسألهٔ دیگری از این گونه، نوسان‌های الکتریکی را در نظر می‌گیریم که در رشته‌ای پدید می‌آیند که مدارهای الکتریکی را به‌هم وصل کرده است (شکل ۶).



شکل ۶

حالت چنین رشته‌ای با انتخاب n عدد u_1، u_2، ...، u_n بیان می‌شود که در آن، u_i عبارت است از ولتاژ در خازن iام مدار. رشته مجموعهٔ n عدد

$$(u_1, u_2, \dots, u_n)$$

تشکیل برداری از یک فضای n بعدی را می‌دهد.

اکنون یک خط دو سیمی در نظر می‌گیریم، یعنی خطی که شامل دو سیمی باشد که ظرفیت و خودالقایی محدودی دارند که در طول سیم پراکنده باشند. حالت الکتریکی خط با تابعی مثل $u(x)$ بیان می‌شود که پراکندگی ولتاژ در طول خط را می‌دهد. این تابع، برداری از فضای بی‌نهایت بعدی تابع‌هاست که در فاصلهٔ (a , b) داده شده‌اند.

n تغییر می‌کند. به‌همین‌ترتیب، بردار فضای بی‌نهایت بعدی به‌عنوان تابع $f(x)$ تعریف می‌شود که در آن، x از a تا b تغییر می‌کند.

بنابراین، بردارها و ضرب بردارها در عدد، همچون تابع‌ها و ضرب تابع‌ها در عدد تعریف می‌شود.

طول بردار f در فضای n بعدی با دستور $\sqrt{\sum_{i=1}^{n} f_i^2}$ بیان می‌شود. از آن جا که در تابع‌ها، نقش مجموع به‌عهدهٔ انتگرال است، بنابراین طول بردار $f(x)$ در فضای هیلبرت، با این دستور داده می‌شود:

$$\sqrt{\int_a^b f^2(x)\,dx} \tag{۵}$$

فاصلهٔ بین دو نقطهٔ f و g در فضای n بعدی، به‌عنوان طول بردار $f-g$، یعنی به‌عنوان $\sqrt{\sum_{i=1}^{n}(f_i-g_i)^2}$ تعریف می‌شود. به‌همین‌ترتیب «فاصلهٔ» بین عضوهای $f(t)$ و $g(t)$ در فضای تابعی برابر است با:

$$\sqrt{\int_a^b [f(t)-g(t)]^2 dt}$$

عبارت $\int_a^b [f(t)-g(t)]^2$ را، میانگین مربع انحراف تابع‌های $f(t)$ و $g(t)$ گویند. بنابراین، مقیاس فاصلهٔ دو عضو فضای هیلبرت، عبارت است از میانگین مربع انحراف آن‌ها.

به‌تعریف زاویهٔ بین دو بردار می‌پردازیم. در فضای n بعدی، زاویه φ، بین بردارهای $f=\{f_i\}$ و $g=\{g_i\}$ با این دستور به‌دست می‌آید:

$$\cos\varphi=\frac{\sum_{i=1}^{n} f_i g_i}{\sqrt{\sum_{i=1}^{n} f_i^2}\sqrt{\sum_{i=1}^{n} g_i^2}}$$

در فضای بی‌نهایت بعدی، مجموع‌ها به‌انتگرال‌های متناظر خود تبدیل می‌شوند و زاویهٔ φ، بین دو بردار f و g در فضای هیلبرتی با دستور

$$\cos\varphi = \frac{\int_a^b f(t)g(t)dt}{\sqrt{\int_a^b f^2(t)dt}\sqrt{\int_a^b g^2(t)dt}} \qquad (۶)$$

بیان می‌شود. چنین عبارتی را وقتی می‌توان کسینوس زاویه‌ای مثل φ به‌شمار آورد که، کسر سمت راست برابری، از لحاظ قدر مطلق کوچکتر از واحد باشد، یعنی وقتی که

$$\left| \int_a^b f(t)g(t)\,dt \right| \leq \sqrt{\int_a^b f^2(t)\,dt}\sqrt{\int_a^b g^2(t)\,dt} \qquad (۷)$$

و این نابرابری، در واقع، برای هر دو تابع $f(t)$ و $g(t)$ برقرار است. این نابرابری نقش مهمی در آنالیز دارد و به «نابرابری *کوشی ـ بونیا کووسکی*» معروف است. اثبات درستی این نابرابری را می‌آوریم.

دو تابع $f(x)$ و $g(x)$ را، که متحد با صفر نیستند، در فاصلهٔ $(a\,, b)$ در نظر می‌گیریم. عددهای دل‌خواه λ و μ را انتخاب می‌کنیم و این عبارت را تشکیل می‌دهیم:

$$\int_a^b [\lambda f(x) - \mu g(x)]^2\, dx$$

چون تابع $[\lambda f(x) - \mu g(x)]^2$ که در زیر علامت انتگرال قرار دارد، غیر منفی است، بنابراین

$$\int_a^b [\lambda f(x)-\mu g(x)]^2 dx \geq 0$$

یعنی

$$2\lambda\mu \int_a^b f(x)g(x)dx \leq \lambda^2 \int_a^b f^2(x)\,dx + \mu^2 \int_a^b g^2(x)dx$$

برای سادگی کار، این نمادها را می‌پذیریم:

$$\left| \int_a^b f(x)g(x)dx \right| = C,\ \int_a^b f^2(x)\,dx = A,\ \int_a^b g^2(x)\,dx = B \qquad (۸)$$

با این نمادها، نابرابری به‌این صورت درمی‌آید[۱]:

---

۱. به خاطر دل‌خواه بودن علامت λ یا μ می‌توانیم به‌عنوان C، قدر مطلق انتگرال را انتخاب کنیم.

$$2\lambda\mu C \leq \lambda^2 A + \mu^2 B \tag{9}$$

این نابرابری برای هر مقدار λ و μ درست است؛ بنابراین می‌توان فرض کرد:

$$\lambda=\sqrt{\frac{C}{A}}\ ,\ \mu=\sqrt{\frac{C}{B}} \tag{10}$$

که اگر این مقدارهای λ و μ را در نابرابری (۹) قرار می‌دهیم، به‌دست می‌آید:

$$\frac{C}{\sqrt{AB}} \leq 1$$

و اگر به‌جای A، B و C، مقدارشان را از دستورهای (۸) قرار دهیم، سرانجام به‌نابرابری **کوشی - بونیاکووسکی** می‌رسیم.

در هندسه، حاصل‌ضرب اسکالر بردارها، به‌عنوان حاصل‌ضرب طول آن‌ها در کسینوس زاویهٔ بین آن‌ها تعریف می‌شود. طول بردارهای f و g، در این حالت، برابرند با

$$\sqrt{\int_a^b f^2(x)dx}\ ,\ \sqrt{\int_a^b g^2(x)dx}$$

کسینوس زاویهٔ بین آن‌ها هم، با این دستور بیان می‌شود:

$$\cos\varphi = \frac{\int_a^b f(x)\ g(x)\,dx}{\sqrt{\int_a^b f^2(x)dx}\ ,\ \sqrt{\int_a^b g^2(x)dx}}$$

اگر این عبارت‌ها را در هم ضرب کنیم، به‌دستوری می‌رسیم که معرف حاصل‌ضرب اسکالر دو بردار در فضای هیلبرت است:

$$(f\ ,\ g)=\int_a^b f(x)g(x)dx \tag{11}$$

از این دستور نتیجه می‌شود که حاصل‌ضرب اسکالر بردار f در خودش، برابر توان دوم طول آن است.

اگر حاصل‌ضرب اسکالر بردارهای غیر صفر f و g برابر صفر باشد، به‌این معناست که $\cos\varphi = 0$، یعنی زاویهٔ φ با تعریفی که برای آن کردیم، برابر ۹۰ درجه است. بنابراین، اگر برای تابع f و g داشته باشیم:

$$(f\,;\,g)=\int_a^b f(x)g(x)dx=0$$

گویند این تابع‌ها متعامدند (یا برهم عمودند یا ارتوگونال‌اند).

در فضای هیلبرت هم، مثل فضای n بعدی، قضیهٔ **فیثاغورس** درست است (بند ۱ را ببینید). اگر $f_1(x)$، $f_2(x)$، ...، $f_n(x)$ و N تابع دو به‌دو متعامد باشند و

$$f(x)=f_1(x)+f_2(x)+\cdots+f_N(x)$$

مجموع آن‌ها باشد، آن وقت توان دوم طول f برابر است با مجموع توان‌های دوم طول‌های f_1، f_2، ...، f_N.

چون در فضای هیلبرت، طول بردارها به‌یاری انتگرال‌ها داده می‌شود، بنابراین قضیهٔ فیثاغورس، در این حالت، این طور بیان می‌شود:

$$\int_a^b f^2(x)dx=\int_a^b f_1^2(x)dx+\int_a^b f_2^2(x)dx+\cdots+\int_a^b f_N^2(x)\,dx \qquad (12)$$

اثبات این قضیه، هیچ تفاوتی با آن چه در بند ۱، برای اثبات همین قضیه در فضای n بعدی آوردیم، ندارد.

هنوز روشن نکرده‌ایم، کدام تابع‌ها را باید به‌عنوان بردارهایی از فضای هیلبرت به‌حساب آورد. همهٔ تابع‌هایی که برای آن‌ها انتگرال $\int_a^b f^2(x)dx$ معنا دارد، از این‌گونه‌اند. به نظر می‌رسد، باید خود را به تابع‌های پیوسته‌ای که، برای آن‌ها $\int_a^b f^2(x)dx$ همیشه وجود دارد، محدود کرد. ولی درواقع، نظریهٔ فضای هیلبرت وقتی کامل می‌شود که انتگرال $\int_a^b f^2(x)dx$ را به‌مفهوم کلی آن، یعنی به‌مفهوم انتگرال **له‌بک** در نظر بگیریم (بخش پانزدهم را ببینید).

این تعمیم مفهوم انتگرال (و متناظر با آن تعمیم رده تابع‌هایی که بررسی می‌کنیم)، برای آنالیز تابعی و هم برای پایه‌گذاری محاسبه‌های دیفرانسیلی و انتگرالی و برای محکم‌تر کردن نظریهٔ عددهای حقیقی ضرورت دارد. به‌این‌ترتیب، تعمیم مفهوم عادی انتگرال، که در آغاز سدهٔ بیستم و در رابطه با پیشرفت نظریهٔ تابع‌های، متغیر حقیقی پدید آمد، برای آنالیز تابعی و شاخه‌هایی از ریاضیات که به‌آن مربوط‌اند، بسیار اهمیت دارد.

۳. تجزیه به‌دستگاه تابع‌های متعامد (ارتوگونال)

تعریف دستگاه تابع‌های متعامد و نمونه‌های آن. اگر دو بردار e_1 و e_2 را به‌طول واحد، روی صفحه و عمود برهم انتخاب کنیم. (شکل ۷)، آن وقت هر بردار واقع بر این صفحه را می‌توان در جهت این دو بردار تجزیه کرد، یعنی به‌این صورت بیان کرد:

$$f = a_1 e_1 + a_2 e_2$$

که در آن a_1 و a_2 عددهایی هستند برابر با طول تصویرهای بردار f روی محورهای موازی و هم جهت با e_1 و e_2. چون تصویر f روی محور، برابر است با حاصل‌ضرب طول f در کسینوس زاویه‌ای که f با محور می‌سازد، بنابراین با به‌یاد آوردن تعریف حاصل‌ضرب اسکالر، می‌توان نوشت:

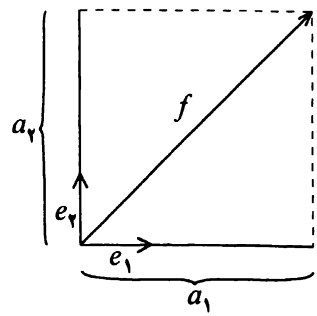
$$a_1 = (f,\ e_1)\ ,\ a_2 = (f,\ e_2)$$

به‌همین‌ترتیب، اگر سه بردار واحد e_1، e_2 و e_3 را که دو به‌دو متعامدند، در فضای سه بعدی انتخاب کنیم، آن وقت می‌توان بردار دل‌خواه f در این فضا را به‌صورت

$$f = a_1 e_1 + a_2 e_2 + a_3 e_3$$

نشان داد که در آن داشته باشیم:

$$a_k = (f,\ e_k) \quad (k = 1,\ 2,\ 3)$$



شکل ۷

در فضای هیلبرت هم می‌توان دستگاه و بردارهای دو به دو متعامد در این فضا، یعنی $\varphi_1(x)$ ، $\varphi_2(x)$ ، ... ، $\varphi_n(x)$ ، ...را در نظر گرفت.

چنین دستگاه‌هایی از تابع‌ها را، دستگاه‌های تابع‌های متعامد (یا اُرتوگونال یا عمود برهم) می‌نامند که در آنالیز نقش عمده‌ای به‌عهده دارند، به‌این گونه دستگاه‌ها در بسیاری از مسأله‌های مربوط به‌فیزیک‌ریاضی، معادله‌های انتگرالی، محاسبه‌های تقریبی، نظریهٔ تابع‌های با متغیر حقیقی و غیره، برخورد می‌کنیم. تنظیم و به‌هم‌پیوستن همهٔ مفهوم‌هایی که به‌این گونه دستگاه‌ها مربوط می‌شد، یکی از انگیزه‌های ریاضی‌دانان آغاز سدهٔ بیستم، برای ساختن مفهوم کلی فضای هیلبرت بود.

تعریف‌های دقیق را می‌آوریم. دستگاه تابع‌های

$$\varphi_1(x)\ ,\ \varphi_2(x)\ ,\ \dots\ ,\ \varphi_n(x)\ ,\ \dots$$

را وقتی ***متعامد*** گویند که هر دو تابع دل‌خواه از این دستگاه نسبت به‌هم متعامد باشند، یعنی وقتی که داشته باشیم:

$$\int_a^b \varphi_i(x)\varphi_k(x)dx = 0 \qquad (i \neq k) \tag{۱۳}$$

در فضای سه‌بعدی، شرط کردیم که بردارهای دستگاه، طولی برابر واحد داشته باشند. با به یاد آوردن تعریف طول بردار، می‌بینیم این شرط در حالت فضای هیلبرت، به این صورت نوشته می‌شود:

$$\int_a^b \varphi^2{}_k(x)dx = 1 \tag{۱۴}$$

دستگاهی از تابع که با شرط‌های (۱۳) و (۱۴) سازگار باشد، دستگاه یکا متعامد (ارتونرمال) نامیده می‌شود.

مثال‌هایی از این گونه دستگاه‌ها می‌آوریم:

۱. این دنبالهٔ تابع‌ها را دربازهٔ $(-\pi\ ,\ \pi)$ در نظر می‌گیریم:

$$1\ ,\ \cos x\ ,\ \sin x\ ,\ \cos 2x\ ,\ \sin 2x\ ,\ \dots\ ,\ \cos nx\ ,\ \sin nx\ ,\ \dots$$

هر دو تابع از این دنباله، نسبت به‌هم متعامدند. درستی این حکم، به‌سادگی و با محاسبهٔ انتگرال‌های متناظر روشن می‌شود. توان دوم طول بردار در فضای هیلبرت، عبارت است از

انتگرال توان دوم تابع. بنابراین، توان دوم طول بردارها در دنبالهٔ

$$1\,,\ \cos x\,,\ \sin x\,,\ \cos 2x\,,\ \sin 2x\,,\ \dots\,,\ \cos nx\,,\ \sin nx\,,\ \dots$$

عبارت است از این انتگرال‌ها:

$$\int_{-\pi}^{\pi} dx = 2\pi\ ,\ \int_{-\pi}^{\pi} \cos^2 nx dx = \pi\ ,\ \int_{-\pi}^{\pi} \sin^2 nx dx = \pi$$

یعنی دنبالهٔ بردارهای ما، دنباله‌ای متعامد (ارتوگونال) است، ولی یکا متعامد (نرمال) نیست. طول نخستین بردار دنباله برابر $\sqrt{2\pi}$ است، در حالی که هریک از بردارهای دیگر دنباله، طولی برابر $\sqrt{\pi}$ دارد. اگر هریک از بردارها را بر طول آن تقسیم کنیم، دستگاهی یکا متعامد از تابع‌های مثلثاتی به‌دست می‌آید:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \dots$$

این دستگاه، از نظر تاریخی، یکی از نخستین و مهم‌ترین نمونه‌ها از دستگاه‌های متعامد است. این دستگاه در کارهای *اولر*، *د.برنولی* و *دالامبر* در رابطه با مسألهٔ مربوط به نوسان‌های سیم پدید آمد. بررسی این دستگاه، در پیشرفت تمامی آنالیز، نقشی جدی داشت (بند ۱ را در بخش دوازدهم در جلد دوم ببینید).

پدید آمدن دستگاه متعامد تابع‌های مثلثاتی در رابطه با مسألهٔ نوسان‌های سیم، تصادفی نبود. هر مسألهٔ مربوط به نوسان‌های کوچک محیط منجر به دستگاهی از تابع‌های متعامد می‌شد که، به اصطلاح، پاسخ‌گوی نوسان‌های خاص دستگاه مفروض بود (بند ۴ را ببینید). از جمله، در رابطه با مسألهٔ مربوط به نوسان‌های کره، تابع‌های کروی، در رابطه با مسألهٔ مربوط به نوسان‌های استوانه، تابع‌های استوانه‌ای پدید می‌آید، و غیره.

۲. می‌توان مثالی از دستگاه متعامد تابع‌ها آورد، به نحوی که هر تابع یک چندجمله‌ای باشد. این مثال عبارت است از دنبالهٔ چندجمله‌ای‌های *لژاندر*

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n}$$

یعنی (با دقت تا ضریب ثابت)، $P_n(x)$ عبارت است از مشتق مرتبهٔ nام از $(x^2 - 1)^n$. چند

جملهٔ اول این دنباله را می‌نویسیم،

$$P_0(x)=1\ ;$$

$$P_1(x)=x\ ;$$

$$P_2(x)=\frac{1}{2}(3x^2-1)\ ;$$

$$P_3(x)=\frac{1}{2}(5x^3-3x)$$

روشن است که، به طور کلی $P_n(x)$ یک چندجمله‌ای از درجهٔ n است. به خواننده پیشنهاد می‌کنیم، خود به تحقیق این مطلب بپردازد که، این چندجمله‌ای‌ها، در بازهٔ $(-1\ ,\ 1)$، دنباله‌ای متعامد را تشکیل می‌دهند.

نظریهٔ کلی چندجمله‌ای‌های متعامد را، پ. ل چیشِف، ریاضی‌دان بزرگ روس، در نیمهٔ دوم سدهٔ نوزدهم تکامل داد.

تجزیه به وسیلهٔ دستگاه متعامد تابع‌ها. به همان ترتیب که در فضای سه‌بعدی، می‌توان هر بردار را به صورت ترکیبی خطی از سه بردار واحد و دو به دو متعامد e_1 ، e_2 ، e_3 نمایش داد:

$$f=a_1e_1+a_2e_2+a_3e_3$$

در فضای تابعی هم این مسأله مطرح می‌شود که، تابع دل‌خواه f را به یک رشته تابع‌های یکا متعامد تجزیه کنیم، یعنی تابع f را به این صورت درآوریم.

$$f(x)=a_1\varphi_1(x)+a_2\varphi_2(x)+\cdots+a_n\varphi_n(x)+\cdots \qquad (15)$$

در ضمن، هم‌گرایی رشتهٔ (۱۵) به سمت تابع f ، به مفهوم فاصلهٔ بین عضوها در فضای هیلبرت گرفته می‌شود، و این به معنای آن است که میانگین مربع انحراف مجموع جزئی رشته، یعنی

$$S_n(t)=\sum_{k=1}^{n}a_k\varphi_k(t)$$

از تابع $f(t)$، وقتی $n\to\infty$، به سمت صفر میل کند، یعنی

$$\lim_{n\to\infty}\int_a^b[f(t)-S_n(t)]^2dt=0 \qquad (16)$$

این هم‌گرایی را «هم‌گرایی در میانگین» می‌نامند.

تجزیهٔ به دستگاهی از تابع‌های متعامد اغلب در آنالیز پیش می‌آید و روشی اساسی برای حل مسأله‌های فیزیک ریاضی است. از جمله، اگر دستگاه متعامد، دستگاهی از تابع‌های مثلثاتی دربازهٔ $(\pi\ ,\ -\pi)$ باشد:

$$1\ ,\ \cos x\ ,\ \sin x\ ,\ \cos 2x\ ,\ \sin 2x\ ,\ \dots,\ \cos nx\ ,\ \sin nx\ ,\ \dots$$

آن وقت، چنین تجزیه‌ای، همان تجزیهٔ تابع‌ها به رشتهٔ مثلثاتی است[1]:

$$f(x) = a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + b_2\sin 2x + \cdots$$

فرض می‌کنیم، تجزیه (۱۵)، برای هر تابع f از فضای هیلبرت، ممکن باشد و ضریب‌های a_n را در این تجزیه پیدا می‌کنیم. برای این منظور، دو طرف برابری را، به صورت اسکالر، در تابعی مثل φ_m از خود دستگاه، ضرب می‌کنیم. به این برابری می‌رسیم:

$$(f\ ,\varphi_m) = a_1(\varphi_1\ ,\ \varphi_m) + a_2(\varphi_2\ ,\ \varphi_m) + \cdots + a_m(\varphi_m\ ,\ \varphi_m) + a_{m+1}(\varphi_{m+1}\ ,\ \varphi_m) + \cdots$$

با توجه به این که، برای $m \neq n$ داریم $(\varphi_m\ ,\ \varphi_n) = 0$ و $(\varphi_m\ ,\ \varphi_m) = 1$، ضریب a_m به دست می‌آید:

$$a_m = (f\ ,\ \varphi_m)\ ,\ (m = 1\ ,\ 2\ ,\ \dots)$$

می‌بینیم، مثل حالت فضای سه‌بعدی (آغاز همین بند را ببینید)، ضریب a_m برابر است با تصویر f روی جهت‌های بردارهای φ_k.

با توجه به تعریف ضرب اسکالر، می‌توانیم ضریب‌های تجزیهٔ تابع $f(x)$ را به تابع‌های یکا متعامد $\varphi_1(x)$ ، $\varphi_2(x)$ ، ... ، $\varphi_n(x)$ ، ... به دست آوریم:

$$f(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x) + \cdots \qquad (17)$$

این محاسبه، به یاری این دستور انجام می‌گیرد:

$$a_m = \int_a^b f(t)\varphi_m(t)dt \qquad (18)$$

---

۱. در مسأله‌های مختلفی از فیزیک و ضمن تجزیه نوسان‌ها به مولفه‌های همساز (هارمونیک)، به چنین تجزیه‌هایی برخورد می‌کنیم. بند ۵ از بخش ششم (جلد دوم) را ببینید.

برای نمونه، دستگاه تابع‌های مثلثاتی یکا متعامد را در نظر می‌گیریم:

$$\frac{۱}{\sqrt{۲\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos ۲x}{\sqrt{\pi}}, \frac{\sin ۲x}{\sqrt{\pi}}, \ldots$$

در این صورت داریم:

$$f(x)=\frac{a_{۰}}{۲}+\sum_{n=۱}^{\infty} a_n \cos nx + b_n \sin nx$$

که در آن

$$a_{۰}=\frac{۱}{\pi}\int_{-\pi}^{\pi} f(x)dx\ ,\ a_n=\frac{۱}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx\ ,$$

$$b_n=\frac{۱}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx$$

به این ترتیب، دستور محاسبهٔ ضریب‌های تجزیهٔ تابع به رشتهٔ مثلثاتی، البته به شرطی که این تجزیه ممکن باشد، به دست می‌آید[۱].

توانستیم ضریب‌های تجزیه تابع $f(x)$ را به دستگاه متعامد تابع‌ها، با فرض ممکن بودن این تجزیه، به صورت (۱۸) پیدا کنیم. ولی دستگاهی از بی‌نهایت تابع متعامد $\varphi_۱$، $\varphi_۲$، ... ، φ_n، ...، ممکن است، برای این که بتوان هر تابع از فضای هیلبرت را تجزیه کرد، کافی نباشد. برای این که این تجزیه ممکن باشد، باید دستگاه تابع‌های متعامد، با شرطی اضافی، که شرط کامل بودن نامیده می‌شود. سازگار باشد.

دستگاه متعامد تابع‌ها را *کامل* گویند، وقتی که نتوان به آن، ولو یک تابع غیر صفر هم اضافه کرد، که با همهٔ تابع‌های دستگاه متعامد باشد.

به سادگی می‌توان مثالی از دستگاه متعامد پیدا کرد که کامل نباشد. برای این منظور می‌توان یک دستگاه متعامد، از جمله دستگاه تابع‌های مثلثاتی را در نظر گرفت و سپس، یکی از تابع‌های آن را کنار گذاشت، مثل تابع $\cos x$ را. به این دستگاه که شامل بی‌نهایت تابع است، می‌رسیم:

$$۱\ ,\ \sin x\ ,\ \cos ۲x\ ,\ \sin ۲x\ ,\ \ldots,\ \cos nx\ ,\ \sin nx\ ,\ \ldots$$

که متعامد بودن خود را حفظ می‌کند، ولی البته کامل نیست، زیرا تابع $\cos x$ را که حذف

۱. دربارهٔ رشته‌های مثلثاتی، بند ۷ از بخش دوازدهم (جلد دوم) را هم ببینید.

کرده‌ایم، نسبت به همهٔ تابع‌های دستگاه متعامد است.

اگر دستگاهی از تابع‌ها کامل نباشد، آن وقت نمی‌توان هر تابعی از فضای هیلبرت را نسبت به آن‌ها تجزیه کرد. در واقع، اگر تابع غیرصفر $f_0(x)$، را که نسبت به همهٔ تابع‌های دستگاه متعامد باشد، به یاری این دستگاه تجزیه کنیم، با توجه به دستور (۱۸)، همهٔ ضریب‌ها صفر می‌شوند. در حالی که تابع $f_0(x)$ برابر صفر نیست.

این قضیه درست است: اگر دستگاه یکا متعامد کامل تابع‌های $\varphi_1(x)$ ، $\varphi_2(x)$ ، .... ، $\varphi_n(x)$ ، ... در فضای هیلبرت مفروض باشد، آن وقت هر تابع $f(x)$ را می‌توان به رشته‌ای برحسب تابع‌های این دستگاه و تجزیه کرد[1]:

$$f(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x) + \cdots$$

در ضمن، ضریب a_n این تجزیه برابر است با تصویر بردارهای f روی عضوهای دستگاه یکا متعامد:

$$a_n = (f,\ \varphi_n) = \int_a^b f(x)\varphi_n(x)dx$$

به یاری قضیهٔ فیثاقورس که در بند ۲ برای فضای هیلبرت آوردیم، می‌توان رابطهٔ جالبی بین ضریب‌های a_k و تابع $f(x)$ پیدا کرده $r_n(x)$ را اختلاف بین $f(x)$ و مجموع n جملهٔ اول رشتهٔ آن می‌گیریم، یعنی

$$r_n(x) = f(x) - [a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)]$$

تابع $r_n(x)$، نسبت به $\varphi_1(x)$ ، $\varphi_2(x)$ ، .... ، $\varphi_n(x)$ متعامد است. برای نمونه، متعامد بودن آن را نسبت به $\varphi_1(x)$ آزمایش می‌کنیم، یعنی ثابت می‌کنیم:

$$\int_a^b r_n(x)\varphi_1(x)dx = 0$$

داریم[2]:

$$\int_a^b r_n(x)\varphi_1(x)dx = \int_a^b [f(x) - a_1\varphi_1(x) - a_2\varphi_2(x) - \cdots - a_n\varphi_n(x)]\varphi_1(x)dx =$$
$$= \int_a^b f(x)\varphi_1(x)dx - a_1\int_a^b \varphi_1^2(x)dx^2$$

---

۱. این رشته نسبت به مجموع خود، به مفهومی وابسته است که با دستور (۱۶) مشخص شده است.

۲. انتگرال‌های از نوع $\int_a^b \varphi_1(x)\varphi_k(x)dx\ (k\neq 1)$ برابر صفرند، زیرا تابع $\varphi_k(x)$ نسبت به تابع $\varphi_1(x)$ متعامد است.

چون

$$a_1 = \int_a^b f(x)\varphi_1(x)dx, \int_a^b \varphi_1^2(x)dx = 1$$

بنابراین، نتیجه می‌شود: $\int_a^b r_n(x)\varphi_1(x)dx = 0$.

به این ترتیب، در برابری

$$f(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_n\varphi_n(x) + r_n(x) \quad (19)$$

جمله‌های جداگانه در سمت راست، نسبت به هم متعامدند. و این، با توجه به قضیهٔ فیثاقورس که در بند ۲ آوردیم، به معنای آن است که توان دوم طول بردار $f(x)$برابر است با مجموع توان‌های دوم طول‌های جمله‌های سمت راست برابری (۱۹). یعنی

$$\int_a^b f^2(x)dx = \int_a^b [a_1\varphi_1(x)]^2dx + \cdots + \int_a^b [a_n\varphi_n(x)]^2dx + \int_a^b r_n^2(x)dx$$

چون دستگاه تابع‌های φ_1 ، φ_2 ، .... ، φ_n ، ... ، یکا متعامدند [برابری (۱۴)]، بنابراین

$$\int_a^b f^2(x)dx = a_1^2 + a_2^2 + \cdots + a_n^2 + \int_a^b r_n^2(x)dx \quad (20)$$

رشتهٔ $\sum_{k=1}^{\infty} a_k\, \varphi_k(x)$ در میانگین هم‌گراست. و این به معنای آن است که

$$\int_a^b [f(x) - a_1\varphi_1(x) - \cdots - a_n\varphi_n(x)]^2dx \to 0$$

و از آن‌جا

$$\int_a^b r^2{}_n(x)dx \to 0$$

ولی در این صورت، از دستور (۲۰) به دست می‌آید[1]:

$$\sum_{k=1}^{\infty} a_k^2 = \int_a^b f^2(x)dx \quad (21)$$

۱. دستور (۲۱)، از نظر هندسی به این معنا است که، توان دوم طول یک بردار در فضای هیلبرت، برابر است با مجموع توان‌های دوم تصویرهای آن روی دستگاهی کامل از جهت‌های دو به دو عمود برهم.

این برابری حاکی از آن است که، انتگرال توان دوم تابع برابر است با مجموع توان‌های دوم ضریب‌های تجزیهٔ آن برحسب دستگاه بسته و متعامد تابع‌ها. شرط (۲۱)، اگر برای هر تابع از فضای هیلبرت وجود داشته باشد، شرط «کامل بودن» نامیده می‌شود.

به این موقعیت مهم هم توجه کنیم. کدام عددهای a_k می‌توانند ضریب‌های تجزیهٔ تابعی از فضای هیلبرت باشند؟ برابری (۲۱) روشن می‌کند که، برای این منظور، باید رشتهٔ $\sum_{k=1}^{\infty} a_k^2$ هم‌گرا باشد. این شرط، کافی هم هست. یعنی برای این که دنبالهٔ عددهای a_k، دنبالهٔ ضریب‌های تجزیه نسبت به تابع‌های متعامد از فضای هیلبرت باشند، لازم و کافی است که رشته $\sum_{k=1}^{\infty} a_k^2$ هم‌گرا باشد.

یادآوری می‌کنیم، این قضیه، برای حالتی هم که فضای هیلبرت، به عنوان مجموعه‌ای از همهٔ تابع‌هایی در نظر گرفته شود که توان دوم آن‌ها، به مفهوم له‌بگ انتگرال‌پذیرید (بند ۲ را ببینید)، درست است. اگر در فضای هیلبرت خود را تنها در تابع‌های پیوسته محدود کنیم، آن وقت، در پاسخ به این پرسش که کدام عددهای a_k می‌توانند ضریب‌های تجزیه باشند، باید گفت که هیچ شرطی لازم نیست.

ملاحظه‌هایی را که در این جا آوردیم، تنها بخشی از دلیل‌هایی است که ضرورت استفاده از انتگرال‌ها به مفهوم کلی آن‌ها (یعنی به مفهوم لِه‌بگ) را روشن می‌کند.

۴. معادله‌های انتگرالی

در این بند، با یکی از مهم‌ترین و از نظر تاریخی یکی از نخستین بخش‌های آنالیز تابعی، یعنی نظریهٔ معادله‌های انتگرالی آشنا می‌شویم. نظریه‌ای که نقشی جدی در پیشرفت بعدی آنالیز تابعی داشته است. پیشرفت نظریهٔ معادله‌های انتگرالی، به‌جز نیازهای درونی ریاضیات [مثل مسأله‌های مرزی برای معادله‌های با مشتق‌های جزئی (جلد دوم، بخش ششم)]، در مسأله‌های مختلف فیزیک هم، نقش اساسی داشت. در کنار معادله‌های دیفرانسیلی، در سدهٔ بیستم، معادله‌های انتگرالی هم، ابزاری جدی برای بررسی‌های ریاضی مسأله‌های گوناگون فیزیک به شمار می‌رود. در این بند، تلاش می‌کنیم آگاهی‌هایی از نظریهٔ معادله‌های انتگرالی به خواننده بدهیم. آگاهی‌هایی که در این جا مطرح می‌کنیم،

رابطه‌ای کامل با مطالعهٔ نوسان‌های دستگاه‌های کشسان دارند و تا حد زیادی (مستقیم یا غیرمستقیم) در بستگی با آن‌ها پدید آمده‌اند.

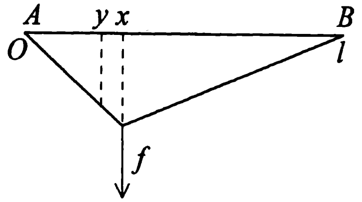
مسألهٔ مربوط به نوسان‌های کوچک دستگاه‌های کشسان. به مسألهٔ نوسان‌های کوچک که در بند ۲ بررسی کردیم، برمی‌گردیم. معادله‌هایی را پیدا می‌کنیم که این‌گونه نوسان‌ها را توضیح می‌دهند. برای سادگی بحث، فرض می‌کنیم با نوسان‌های دستگاه کشسان خطی سروکار داشته باشیم. به عنوان نمونهٔ این دستگاه، می‌توان سیمی به طول l (شکل ۸) یا میله‌ای کشسان (شکل ۹) را در نظر گرفت. فرض می‌کنیم، دستگاه کشسان ما در حالت تعادل، روی پاره‌خط راست Ol از محور Ox واقع باشد. به نقطهٔ x نیروی واحدی وارد می‌کنیم. زیر تأثیر این نیرو، همهٔ نقطه‌های دستگاه دچار انحراف می‌شوند. انحرافی را که در نقطه y پدید می‌آید (شکل ۱۸)، $k(x , y)$ می‌نامیم.

تابع $k(x , y)$، تابعی از دو نقطه است: نقطهٔ x که نیرو بر آن وارد شده است، و نقطه y که میزان انحراف آن را اندازه می‌گیریم. این تابع را، تابع اثر (یا تابع گرین) می‌نامیم.

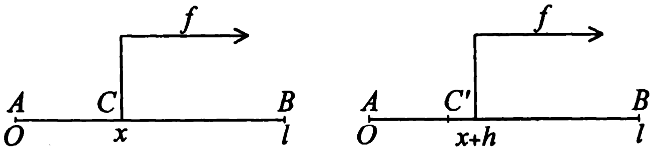
با توجه به قانون بقای انرژی، می‌توان ویژگی مهم تابع اثر $k(x , y)$ را نتیجه گرفت، به اصطلاح قانون متقابل: انحرافی که در نقطهٔ y زیر تأثیر نیروی وارد بر نقطهٔ x پدید می‌آید، برابر است با انحرافی که در نقطهٔ x زیر تأثیر همان نیرو، وقتی به نقطهٔ y وارد می‌شد به زبان دیگر، این قانون به این معنا است که

$$k(x , y) = k(y , x) \tag{۲۲}$$

برای نمونه، تابع تأثیر را برای نوسان‌های طولی میلهٔ کشسان پیدا می‌کنیم (روی شکل ۸ چیز دیگری نشان داده شده است: عرضی، جابه‌جایی). میلهٔ AB به طول l را در نظر می‌گیریم که در دو انتهای خود محکم شده باشد (شکل ۹). در نقطهٔ C نیروی f را قرار



شکل ۸



شکل ۹

می‌دهیم که در جهت B عمل می‌کند. با اثر گذاشتن این نیرو، میله تغییر شکل می‌دهد و نقطهٔ C به نقطهٔ C' می‌رود. مقدار جابه جایی نقطهٔ C را h می‌نامیم. در آغاز h را پیدا می‌کنیم. به‌کمک h می‌توانیم جابه جایی y را به هر نقطه پیدا کنیم. برای این منظور از قانون هوک استفاده می‌کنیم که می‌گوید: نیرو متناسب با کشش (یعنی نسبت مقدار جابه جایی به طول) است. به ازای فشار هم، نسبت مشابهی وجود دارد.

در اثر نیروی f، بخش AC از میله کش می‌آید. نیروی عکس‌العملی را که پیش می‌آید، با T_1 نشان می‌دهیم، در ضمن بخش CB از میله فشرده می‌شود و نیروی عکس‌العمل T_2 را پدید می‌آورد. با توجه به قانون هوک

$$T_1=\chi\frac{h}{x}\ ,\quad T_2=\chi\frac{h}{l-x}$$

که در آن χ ضریب نسبت است و به ویژگی کشسانی میله بستگی دارد. شرط تعادل نیرو، که بر نقطهٔ C عمل می‌کند، می‌دهد:

$$f=\chi\frac{h}{x}+\chi\frac{h}{l-x}\ \Rightarrow f=\frac{\chi lh}{x(l-x)}$$

از آن جا

$$h=\frac{f}{\chi l}x(l-x)$$

برای این که جابه جایی که در نقطه‌ای مثل y واقع بر میلهٔ AC، یعنی برای $y<x$، به وجود می‌آید، پیدا کنیم، با توجه به قانون هوک نتیجه می‌گیریم، به ازای انبساط میله، انبساط‌پذیری نسبی (یعنی نسبت جابه جایی نقطه به فاصلهٔ آن از نقطهٔ بی‌حرکت)، به موضع نقطه بستگی ندارد. جابه جایی نقطهٔ y را با k نشان می‌دهیم. در این صورت، اگر جابه جایی نسبی را در نقطه‌های x و y برابر قرار دهیم، به دست می‌آید:

$$\frac{k}{y}=\frac{h}{x}$$

از آن جا

$$k = h\frac{y}{x} = \frac{f}{\chi l}y(l - x)\ ,\quad (y<x)$$

به همین ترتیب، اگر نقطه روی پاره‌خط راست CB باشد $(y>x)$، به دست می‌آید:

$$k = h\frac{l-y}{l-x} = \frac{f}{\chi l}x(l - y)$$

اگر به یاد آوریم که تابع اثر، یعنی $k(x\ ,\ y)$ عبارت است از انحراف در نقطهٔ y در نتیجهٔ عمل نیروی واحد در نقطهٔ x قرار دارد، آن وقت برای نوسان‌های طولی میلهٔ کشسان، تابع اثر به این صورت در می‌آید:

$$k(x\ ,\ y) = \begin{cases} \frac{۱}{\chi l}y(l - x) & ,\quad (y < x) \\ \frac{۱}{\chi l}x(l - y) & ,\quad (y > x) \end{cases}$$

می‌توان، کم و بیش شبیه همین حالت، تابع اثر را برای سیم پیدا کرد. اگر کشش سیم را T و طول آن را l بگیریم، آن وقت با اثر گذاشتن نیروی واحدی که در نقطهٔ x قرار دارد، آن وقت سیم به صورتی مثل شکل ۸ در می‌آید و جابه جایی $k(x\ ,y)$ در نقطهٔ y با این دستور داده می‌شود:

$$k(x\ ,y) = \begin{cases} \frac{۱}{Tl}x(l-y), & (x<y) \\ \frac{۱}{Tl}y(l-x), & (x>y) \end{cases}$$

که همان تابع اثر، برای میله است که پیش از این به دست آوردیم.

از تابع اثر می‌توان برای بیان انحراف دستگاه از موقعیت تعادل، برای حالتی که نیرویی با تراکم $f(y)$، به صورتی پراکنده ولی پیوسته روی دستگاه عمل می‌کند، استفاده کرد. چون در بازهٔ به طول Δy، نیروی $f(y)\Delta y$ عمل می‌کند و می‌توانیم به تقریب آن را در نقطهٔ y متمرکز بدانیم، بنابراین زیر تأثیر این نیرو انحراف $k(x\ ,\ y)\,f(y)\Delta y$ در نقطهٔ x به وجود می‌آید و انحراف، زیر تأثیر تمامی بار، به تقریب برابر است با مجموع

$$\sum k(x\ ,\ y)f(y)\Delta y$$

وقتی به ازای $\Delta y \to 0$، به سمت حد برویم، انحراف $u(x)$ در نقطهٔ x، زیر تأثیر نیروی $f(y)$ که در طول دستگاه پراکنده است، با این دستور داده می‌شود:

$$u(x)=\int_a^b k(x,y)f(y)dy \qquad (۲۳)$$

فرض می‌کنیم نیروی بیرونی بر دستگاه کشسان ما تأثیری نداشته باشد. اگر آن را از حالت تعادل خارج کنیم، آن وقت به حرکت درمی‌آید. این حرکت را نوسان آزاد دستگاه گویند.

اکنون به‌کمک تابع اثر $k(x\ ,\ y)$، معادله‌ای را می‌نویسیم که از نوسان‌های آزاد دستگاه کشسان ما پیروی می‌کند. برای این منظور، انحراف از وضع تعادل در نقطهٔ xدر لحظهٔ زمانی t را، $u(x\ ,\ t)$ می‌نامیم. در این صورت، شتاب در نقطهٔ xدر لحظهٔ t برابر $\frac{\partial^2 u(x\ ,\ t)}{\partial t^2}$ می‌شود.

اگر تراکم خطی جسم را ρ بگیریم، یعنی اگر $\rho\, dy$جرم عنصر به طول dy باشد، آن وقت بنابر قانون اصلی مکانیک، معادلهٔ حرکت را، با تبدیل نیروی $f(y)dy$، در (۲۳)، به حاصل‌ضرب جرم در شتاب $\rho dy \frac{\partial^2 u(y\ ,\ t)}{\partial t^2}$ و با علامت منفی، به دست می‌آوریم.

بنابراین، معادلهٔ نوسان‌های آزاد، به این صورت است:

$$u(x\ ,\ t)=-\int_a^b k(x\ ,\ y)\frac{\partial^2 u(y\ ,\ t)}{\partial t^2}\ \rho dy$$

نوسان‌های همساز (هارمونیک) دستگاه کشسان، یعنی حرکت‌هایی که برای آن‌ها داشته باشیم:

$$u(x\ ,\ t)=u(x)\sin\omega t$$

نقش مهمی در نظریهٔ نوسان‌ها به عهده دارند. این نوسان‌ها دارای این ویژگی هستند که هر نقطهٔ آن نوسانی همساز (حرکتی با قانون سینوسی) با بسامدی (فرکانسی) برابر ω دارد و، در ضمن، این بسامد برای همهٔ نقطه‌های xیکی است.

خواهیم دید که هر نوسان آزاد، می‌تواند ترکیبی از نوسان‌های همساز باشد.

اگر در معادلهٔ نوسان‌های آزاد قرار دهیم:

$$u(x\ ,\ t)=u(x)\sin\omega t$$

و به $\sin\omega t$ ساده کنیم، به این معادله برای تعیین تابع $u(x)$ می‌رسیم:

$$u(x)=\rho\omega^{۲}\int_a^b k(x\,,\ y)u(y)dy \qquad (۲۴)$$

این معادله را، معادلهٔ انتگرالی همگن برای تابع $u(x)$ گویند.

روشن است، معادلهٔ (۲۴)، به ازای هر مقدار ω، جواب $u(x)\equiv ۰$ را هم می‌دهد که برای ما جالب نیست و پاسخ‌گوی حالت سکون است. مقدارهایی از ω که به ازای آنها، معادلهٔ (۲۴) جواب مخالف صفر داشته باشد، ویژه بسامدهای دستگاه نامیده می‌شوند.

چون برای همهٔ مقدارهای ω جواب غیرصفر به دست نمی‌آید، دستگاه تنها با بسامدهای معینی، می‌تواند نوسان‌های آزاد داشته باشد. کوچکترین آن‌ها را آهنگ اصلی دستگاه و بقیه را آهنگ فرعی آن می‌گویند.

هر دستگاه دارای دنبالهٔ بی‌پایان از ویژه‌بسامدها است که به آن «طیف بسامدها» گویند:

$$\omega_{۱}\,,\ \omega_{۲}\,,\ \dots,\ \omega_n\,,\ \dots$$

جواب غیرصفر $u_n(x)$ از معادلهٔ (۲۴)، که متناظر با ویژه‌بسامد ω_n است صورتی از نوسان ویژه‌متناظر با آن را به ما می‌دهد.

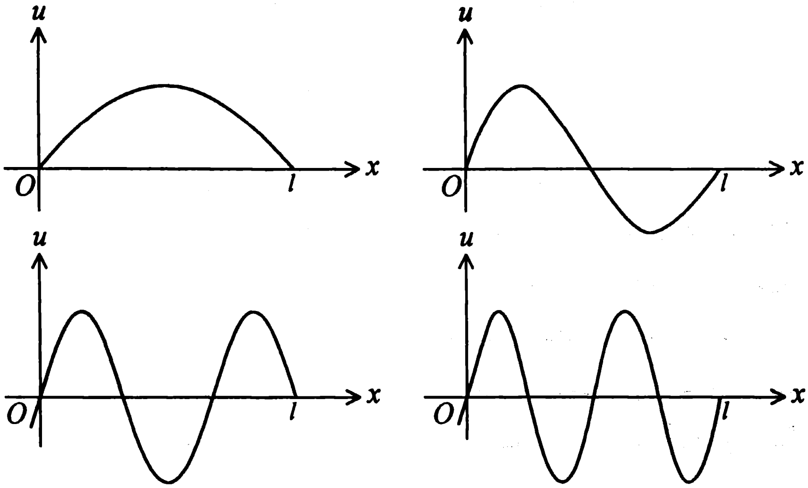
برای نمونه، اگر دستگاه کشسان سیمی باشد که بین نقطه‌های O و l کشیده و در این نقطه‌ها محکم شده باشد، آن‌وقت بسامدهای ممکن نوسان‌های ویژه برای چنین دستگاهی برابرند با

$$a\frac{\pi}{l}\,,\ ۲a\frac{\pi}{l}\,,۳a\frac{\pi}{l}\,,\ \dots,\ na\frac{\pi}{l}\,,\ \dots$$

که در آن، a عبارت است از ضریب مربوط به چگالی و کشش‌پذیری سیم، یعنی $a=\sqrt{\frac{T}{\rho}}$. در این‌جا، آهنگ اصلی $\omega_{۱}=a\frac{\pi}{l}$ است و آهنگ‌های فرعی عبارت‌اند از $\omega_{۲}=۲\omega_{۱}\,,\ \omega_{۳}=۳\omega_{۱}$ ، ... ، $\omega_n=n\omega_{۱}$. شکل‌های نوسان‌های همساز متناظر با این برابری‌ها داده می‌شود:

$$u_n(x)=\sin\frac{na\pi}{l}x$$

و به صورتی درمی‌آیند که برای حالت‌های $n=۱$، $n=۲$، $n=۳$ و $n=۴$ در شکل ۱۰ نشان داده شده است.



شکل ۱۰

بررسی ما مربوط به نوسان‌های آزاد دستگاه کشسان بود. اگر در زمان حرکت، نیروی همساز بیرونی بر دستگاه کشسان اثر بگذارد، آن‌وقت برای تعیین نوسان‌های همسازی که زیر تأثیر این نیرو قرار دارند، برای تابع $u(x)$ به این معادلهٔ انتگرالی ناهمگن می‌رسیم:

$$u(x)=\rho\omega^2\int_a^b k(x\ ,\ y)u(y)dy+h(x) \qquad (۲۵)$$

ویژگی‌های معادله‌های انتگرالی. تا این‌جا، با نمونه‌هایی از معادله‌های انتگرالی آشنا شدیم:

$$f(x)=\lambda\int_a^b k(x\ ,\ y)f(y)dy \qquad (۲۶)$$

و

$$f(x)=\lambda\int_a^b k(x\ ,\ y)f(y)dy+h(x) \qquad (۲۷)$$

که اولی ضمن حل مسألهٔ مربوط به نوسان‌های آزاد دستگاه کشسان، و دومی ضمن بررسی نوسان‌های شرطی، یعنی نوسان‌هایی که زیر تأثیر نیروی بیرونی پدید می‌آیند، به دست آمد.

تابع مجهول در این معادله‌ها، تابع $f(x)$ است، تابع مفروض $k(x\ ,\ y)$ را هستهٔ انتگرالی گویند.

معادلهٔ (۲۷)، ***معادلهٔ انتگرالی خطی ناهمگن*** و معادلهٔ (۲۶) ***معادلهٔ انتگرالی خطی همگن*** است. معادلهٔ همگن از معادلهٔ ناهمگن، به ازای $h(x)=0$ به دست می‌آید.

روشن است که معادلهٔ همگن، همیشه دارای جواب صفر است، یعنی $f(x) = 0$. بین جواب‌های معادله‌های انتگرالی همگن و ناهمگن، رابطهٔ تنگاتنگی وجود دارد. از جمله، این قضیه درست است: اگر معادلهٔ انتگرالی همگن، تنها جواب صفر را داشته باشد، آن وقت معادلهٔ ناهمگن نظیر آن، برای هر تابع $h(x)$، قابل حل است.

اگر برای مقداری از λ، معادلهٔ همگن جوابی برای $f(x)$ غیر از صفر داشته باشد، آن وقت این مقدار λ را ***ویژه‌مقدار*** و جواب $f(x)$ متناظر با آن را ***ویژه‌تابع*** گویند. دیدیم وقتی معادلهٔ انتگرالی، نوسان آزاد یک دستگاه کشسان را شرح می‌دهد، ویژه‌مقدار بستگی کاملی به بسامدهای نوسان‌های دستگاه دارد (یعنی $\lambda = \rho\omega^2$). ویژه‌تابع‌ها هم، نوع نوسان‌های همساز متناظر را می‌دهند.

برای مسأله‌هایی دربارهٔ نوسان‌ها از قانون بقای انرژی داریم:

$$k(x, y) = k(y, x) \tag{۲۸}$$

هسته‌ای که با شرط (۲۸) سازگار باشد، ***متقارن*** نامیده می‌شود.

برای معادله‌ای که هستهٔ متقارن داشته باشد، ویژه‌تابع‌ها و ویژه‌مقدارها، دارای یک رشته ویژگی‌های مهم‌اند. از جمله، برای چنین معادله‌هایی، همیشه دنباله‌ای از ویژه‌مقدارهای حقیقی وجود دارد:

$$\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$$

هر ویژه‌مقدار، با یک یا چند ویژه‌تابع متناظر است. در ضمن، ویژه‌تابع‌ها، برای ویژه‌مقدارهای مختلف همیشه نسبت به هم متعامدند[1].

به این ترتیب، برای هر معادلهٔ انتگرالی با هستهٔ متقارن، دستگاه ویژه‌تابع‌ها، دستگاهی از تابع‌های متعامد است. پرسش پیش می‌آید چه موقع این دستگاه کامل است، یعنی چه موقع می‌توان هر تابع از فضای هیلبرت را به رشته‌ای از دستگاه ویژه‌تابع‌های معادله انتگرالی تجزیه کرد؟ در حالت خاص، اگر معادلهٔ

$$\int_a^b k(x,y) f(y)\,dy = 0 \tag{۲۹}$$

تنها به ازای $f(y) \equiv 0$ برقرار باشد، آن وقت دستگاه ویژه‌تابع‌های معادلهٔ انتگرالی

---

۱. این حکم را در بند بعد ثابت خواهیم کرد.

$$\lambda \int_a^b k(x\ ,\ y)f(y)dy = f(x)$$

یک دستگاه متعامد کامل است[۱].

به این ترتیب، در این حالت، هر تابع انتگرال‌پذیر $f(x)$، می‌تواند به رشته‌ای از ویژه‌تابع‌ها تجزیه شود. با بررسی معادله‌های انتگرالی، روشی کلی و نیرومند برای اثبات بسته بودن دستگاه‌های متعامد مختلف و مهم یعنی تجزیه‌پذیر بودن تابع‌ها به رشته‌ای با تابع‌های متعامد به دست می‌آید. با این روش می‌توان کامل بودن دستگاه‌های شامل تابع‌های مثلثاتی، تابع‌های استوانه‌ای، تابع‌های کروی و بسیاری از دستگاه‌های تابع‌های مهم دیگری را ثابت کرد.

این که هر تابع دل‌خواه را می‌توان به رشته‌ای از ویژه‌تابع‌ها تجزیه کرد، در حالت نوسان‌ها به این معنی است که هر نوسان را می‌توان به مجموع نوسان‌های همساز تجزیه کرد. این‌گونه تجزیه، روشی است که به طور گسترده‌ای برای حل مسأله‌های مربوط به نوسان‌ها در زمینه‌های مختلف مکانیک و فیزیک به‌کار می‌رود (نوسان‌های جسم‌های کشسان، نوسان‌های مربوط به صداشناسی، موج‌های الکترومغناطیسی و غیره).

پیشرفت نظریهٔ معادله‌های انتگرالی خطی، انگیزه‌ای برای ساختن نظریهٔ عملگرهای خطی بود که در واقع، نظریهٔ معادله‌های انتگرالی خطی، به طور کامل در آن نفوذ کرده است. البته، روش‌های کلی نظریهٔ اُپراتورهای خطی هم، به نوبهٔ خود، به پیشرفت بعدی معادله‌های انتگرالی یاری فراوان رساند.

۵. عمل‌گرها یا اُپراتورهای خطی و تکامل بعدی آنالیز تابعی

دربند پیشین دیدیم، مسأله‌های مربوط به نوسان‌های دستگاه‌های کشسان، منجر به مسألهٔ

---

۱. در حالتی که $k(x\ ,\ y)$ تابع اثر برای دستگاه کشسان است، شرط (۲۹) معنای فیزیکی ساده‌ای پیدا می‌کند. درواقع [دستور (۲۳) را ببینید] ، دیدیم زیر تأثیر نیروی $f(y)$ که در طول دستگاه پراکنده است، انحراف دستگاه واقع از وضع تعادل با دستور

$$u(x) = \int_a^b k(x\ ,\ y)\, f(y)dy$$

بیان می‌شود. بنابراین، شرط (۲۹) به این معنی است که، هر نیروی مخالف صفر، دستگاه را از حالت تعادل خارج می‌کند.

مربوط به جست‌وجوی ویژه‌مقدارها و ویژه‌تابع‌های معادله‌های انتگرالی شد. یادآوری این مطلب لازم است که این مسأله‌ها را می‌توان به جست‌وجوی ویژه‌مقدارها و ویژه‌تابع‌های معادله‌های دیفرانسیلی خطی هم منجر کرد (بند ۵ از بخش ششم را در جلد دوم ببینید). بسیاری از مسأله‌های دیگر فیزیک هم، به مسألهٔ محاسبهٔ ویژه‌مقدارها و ویژه‌تابع‌های معادله‌های دیفرانسیلی خطی و معادله‌های انتگرالی خطی منجر می‌شوند.

باز هم نمونه‌ای می‌آوریم. در مهندسی رادیو (رادیوتکنیک)، برای انتقال موج‌های الکترومغناطیسی با بسامدهای بالا، به فراوانی از وسیله‌ای به نام «موج‌بر» استفاده می‌کنند. یعنی لوله‌های فلزی توخالی که در درون آن‌ها، موج‌های الکترومغناطیسی پراکنده‌اند. معلوم است، به وسیلهٔ موج‌بر می‌توان نوسان‌های الکترومغناطیسی با طول موج‌های نه‌چندان بزرگ را منتشر کرد. جست‌وجوی طول موج‌های بحرانی، منجر به مسألهٔ مربوط به‌ویژه‌مقدارهای یک معادلهٔ دیفرانسیلی می‌شود.

به‌جز این، در جبر خطی، در نظریهٔ معادله‌های دیفرانسیلی عادی، در مسأله‌های مربوط به پایداری و مقاومت و غیره هم، با ویژه‌مقدارها برخورد می‌کنیم.

لازم است همهٔ این موضوع‌های مشابه را، از دیدگاه کلی و واحدی بررسی کنیم. چنین دیدگاهی را، نظریهٔ کلی عمل‌گرها (اُپراتورها) تامین می‌کند. به بسیاری از پرسش‌های مربوط به‌ویژه‌تابع‌ها و ویژه‌مقدارها که در مسأله‌های مشخص مختلف پیش می‌آید، تنها می‌توان در پرتو نظریهٔ کلی عمل‌گرها، به طور کامل پاسخ داد. بنابراین، نظریهٔ کلی عمل‌گرها، علاوه بر جهت‌های دیگر در این جهت هم می‌تواند ابزاری سودمند و ثمربخش باشد.

در تکامل بعدی نظریهٔ عمل‌گرها، مکانیک کوانتایی نقشی سرنوشت‌ساز داشت، چرا که در آن، به صورت گسترده‌ای از روش‌های نظریهٔ عمل‌گرها استفاده می‌شود. ابزار اصلی ریاضی در مکانیک کوانتایی عبارت است از نظریهٔ عمل‌گرهای «خودمزدوج» یا «خودالحاقی[1]». طرح مسأله‌هایی از ریاضیات، که در مکانیک کوانتایی پدید می‌آید، چه در گذشته و چه بعد از این، مهم‌ترین انگیزه برای پیشرفت آنالیز تابعی است.

دیدگاه نظریهٔ عمل‌گرها نسبت به معادله‌های دیفرانسیلی و معادله‌های انتگرالی برای تکامل روش‌های عملی و تقریبی حل این معادله‌ها هم، بسیار سودمند است.

---

1. Self - Conjugate یا Self - adjoint

مفهوم‌های اصلی نظریهٔ عمل‌گرها. به طرح تعریف‌ها و مفهوم‌های اصلی مربوط به نظریهٔ عمل‌گرها می‌پردازیم.

در آنالیز با مفهوم تابع برخورد می‌کنیم. این مفهوم، به ساده‌ترین بیان خود، عبارت است از یک تناظر، که هر عدد x را (به نام متغیر مستقل) به عدد متناظر خود y (به نام تابع) مربوط می‌کند. با پیشرفت آنالیز، ضرورت بررسی تناظرهایی از نوع کلی‌تر به وجود آمد.

از این گونه تناظرهای کلی، از جمله در حساب وردشی (بخش هشتم، جلد دوم) بررسی می‌شود که مقابل هر تابع یک عدد قرار می‌گیرد. اگر هر تابع در تناظر با یک عدد قرار گیرد، می‌گوییم یک تابعک (فونکسیونل) به ما داده شده است. برای نمونهٔ تابعک، می‌توان از تناظر هر تابع $y=f(x)$ $(a \leq x \leq b)$ با طول کمان نمودار آن نام برد. می‌توانیم نمونهٔ دیگری از تابعک را، تناظر هر تابع $y=f(x)$ $(a \leq x \leq b)$ با انتگرال معین $\int_a^b f(x)dx$ بدانیم.

اگر $f(x)$ را نقطه‌ای از فضای بی‌نهایت بعدی بگیریم، آن وقت تابعک چیزی جز تابعی از نقطه‌های فضای بی‌نهایت بعدی نیست. از این دیدگاه، در حساب وردشها، مسأله‌های مربوط به جست‌وجوی ماکزیمم‌ها و می‌نیمم‌های تابع‌های نقطه‌های فضای بی‌نهایت‌بعدی بررسی می‌شود.

برای این که تعریفی برای تابعک پیوسته پیدا کنیم، باید پیش از آن، نزدیکی دو نقطه را در فضای بی‌نهایت‌بعدی تعریف کنیم. در بند ۲، فاصلهٔ بین دو تابع [یعنی بین نقطه‌های $f(x)$ و $g(x)$ از فضای بی‌نهایت‌بعدی] را به عنوان

$$\sqrt{\int_a^b [f(x)-g(x)]^2 dx}$$

تعریف کردیم. این روش تعیین فاصله در فضای بی‌نهایت بعدی، اغلب به‌کار می‌رود، ولی البته تنها امکان برای معرفی فاصله نیست. در جاهای دیگر، ممکن است روش‌های دیگری برای معرفی فاصلهٔ بین تابع‌ها، مناسب‌تر باشد. از جمله، در مسأله‌های مربوط به نظریهٔ تقریب تابع‌ها (بند ۳ از بخش هفتم را در جلد دوم ببینید)، فاصلهٔ بین تابع‌ها را، به یاری مفهوم نزدیکی دو تابع $f(x)$ و $g(x)$ مشخص کردیم، یعنی با دستور

$$\max |f(x) - g(x)|$$

روش‌های دیگری هم برای معرفی فاصلهٔ بین تابع‌ها وجود دارد که ضمن بررسی

تابعک‌ها در حساب وردشها به‌کار می‌رود.

شیوه‌های گوناگون معرفی فاصلهٔ بین تابع‌ها، ما را به فضاهای بی‌نهایت بعدی گوناگون می‌رساند.

به این ترتیب، فضاهای بی‌نهایت‌بعدی (فضاهای تابعی)، با ذخیره‌ای که از تابع‌ها دارند، و با تعریفی که از فاصلهٔ بین این تابع‌ها می‌شود، با هم اختلاف پیدا می‌کنند. برای نمونه، اگر مجموعهٔ همهٔ تابع‌های انتگرال‌پذیر را در نظر بگیریم و فاصلهٔ بین تابع‌ها را با دستور

$$\sqrt{\int_a^b [f(x)-g(x)]^{۲} dx}$$

تعریف کنیم، آن وقت به فضای هیلبرت می‌رسیم که در بند ۲ از آن یاد کردیم. ولی اگر مجموعهٔ همهٔ تابع‌های پیوسته را در نظر بگیریم و فاصلهٔ بین تابع‌ها را با دستور

$$\max |f(x) - g(x)|$$

تعریف کنیم، آن وقت به فضایی می‌رسیم که آن را فضای (C) می‌نامند.

ضمن بررسی معادله‌های انتگرالی با عبارت‌هایی به صورت

$$g(x)= \int_a^b k(x , y)f(y)dy$$

برخورد می‌کنیم. با معلوم بودن هستهٔ $k(x , y)$، این برابری به صورت قانونی در می‌آید که به یاری آن، هر تابع $f(x)$در تناظر با تابع دیگر $g(x)$قرار می‌گیرد. این گونه تناظر، که تابع fرا به تابع دیگر gمربوط می‌کند، عمل‌گر (یا اُپراتور) گویند.

وقتی از عمل‌گر خطی A در فضای هیلبرت صحبت می‌کنیم، به معنای وجود قانونی است که به یاری آن، هر تابع fمتناظر با تابع g قرار گیرد. این تناظر ممکن است برای همهٔ تابع‌های فضای هیلبرت ممکن نباشد. در این صورت، مجموعهٔ تابع‌های f را، که برای آن‌ها تابع $g=Af$ وجود دارد، **حوزهٔ تعریف** یا **دامنهٔ** عمل‌گر A نامیده می‌شود (به همین ترتیب، **حوزه مقدارها** یا برد تابع‌ها هم تعریف می‌شود). خود این تناظر را، این طور نشان می‌دهند:

$$g=Af \tag{۳۰}$$

خطی بودن عمل‌گر به این معنی است که مجموع تابع‌های f_1و f_2 در تناظر با مجموع Af_1

و $Af_۲$ است؛ در ضمن، حاصل‌ضرب تابع f در عدد λ متناظر با تابع λAf قرار می‌گیرد، یعنی

$$A(f_۱+f_۲)=Af_۱+Af_۲ \tag{۳۱}$$

و

$$A(\lambda f)=\lambda Af \tag{۳۲}$$

اغلب، شرط پیوستگی را هم برای عمل‌گرهای خطی در نظر می‌گیرند، یعنی اگر دنبالهٔ تابع‌های f_n به سمت f هم‌گرا باشند، دنبالهٔ Af_n هم به سمت تابع Af هم‌گرا می‌شود.

نمونه‌هایی از عمل‌گرهای خطی را می‌آوریم.

۱°. هر تابع $f(x)$ را متناظر با تابع $g(x)=\int_a^x f(t)dt$، یعنی انتگرال نامعین تابع f قرار می‌دهیم. خطی بودن این عمل‌گر، با توجه به ویژگی عادی انتگرال روشن می‌شود: انتگرال مجموع برابر است با مجموع انتگرال‌ها؛ اگر تابع دارای ضریب ثابتی باشد، می‌توان این ضریب را از زیر علامت انتگرال بیرون آورد.

۲°. هر تابع دیفرانسیل‌پذیر $f(x)$ را متناظر با مشتق آن $f'(x)$ قرار می‌دهیم. این عمل‌گر با حرف D نشان می‌دهند، یعنی

$$f'(x)=Df(x)$$

توجه کنیم، این عمل‌گر برای همهٔ تابع‌های فضای هیلبرت معین نیست و تنها تابع‌هایی را در بر می‌گیرد که دارای مشتق باشند و در ضمن، به فضای هیلبرت تعلق داشته باشند. در این صورت، چنین تابع‌هایی، دامنهٔ عمل‌گر مفروض را تشکیل می‌دهند.

۳°. مثال‌های ۱ و ۲ نمونه‌هایی از عمل‌گرهای خطی در فضای بی‌نهایت‌بعدی بودند. عمل‌گرهای خطی در فضاهای با تعداد بعدهای محدود را در بخش‌های دیگری از این کتاب دیده‌ایم. از جمله، در بخش سوم (جلد اول)، تبدیل‌های آفین را بررسی کردیم. اگر تبدیل آفین صفحه یا فضا، مبداء مختصات را در جای خود نگه دارد، آن وقت مثالی برای عمل‌گر فضای دوبعدی، و شبیه آن برای فضای سه‌بعدی است. در بخش شانزدهم هم، که از تبدیل خطی فضای n بعدی صحبت کردیم، در واقع، عمل‌گر خطی در فضای n بعدی بود.

۴°. در معادله‌های انتگرالی، با خانواده‌ای از عمل‌گرها در فضای تابعی سر و کار داریم که خانواده‌ای مهم و اساسی از عمل‌گرها است و کاربردهای فراوانی در آنالیز دارد که آن‌ها را عمل‌گرهای انتگرالی گویند. اگر به تابع معین $k(x, y)$ برگردیم، آن وقت تابع

$$g(x) = \int_a^b k(x, y) f(y) dy$$

هر تابع f را در تناظر با تابعی مثل g قرار می‌دهد. این تبدیل را، به صورت نمادین، می‌توان این طور نشان داد:

$$g = Af$$

عمل‌گر A، در این حالت، عمل‌گر انتگرالی نامیده می‌شود. عمل‌گرهای انتگرالی را، به فراوانی، می‌توان پیدا کرد.

در بند ۴، دربارهٔ این معادلهٔ انتگرالی ناهمگن صحبت کردیم:

$$f(x) = \lambda \int_a^b k\,(x\ ,\ y) f(y) dy + h(x)$$

با توجه به نمادی که برای عمل‌گرها به‌کار می‌بریم، این معادله را می‌توان این طور نوشت:

$$f = \lambda Af + h \qquad (۳۳)$$

که در آن λ عددی مفروض، h تابعی مفروض (برداری از فضای بی‌نهایت‌بعدی) و f تابع مجهول است. معادلهٔ همگن هم، با همین نمادها به این صورت نوشته می‌شود:

$$f = \lambda Af \qquad (۳۴)$$

قضیه‌های سنتی مربوط به معادله‌های انتگرالی و از جمله قضیه‌ای که در بند ۴ دربارهٔ رابطهٔ جواب‌های معادله‌های انتگرالی ناهمگن با معادله‌های انتگرالی همگن نظیر آن‌ها تنظیم کردیم، برای هر معادلهٔ عمل‌گر صادق نیست. با وجود این، می‌توان شرط‌هایی کلی برای عمل‌گر A در نظر گرفت که این قضیه به ازای آن‌ها درست باشد.

این شرط‌ها، با اصطلاح‌های توپولوژیک تنظیم شده‌اند و حاکی از آن هستند که عمل‌گر A کرهٔ واحد را (یعنی مجموعه‌ای از بردارهایی را که طول آن‌ها از واحد تجاوز نمی‌کند) به مجموعه‌ای فشرده تبدیل کند.

ویژه‌مقدارها و ویژه‌بردارهای عمل‌گرها. مسألهٔ مربوط به‌ویژه‌مقدارها و ویژه‌تابع‌ها در معادلهٔ انتگرالی، که از راه مسألهٔ مربوط به نوسان‌ها به آن رسیدیم، به این صورت تنظیم می‌شود: مقدارهای λ را طوری پیدا کنید که به ازای آن‌ها، تابع‌های مخالف صفر f را داشته باشیم که

در معادلهٔ

$$f(x)= \lambda\int_a^b k(x\ ,\ y)f(y)dy$$

صدق کنند. مثل قبل، این معادله را می‌توان به این صورت نوشت:

$$f=\lambda Af$$

یا

$$Af=\frac{1}{\lambda}f \tag{۳۵}$$

در این جا، A را به معنای عمل‌گری خطی و دل‌خواه گرفته‌ایم. در این صورت، بردار f که در برابری (۳۵) صدق می‌کند، ویژه‌بردار عمل‌گر A و عدد $\frac{1}{\lambda}$ ویژه‌مقدار متناظر آن نامیده می‌شود.

چون بردار $\frac{1}{\lambda}f$ با بردار f هم‌جهت است (و تنها از نظر ضریب عددی با آن فرق دارد)، بنابراین مسألهٔ جست‌وجوی ویژه‌بردارها می‌توان به عنوان مسألهٔ جست و جوی بردارهای غیرصفری تنظیم کرد که جهت آن‌ها، ضمن تبدیل A، تعییر نمی‌کنند.

چنین دیدگاهی به مسألهٔ مربوط به‌ویژه‌مقدار، این امکان را پدید می‌آورد که مسألهٔ مربوط به‌ویژه‌مقدارهای معادله‌های انتگرالی (اگر A عمل‌گر انتگرالی باشد)، معادله‌های دیفرانسیلی (اگر A عمل‌گر دیفرانسیلی باشد) را با مسألهٔ مربوط به‌ویژه‌مقدارها در جبر خطی (اگر A تبدیلی خطی در فضای با تعداد محدود بعد باشد، بخش ششم و بخش هفتم را ببینید)، پیوند دهیم. برای فضای سه‌بعدی، به این مسأله، ضمن جست‌وجوی محورهای اصلی بیضوی (الیپسوید) برخورد می‌کنیم.

در حالت معادله‌های انتگرالی، یک رشته از ویژگی‌های مهم را دربارهٔ ویژه‌تابع‌ها و ویژه‌مقدارها (مثل حقیقی بودن ویژه‌مقدارها، متعامد بودن ویژه‌تابع‌ها و غیره)، می‌توان به عنوان نتیجه‌ای از متقارن بودن هسته، یعنی برابری $k(x\ ,\ y)=k(y\ ,\ x)$ به دست آورد.

برای عمل‌گر خطی دل‌خواه A در فضای هیلبرت، شبیه این ویژگی، برای به اصطلاح عمل‌گر «خودمزدوجی» (یا «خودالحاقی») وجود دارد.

شرط خودالحاقی عمل‌گر A در حالت کلی، به این معنا است که برای هر دو عضو f_1 و f_2، این برابری برقرار باشد:

$$(Af_1\ ,\ f_2)=(f_1\ ,\ Af_2)$$

که در آن $(Af_1\ ,\ f_2)$ به معنای حاصل‌ضرب اسکالر بردار Af_1 در بردار f_2 است.

شرط خودالحاقی عمل‌گر در مسأله‌های مربوط به مکانیک، اغلب نتیجه‌ای از قانون بقای انرژی است. بنابراین، شرط در عمل‌گرهایی که، درمثل، مربوط به نوسان‌هایی است، که به ازای آن‌ها، انرژی تلف شده نداشته باشیم.

بسیاری از عمل‌گرهایی هم که در مکانیک کوانتایی با آن‌ها سروکار داریم، خودالحاقی‌اند.

تحقیق می‌کنیم، عمل‌گر انتگرالی یا هستهٔ متقارن $k(x\ ,\ y)$، خودالحاقی است. در واقع، در این حالت، Af_1 عبارت است از تابع

$$\int_a^b k(x\ ,\ y)f_1(y)dy$$

بنابراین، حاصل ضرب اسکالر $(Af_1\ ,\ f_2)$ عبارت است از انتگرال حاصل‌ضرب این تابع در f_2، که با این دستور داده می‌شود:

$$(Af_1\ ,\ f_2)=\int_a^b\int_a^b k(x\ ,\ y)f_1(y)f_2(x)\,dydx$$

به همین ترتیب

$$(f_1, Af_2)=\int_a^b\int_a^b k(x, y)f_2(y)f_1(x)dydx$$

و برابری $(Af_1\ ,\ f_2)=(f_1\ ,\ Af_2)$، نتیجهٔ مستقیم متقارن بودن هستهٔ $k(x\ ,\ y)$ است.

عمل‌گرهای خودالحاقی دل‌خواه، ویژگی‌های مهمی دارند که در کاربرد این عمل‌گرها برای حل مسأله‌های مختلف، سودمندند. به‌ویژه معلوم می‌شود که ویژه‌مقدارهای عمل‌گرهای خطی خودالحاقی، همیشه حقیقی‌اند و ویژه‌تابع‌های متناظر با ویژه‌مقدارهای مختلف نسبت به هم متعامدند.

برای نمونه، این گزاره را ثابت می‌کنیم. λ_1 و λ_2 را دو ویژه‌مقدار عمل‌گر A، و f_1 و f_2 را ویژه‌بردارهای متناظر با آن‌ها فرض می‌کنیم. این، به معنای آن است که

$$\begin{aligned} Af_1&=\lambda_1 f_1\\ Af_2&=\lambda_2 f_2 \end{aligned} \tag{۳۶}$$

نخستین برابری (۳۶) را در f_2 و دومی را در f_1 (به صورت اسکالر) ضرب می‌کنیم.

به دست می‌آید:

$$(Af_1 \,,\, f_2) = \lambda_1 (f_1 \,,\, f_2)$$
$$(Af_2 \,,\, f_1) = \lambda_2 (f_2 \,,\, f_1) \qquad (۳۷)$$

چون عمل‌گر A خودالحاقی است، بنابراین

$$(Af_1 \,,\, f_2) = (Af_2 \,,\, f_1)$$

برابری دوم (۳۷) را از برابری اول آن کم می‌کنیم، به دست می‌آید:

$$0 = (\lambda_1 - \lambda_2)(f_1 \,,\, f_2)$$

چون $\lambda_1 \neq \lambda_2$، پس $(f_1 \,,\, f_2) = 0$، یعنی ویژه‌بردارهای f_1 و f_2 متعامدند.

بررسی خودالحاقی عمل‌گرها در بسیاری از پرسش‌ها و مسأله‌های مشخص به روشنی انجام می‌گیرد و به نظریهٔ ویژه‌مقدارها مربوط می‌شود.

دربارهٔ یکی از این مسأله‌ها، یعنی مسألهٔ مربوط به تجزیهٔ به وسیلهٔ ویژه‌تابع‌ها، در حالت طیف پیوسته، بیشتر صحبت می‌کنیم.

برای این که به معنای طیف پیوسته پی‌ببریم، دوباره به مسألهٔ نوسان سیم برمی‌گردیم. پیش از این گفته‌ایم، برای سیم به طول l، ویژه‌بسامدهای نوسان ممکن است به صورت دنباله‌ای از این مقدارها باشد:

$$a\frac{\pi}{l} \,,\; 2a\frac{\pi}{l} \,,\; \dots \,,\; na\frac{\pi}{l} \,,\; \dots$$

نقطه‌های این دنباله را روی محور عددی $O\lambda$ مشخص می‌کنیم. اگر طول l سیم را زیاد کنیم، فاصلهٔ بین هر دو نقطهٔ مجاور دنباله کوچک می‌شود. به نحوی که به صورت متراکم‌تری محور عادی را می‌پوشانند. در حالت حدی، وقتی $l \to \infty$، یعنی برای سیمی به طول بی‌نهایت، ویژه‌بسامدها، نیم محور $\lambda \geq 0$ را پر می‌کنند. در این حالت است که می‌گویند، دستگاه دارای طیف پیوسته است.

گفته‌ایم، تجزیهٔ برحسب ویژه‌تابع‌ها برای سیم به طول l، عبارت است از تجزیه به صورت رشته‌ای برحسب سینوس و کسینوسِ $n\frac{\pi}{l}x$، یعنی به رشتهٔ مثلثاتی

$$f(x) = \frac{a_0}{2} + \sum a_n \cos n\frac{\pi}{l}x + b_n \sin n\frac{\pi}{l}x$$

برای حالت سیم بی‌پایان، دوباره می‌توان ثابت کرد که یک تابع دل‌خواه را، کم و بیش، می‌توان برحسب سینوس و کسینوس تجزیه کرد. ولی از آن‌جا که در این‌جا ویژه‌بسامدها در طول خط راست عددی، به صورتی پیوسته پراکنده‌اند، این تجزیه (یا بسط) به صورت یک رشته در نمی‌آید، بلکه به صورت انتگرال فوریه است:

$$f(x) = \int_{-\infty}^{+\infty} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x] \, d\lambda$$

بسط یا تجزیهٔ به انتگرال فوریه، از سدهٔ نوزدهم شناخته شده بود و برای حل مسأله‌های مختلف فیزیک ریاضی از آن استفاده می‌شد.

با وجود این، مسأله‌های کلی‌تر با طیف پیوسته[1]، پرسش‌های زیادی دربارهٔ بسط تابع‌ها به‌ویژه‌تابع‌ها، بی‌پاسخ مانده بود. تنها با پیدایی نظریهٔ کلی عمل‌گرهای خودالحاقی، پاسخ‌های لازم برای این پرسش‌ها پیدا شد.

به‌گروهی از مسأله‌های دیگر هم، که راه حل خود را براساس نظریهٔ کلی عمل‌گرها پیدا کردند. اشاره کنیم. برای مثال، بررسی نوسان‌ها، ضمن پراکندگی انرژی را باید به این گونه مسأله‌ها مربوط دانست.

در این حالت، دوباره می‌توانیم، نوسان‌های آزاد دستگاه را به صورت $u(x)\varphi(t)$ جست‌وجو کنیم. ولی، برخلاف حالت نوسان‌های بدون پراکندگی انرژی، در این جا تابع $\varphi(t)$ به صورت سادهٔ $\cos\omega t$ نیست، بلکه به صورت $e^{-kt}\cos\omega t$ است که در آن $k>0$. بنابراین، جواب متناظر با آن به صورت $u(x)\,e^{-kt}\cos\omega t$ است. در این حالت، هر نقطهٔ x باز هم نوسان دارد (با بسامد ω)، ولی این نوسان‌ها رو به خاموشی می‌رود، زیرا وقتی t به سمت بی‌نهایت میل کند، دامنهٔ این نوسان‌ها، که دارای ضریب e^{-kt} است، به سمت صفر میل می‌کند.

بهتر است ویژه‌نوسان‌های دستگاه به صورت مختلط نوشته شود: $u(x)e^{-i\lambda t}$ که در آن، عدد λ در حالت نبودن مالش (یا اصطکاک) حقیقی و در حالت وجود مالش مختلط است.

مسألهٔ مربوط به نوسان‌های دستگاه، در حالت پراکنده شدن انرژی، دوباره منجر به‌ویژه‌مقدارها می‌شود، البته نه برای عمل‌گرهای خودالحاقی. برای این عمل‌گرها، ویژه‌مقدارهای مختلط وجود دارد که حاکی از خاموشی نوسان‌های آزاد است.

---

۱. برای نمونه، می‌توان از نوسان‌های محیط کشسان ناهمگن و بسیاری از مسأله‌های مکانیک کوانتایی نام برد.

در سالهای ۱۹۵۰ ـ ۱۹۵۱، م.و.کاریش، با استفاده از روش‌های نظریهٔ عمل‌گرها و هم روش‌های نظریهٔ تابع‌های تحلیلی، به بررسی این گروه از مسأله‌ها پرداخت و کامل بودن دستگاه ویژه‌تابع‌ها را، برای آن‌ها، ثابت کرد.

بستگی آنالیز تابعی با شاخه‌های دیگر ریاضیات و با مکانیک کوانتایی. پیش از این هم اشاره کرده‌ایم که پیدایی مکانیک کوانتایی، انگیزه‌ای جدی برای تکامل آنالیز تابعی بود.

همان طور که نیازهای مکانیک و فیزیک سنتی در سدهٔ هجدهم، محاسبه‌های دیفرانسیلی و انتگرالی را پدید آورد، آنالیز تابعی هم، زیر تأثیر نیرومند فیزیک معاصر و به‌ویژه مکانیک کوانتایی، پدید آمد و تکامل یافت. ابزار اصلی ریاضی برای بررسی‌های مکانیک کوانتایی، شاخه‌هایی از ریاضیات است که در واقع، به آنالیز تابعی مربوط می‌شوند. در این‌جا، خیلی کوتاه دربارهٔ این بستگی صحبت می‌کنیم، زیرا طرح پایه‌های مکانیک کوانتایی، خارج از چارچوب تعیین شده برای این کتاب است.

در مکانیک کوانتایی، حالت دستگاه، به وسیلهٔ یک بردار فضای هیلبرت داده می‌شود. پدیده‌هایی مثل انرژی، تکان، گشتاور مقدار حرکت، به یاری عمل‌گرهای خودالحاقی بررسی می‌شوند. برای نمونه، سطح ممکن انرژی الکترون در اتم، به عنوان ویژه‌مقدارهای عمل‌گر انرژی محاسبه می‌شود. اختلاف این ویژه‌مقدارها بسامدهای اتم نور منتشر شده را می‌دهد و بنابراین ساختار طیف تابش را در جسم مفروض معین می‌کند. در ضمن حالت‌های متناظر الکترون به عنوان ویژه‌تابع‌های عمل‌گر انرژی شرح داده می‌شود.

برای حل مسأله‌های مکانیک کوانتایی، اغلب لازم است ویژه‌مقدارهای عمل‌گرهای مختلف (وبیشتر دیفرانسیلی) را محاسبه کنیم. وقتی با حالت‌های پیچیده‌تری سروکار داشته باشیم، حل دقیق این مسأله‌ها در عمل ممکن نیست. برای حل تقریبی این مسأله‌ها، به طور گسترده‌ای از به اصطلاح «نظریهٔ اختلال» استفاده می‌کنند که امکان می‌دهد از روی ویژه‌مقدارها و ویژه‌تابع‌های عمل‌گر خودالحاقی مثل A، ویژه‌مقدارهای عمل‌گر A_1 را، که خیلی کم با آن اختلاف دارد، پیدا کنند.

جدا از تعیین تقریبی ویژه‌مقدارها، اغلب می‌توان دربارهٔ مسألهٔ مفروض به یاری بررسی کیفی اظهارنظر کرد. این بررسی در مسأله‌های مکانیک کوانتایی بر پایهٔ تقارن‌هایی که در مسألهٔ مفروض وجود دارد، انجام می‌گیرد. نمونهٔ چنین تقارن‌هایی را می‌توان در ویژگی‌های تقارن بلورها، در تقارن کروی اتم، در تقارن نسبت به بازتاب و غیر آن پیدا کرد. از آن جا که

تقارن‌ها یک گروه را تشکیل می‌دهند (بخش بیستم را ببینید)، روش‌های مربوط به نظریهٔ گروه‌ها، می‌تواند بدون محاسبه پاسخ‌گوی برخی از پرسش‌ها باشد، در رده‌بندی طیف‌های اتمی، در تبدیل‌های هسته‌ای و در برخی مسأله‌های دیگر.

به این ترتیب، مکانیک کوانتایی به صورت گسترده‌ای از ابزار ریاضی نظریهٔ عمل‌گرهای خودالحاقی استفاده می‌کند. در ضمن، پیشرفت مکانیک کوانتایی هم، به نوبهٔ خود، موجب تکامل نظریهٔ عمل‌گرها شده و مسأله‌های تازه‌ای را در برابر آن قرار داده و قرار می‌دهد.

تأثیر مکانیک کوانتایی و هم تکامل درونی ریاضیات در آنالیز تابعی، به این‌جا منجر شد که مسأله‌ها و روش‌های جبری، نقشی جدی در آنالیز تابعی پیدا کردند. این کشش به سمت روش‌های جبری که به نوبهٔ خود موجب شکوفایی این روش‌ها شد، فیزیک نظری سدهٔ بیستم را متمایز می‌کند.

در پایان دوباره تاکید می‌کنیم، آنالیز تابعی یکی از شاخه‌های ریاضیات امروزی است که توانسته است تأثیر عظیمی بر شکوفایی و پیشرفت همهٔ جنبه‌های ریاضیات داشته باشد. بستگی آنالیز تابعی با فیزیک و کاربرد آن در معادله‌های دیفرانسیلی، محاسبه‌های تقریبی و استفاده از روش‌های کلی آن در جبر، توپولوژی، نظریهٔ تابع‌های با متغیر حقیقی و بسیاری از شاخه‌های دیگر ریاضیات، نظریهٔ آنالیز تابعی را به یکی از کارآمدترین شاخه‌های جدید ریاضیات تبدیل کرده است.

بخش بیستم

گروه‌ها و دستگاه‌های دیگر جبری

آ.ای.مالتسرف

۱. ورود به‌مطلب

در بخش چهارم (جلد اول)، که به جبر چندجمله‌ای‌ها اختصاص داشت، دربارهٔ مسیرهای اصلی تکامل جبر، جایگاه آن در میان سایر شاخه‌های ریاضیات و دربارهٔ تغییر دیدگاهی که نسبت به خود موضوع جبر پدید آمده‌است، صحبت کردیم. در این بخش می‌خواهیم تصوری از نظریه‌های تازهٔ جبری که البته در سدهٔ نوزدهم پدید آمده‌اند، ولی در سدهٔ بیستم به‌طور جدی تکامل یافته‌اند و تأثیر بسیار زیادی بر بررسی‌های ریاضی داشته‌اند، به‌خواننده بدهیم.

جبر امروزی، مثل جبر کلاسیک، بازهم به مسألهٔ آموزش دربارهٔ عمل‌ها و دربارهٔ قاعده‌های محاسبه می‌پردازد. ولی مثل سابق، خود را به بررسی ویژگی‌های عمل روی عددها محدود نکرده است و می‌خواهد ویژگی‌های عمل را دربارهٔ عنصرهایی با طبیعتی بسیار کلی‌تر در بررسی‌های خود وارد کند. این کشش تازه به سمت بررسی‌های تازه، پاسخ‌گوی نیازهای عمل است. از جمله، در مکانیک نیاز داریم نیروها، سرعت‌ها و دوران‌ها را باهم جمع کنیم. در جبر خطی (بخش شانزدهم را ببینید)، اندیشه‌ها و روش‌هایی برای محاسبه وجود دارد و به‌طور گسترده‌ای درعمل به‌کار می‌روند، در حالی که حوزهٔ عمل آن، ماتریس‌ها، تبدیل‌های خطی و بردارهای فضای n بُعدی است.

نقش اساسی را در جبر جدید، نظریهٔ گروه‌ها به‌عهده دارد که در این بخش، جای بزرگی به‌آن داده شده است. از نظریه‌های جبری دیگری که به‌آن پرداخته‌ایم، نظریهٔ دستگاه‌های فرامختلط است که از نظر تاریخی، مرحلهٔ لازم و مهمی را در تکامل مفهوم عدد تشکیل می‌دهد. البته، این دو نظریه به تنهایی نمی‌توانند چهره و مضمون جبر معاصر را به‌طور کامل نشان دهند، ولی اندیشه‌ها و روش آن را، تاحد زیادی روشن می‌کنند.

نظریهٔ گروه‌ها به‌خاطر نیاز به یافتن ابزاری برای بررسی قانون‌مندی‌های مهمی از دنیای

واقع، همچون قانون‌مندی‌های تقارن‌ها، به‌وجود آمد.

شناخت ویژگی‌های تقارنی یک جسم هندسی یا موضوع دیگری از هندسه و فیزیک، اغلب کلیدی است که درها را به‌روی شناسایی ساختار این جسم‌ها و موضوع‌ها به‌روی ما می‌گشاید. با این که تقارن مفهومی عینی است، پاسخ دقیق و کلی به‌این پرسش که، تقارن یعنی چه و چگونه می‌توان ویژگی‌های تقارن را به صورت کمیتی وارد محاسبه کرد، نیاز به استفاده از نظریهٔ گروه‌ها دارد.

طرح نظریهٔ گروه‌ها تازگی ندارد و به پایان سدهٔ هجدهم و آغاز سدهٔ نوزدهم برمی‌گردد. این نظریه، در آغاز، تنها به‌عنوان وسیله‌ای کمکی برای حل مسأله‌های مربوط بـه مـعادله‌های از درجه‌های بالا به یاری رادیکال‌ها مطرح شد. ضمن بررسی این مسأله، معلوم شد که ویژگی‌های تقارنی ریشه‌های معادله، برای حل تمامی مسأله، موضوعی جدی و اساسی است. در طـول سده‌های ۱۹ و ۲۰، نقش مهمی که قانون‌مندی‌های تقارن دارد، در بسیاری از رشته‌های دانش روشن شد: هندسه، بلورشناسی، فیزیک، شیمی. به‌برکت این روش‌ها و نتیجه‌گیری‌ها، نظریهٔ گروه‌ها توانست گستردگی زیادی به‌دست آورد. از آن‌جا که هر رشته‌ای از دانش، مسأله‌هـای خاص خود را در برابر نظریهٔ گروه‌ها می‌گذاشت، فراوانی و گوناگونی این جنبه‌ها، بـه‌نوبهٔ خود، موجب پیشرفت نظریهٔ گروه‌ها شد و به‌این جا رسید که امروز، به‌یکی از شاخه‌های خاص و مستقل ریاضیات تبدیل شده است: نظریهٔ کلی گروه‌ها، نظریهٔ گروه‌های محدود، نظریهٔ گروه‌های پیوسته، گروه‌های گسستهٔ تبدیل، نـظریهٔ نـمایش گـروه‌ها و مشـخصه‌های آن‌ها. به‌تدریج روش‌ها و مفهوم‌های مربوط به نظریهٔ گروه‌ها چنان تکامل یافت که، بـه‌جز بررسی قانون‌مندی‌های تقارن، برای حل بسیاری از مسأله‌های دیگر هم، آمادگی پیدا کرد.

امروز، مفهوم گروه به‌یکی از مهم‌ترین و کلی‌ترین مفهوم‌های ریاضیات تبدیل شده است و نظریهٔ گروه‌ها، در میان شاخه‌های ریاضیات، جای نمایانی را به‌خود اختصاص داده است. *اِ.س.فِردوروف، اُ.یو.شمیت، ل.س.پون تریاگین* را می‌توان ازجملهٔ کسانی دانست که در تکامل نظریهٔ گروه‌ها و کاربردهای آن، نقش اساسی داشته‌اند.

۲. تقارن و تبدیل

ساده‌ترین صورت‌های تقارن. از ساده‌ترین صورت‌های تقارن که در زندگی روزانه دیده می‌شود،

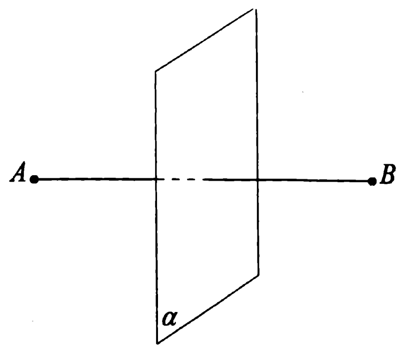
آغاز می‌کنیم. یکی از این صورت‌های تقارن، تقارن آئینه‌ای جسم‌های هندسی یا تقارن نسبت به‌صفحه است.

نقطهٔ A از فضا را قرینهٔ نقطهٔ B نسبت به‌صفحهٔ α گویند (شکل ۱)، وقتی که صفحهٔ α از وسط پاره‌خط راست AB بگذرد و بر آن عمود باشد. همچنین نقطهٔ B هم، تصویر آئینه‌ای نقطهٔ A نسبت به‌صفحهٔ α است. یک جسم فضایی را نسبت به‌صفحه متقارن گویند وقتی که این صفحه جسم را به دو بخش طوری تقسیم کند که هر بخش بازتاب آئینه‌ای بخش دیگر نسبت به صفحه باشد. در این حالت، چنین صفحه‌ای را صفحهٔ تقارن جسم گویند. تقارن آئینه‌ای به فراوانی در طبیعت دیده می‌شود. از جمله، شکل ظاهری بدن انسان، جانوران و پرندگان، اغلب دارای صفحهٔ تقارن‌اند.

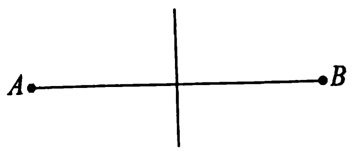
تقارن نسبت به‌خط راست هم، به صورت مشابهی تعریف می‌شود. دو نقطهٔ A و B را قرینهٔ هم نسبت به خط راست گویند، وقتی که این خط راست عمودمنصف پاره‌خط راست AB باشد (شکل ۲). جسم هندسی را متقارن نسبت به خط راست گویند (یا جسم هندسی دارای محور تقارنی از مرتبهٔ دوم است)، وقتی هر نقطهٔ دل‌خواه جسم قرینه‌ای نسبت به‌محور تقارن، منطبق برخود جسم داشته باشد.

هر جسمی که دارای محور تقارن مرتبهٔ دوم باشد، ضمن دوران دور محور تقارن به‌اندازهٔ نیم دور کامل، یعنی ۱۸۰ درجه، برخودش قرار می‌گیرد.

مفهوم محور تقارن را، به‌طور طبیعی، می‌توان تعمیم داد. خط راست را محور تقارن مرتبهٔ nام برای جسم مفروض گویند وقتی که این جسم، ضمن دوران به‌اندازهٔ $\frac{1}{n}$ ۳۶۰ درجه دور محور تقارن، برخودش منطبق شود. برای نمونه، هرم منتظمی که قاعدهٔ آن یک nضلعی



شکل ۱



شکل ۲

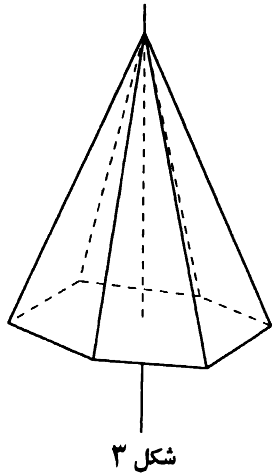
منتظم باشد، دارای محور تقارنی از قرینهٔ n است؛ این محور تقارن، خط راستی است که رأس هرم را به وسط قاعدهٔ آن می‌پیوندد (شکل ۳).

خط راست را محور دوران جسم گویند، وقتی که با دوران جسم دور محور، به اندازهٔ زاویه‌ای دل‌خواه، جسم برخودش قرار گیرد. ازجمله، محور استوانه، محور مخروط و هر قطری از کره، محورهای دوران برای این جسم‌ها هستند. در ضمن محور دوران، محور تقارن جسم از هر مرتبهٔ دل‌خواه است.

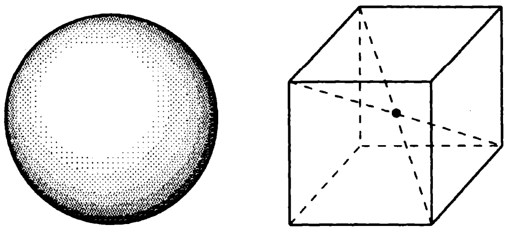
سرانجام، تقارنی مهم، تقارن نسبت به نقطه یا تقارن مرکزی است. دو نقطهٔ A و B را نسبت به مرکز O متقارن گویند، وقتی که پاره‌خط راست AB از نقطهٔ O بگذرد و در آن‌جا نصف شود. نمونهٔ جسم‌هایی که مرکز تقارن دارند، کره و مکعب است که مرکز هریک از آن‌ها، مرکز تقارن جسم است (شکل ۴).

وجود صفحه، محور یا مرکز تقارن برای بررسی جسم و تصور دربارهٔ ویژگی‌های تقارنی آن، بسیار سودمند است.

ولی کاربرد مفهوم تقارن تنها برای بررسی شکل‌های هندسی نیست. برای مثال، همه می‌دانیم، چندجمله‌ای $x_1^3+x_2^3+x_3^3+x_4^3$ نسبت به‌متغیرهای x_1، x_2، x_3، x_4 متقارن است و چندجمله‌ای $x_1^3+x_2^2+x_3^2+x_4^3$ نسبت به‌متغیرهای x_1 و x_4 و هم نسبت به‌متغیرهای x_2 و x_3 متقارن است، ولی متغیرهای x_1 و x_2 نقش‌های متفاوتی دارند، یعنی چندجمله‌ای نسبت به x_1 و x_2 متقارن نیست. از این نمونه‌ها می‌توان به‌فراوانی آورد. به‌این ترتیب، در برابر پرسش مهمی قرار می‌گیریم، تقارن در حالت کلی خود به‌چه معنی است و از دیدگاه ریاضی،



شکل ۳



شکل ۴

چگونه می‌توان رابطه‌ای را متقارن دانست؟ به‌نظر می‌رسد، پاسخ دقیق به‌این پرسش، به مفهوم تبدیل بستگی دارد که از همان بخش اول این کتاب بارها با آن برخورد کرده‌ایم. برای این که بتوانیم تعریفی کلی از تقارن بدهیم که همهٔ حالت‌های مختلف از تقارن جسم فضایی تا تقارن چندجمله‌ای‌ها را دربر بگیرد، باید مفهوم تبدیل را هم به‌صورتی بسیار کلی‌تر تنظیم کرد.

تبدیل. M را مجموعه‌ای باپایان یا بی‌پایان از عضوهای دل‌خواه در نظر می‌گیریم. به‌عنوان نمونه، M می‌تواند مجموعهٔ عددهای ۱، ۲، ...، n، مجموعهٔ متغیرهای مستقل x_1، x_2، x_3، x_4 یا مجموعهٔ همهٔ نقطه‌های یک صفحه باشد. اگر عضوی از مجموعه در تناظری معین با عضو دیگر این مجموعه باشد، آن‌وقت می‌گویند تبدیلی از مجموعهٔ M داده شده است. هر تبدیلی از مجموعهٔ باپایان M را می‌توان به وسیلهٔ جدولی شامل دو سطر داد: در یک سطر، همهٔ عضوهای مجموعه را، به ردیفی دل‌خواه می‌نویسیم و در سطر دوم و زیر هر جمله، جملهٔ متناظر آن را قرار می‌دهیم. برای نمونه، جدول $\begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۳ & ۲ & ۱ \end{pmatrix}$ به‌معنای تبدیلی از مجموعهٔ ۱، ۲، ۳، ۴ است که عددهای ۱، ۲، ۳، ۴ را به عددهای ۲، ۳، ۲، ۱ تبدیل می‌کند. اگر عددهای سطر بالا را به‌ردیف ۳، ۴، ۱، ۲ بنویسیم، می‌توانیم همین تبدیل را به‌صورت $\begin{pmatrix} ۳ & ۴ & ۱ & ۲ \\ ۲ & ۱ & ۲ & ۳ \end{pmatrix}$ بنویسیم.

اگر مجموعهٔ M بی‌پایان ولی «شمارا» باشد، یعنی بتوان برای عضوهای آن شماره‌ای گذاشت، عضوهای آن را در یک سطر می‌نویسیم (درمثل، اگر مجموعه شامل همهٔ عددهای طبیعی ۱، ۲، ۳، ... باشد) و تبدیل را به صورتی مشابه مجموعهٔ باپایان نشان داد.

ضمن بررسی تبدیل‌ها، باید برای هر تبدیل نامی در نظر گرفت. تبدیل‌ها را به‌طور ساده با حرف‌های A، B و غیره نشان می‌دهند؛ درضمن، اگر تبدیل مجموعهٔ M را با حرف A نشان داده‌ایم، آن‌وقت تصویر m، یعنی عضو متناظر m را بعداز تبدیل به mA نشان می‌دهیم، که در آن، m عضوی از مجموعهٔ M و mA تبدیل شدهٔ آن، به‌ازای تبدیل A است.

فرض کنید: $A=\begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۳ & ۲ & ۱ \end{pmatrix}$ در این صورت

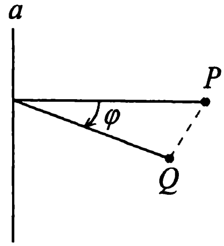
$$۱A = ۲\ ,\ \ ۲A = ۳\ ,\ \ ۳A = ۲\ ,\ \ ۴A = ۱$$

از برخی تبدیل‌ها، که در هندسه نقش مهمی به‌عهده دارند، یاد می‌کنیم.

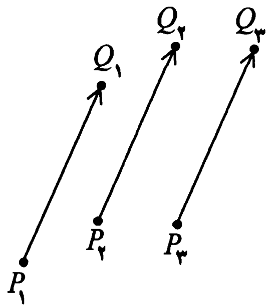
خط راستی مثل a را در فضا انتخاب می‌کنیم و هر نقطهٔ P از فضا را متناظر با نقطهٔ Q قرار می‌دهیم که از راه دوران نقطهٔ P دور محور a به‌اندازهٔ زاویهٔ ثابتی مثل φ به‌دست آمده باشد (شکل ۵). به‌این ترتیب تبدیلی از مجموعهٔ همهٔ نقطه‌های فضا را تعریف کرده‌ایم که دوران فضا دور محور a به‌اندازهٔ زاویه φ نامیده می‌شود.

یادآوری می‌کنیم، واژهٔ «دوران» در مکانیک به‌معنای روندی است که در نتیجهٔ آن، نقطه‌های جسم در موقعیت تازه‌ای قرار می‌گیرند. ولی در این جا، اصطلاح «دوران» به‌معنای تبدیلی از فضا است. درضمن، اگر از روند حرکت جدا شویم و تنها نتیجهٔ نهایی را در نظر بگیریم، به‌تناظر نقطهٔ در حالت آغازین خود با نقطه‌ای که در موقعیت نهایی قرار دارد، می‌رسیم.

یکی دیگر از تبدیل‌های مهم فضا، عبارت است از انتقال موازی همهٔ نقطه‌ها در جهت مفروض و به فاصلهٔ مفروض. باتوجه به شکل ۶ روشن می‌شود که، اگر جهت انتقال و فاصلهٔ انتقال معلوم باشد، هر نقطهٔ دل‌خواهی مثل $P_۱$، $P_۲$ یا $P_۳$، متناظر با نقطهٔ دیگری از فضا، مثل $Q_۱$، $Q_۲$ یا $Q_۳$ می‌شود. اگر انتقال موازی برای یک نقطه معلوم باشد، برای همهٔ



شکل ۵



شکل ۶

نقطه‌های فضا معلوم خواهد بود.

پیش از این دربارهٔ صفحه، محور یا مرکز تقارن یک شکل فضایی صحبت کردیم. هریک از این مفهوم‌ها متناظر با تبدیل معینی از فضا است: بازتاب نسبت به صفحه، دوران نسبت به خط راست و بازتاب نسبت به مرکز. ازجمله، بازتاب نسبت به صفحه تبدیلی است که هر نقطه از فضا را با قرینهٔ آن نسبت به صفحه متناظر می‌کند. به‌همین ترتیب، دوران نسبت به خط راست یا بازتاب نسبت به مرکز هم تعریف می‌شود.

دربارهٔ تبدیل‌های فضا صحبت کردیم. تبدیل‌های مشابهی برای صفحه وجود دارد: دوران صفحه دور نقطه‌ای از آن و به‌اندازهٔ زاویه‌ای مفروض، انتقال صفحه روی خودش و در جهتی مفروض، بازتاب نسبت به خط راستی که روی صفحه واقع است؛ همهٔ این‌ها به‌صورت مشابهی تعریف می‌شوند و بسیار عینی‌تر از نمونه‌های مشابه خود در فضا هستند.

تبدیل‌های یک‌به‌یک. ضمن بررسی همهٔ گونه‌های تبدیل یک مجموعه، قبل از هرچیز، تفاوتی عمیق بین نگاشت‌های یک‌به‌یک مجموعه به روی خودش، و نگاشت‌هایی که یک‌به‌یک نیستند، به‌چشم می‌خورد. تبدیل A از مجموعهٔ M، یعنی نگاشت این مجموعه روی خودش، تناظر یک‌به‌یک (یا نگاشت یک‌به‌یک و به روی) می‌نامیم، به‌شرطی که نه‌تنها هر عضو مجموعهٔ M متناظر عضو واحد و معینی از مجموعهٔ M باشد (چیزی که در تعریف تبدیل وجود دارد)، بلکه به شرطی که به‌جز آن، برای هر عضو y از مجموعهٔ M، یک و تنها یک عضو x باشد که به‌عضو y تبدیل شود. به‌زبان دیگر، تبدیل A به‌شرطی یک تناظر یک‌به‌یک است که «معادلهٔ» $xA=y$ برای هر y از M، یک و تنها یک «جواب» برای x از M داشته باشد.

همهٔ تبدیل‌هایی از فضا که در این‌جا یاد کردیم ـ بازتاب‌ها، دوران‌ها و انتقال‌ها ـ تناظر یک‌به‌یک بودند. زیرا در آن‌ها، نه‌تنها برای هر نقطهٔ X نقطه‌ای وجود دارد که X به‌آن تبدیل می‌شود، بلکه همچنین نقطه‌ای یگانه وجود دارد که به X تبدیل می‌شود.

به‌سادگی می‌توان مثال‌هایی آورد که نمونه تبدیل‌هایی باشند که تناظر یک‌به‌یک نیستند: تبدیل مجموعهٔ عددهای ۱، ۲، ۳، ۴ که در جدول $\begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۱ & ۲ & ۳ \end{pmatrix}$ داده شده است، تناظر یک‌به‌یک نیست، زیرا هیچ عضوی از مجموعه، به‌عضو ۴ تبدیل نمی‌شود. تبدیل مجموعهٔ همهٔ عددهای طبیعی ۱، ۲، ۳، ... به‌صورتی که در جدول

$$\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ & ۶ & \dots \\ ۱ & ۱ & ۲ & ۲ & ۳ & ۳ & \dots \end{pmatrix}$$

معرفی شده است، بازهم تناظر یک‌به‌یک نیست. درست است که در این‌جا، برای هر عدد n، عدد $2n$ وجود دارد که به‌آن تبدیل می‌شود، ولی عدد $2n$ از نظر این ویژگی یگانه نیست، زیرا در این تبدیل هم $2n$ و هم $2n-1$ تبدیل به عدد n می‌شود. به‌طور کلی، برای تبدیلی که به‌وسیلهٔ جدول داده شده است، به‌سادگی می‌توان معیاری برای تشخیص این که تبدیل، تناظر یک‌به‌یک است یا نه، پیدا کرد. برای این منظور، لازم و کافی است، در سطر دوم جدول، همهٔ عضوهای مجموعه و در ضمن از هر کدام تنها یکبار وجود داشته باشد.

گاهی در ریاضیات توجه به تبدیل‌هایی هم که تناظر یک‌به‌یک نیستند، لازم می‌شود. ازجمله می‌دانیم، عمل تصویر کردن فضا بر صفحه چه ارزش زیادی دارد. این تبدیل، تناظر یک‌به‌یک نیست، زیرا در این تبدیل، هر نقطهٔ صفحه تصویری از یک رشته نقطه‌های فضا است. ولی در بیشتر حالت‌ها تنها با تناظرهای یک‌به‌یک سروکار داریم؛ این‌گونه تبدیل‌ها به‌ویژه نقشی اساسی در بررسی روندهای فیزیکی دارند که در آن‌ها، عضوهای دستگاهی که مطالعه می‌شود، نه به‌هم می‌آمیزند، نه از میان می‌روند و نه به‌وجود می‌آیند.

از این به بعد، هرجا از تبدیل نام می‌بریم، منظورمان تبدیل یک‌به‌یک است و گاه و به‌ویژه وقتی صحبت برسر مجموعه‌های باپایان باشد، به‌جای «تبدیل» از واژهٔ «جایگشت» هم استفاده می‌کنند.

برای هر تبدیل A (تناظر یک‌به‌یک) از مجموعهٔ M به روی خودش، به‌سادگی می‌توان تبدیل وارونِ A (A^{-1}) را تعریف کرد. اگر تبدیل A، عضوی مثل x از مجموعهٔ M را به عضو y منجر می‌کند، آن وقت تبدیلی را که در آن، عضو y به‌عضو x منجر می‌شود، تبدیل وارون

A نام دارد و با نماد A^{-1} نشان داده می‌شود. به‌عنوان نمونه

اگر $A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$،

آن‌وقت $A^{-1} = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$،

اگر A دوران فضا دور محوری به زاویهٔ φ باشد، آن‌وقت A^{-1} دوران فضا دور همان محور و به زاویهٔ φ، ولی در جهت عکس، است؛ و غیره.

گاهی پیش می‌آید که تبدیل وارون بر تبدیل مستقیم منطبق است. این ویژگی را، به‌خصوص، برای تقارن (بازتاب) نسبت به صفحه یا نقطه، در فضا می‌توان دید. تبدیل $A=\begin{pmatrix} 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}$ هم این ویژگی را دارد، زیرا

$$A^{-1} = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

یادآوری می‌کنیم دربارهٔ تبدیل‌هایی که تناظر یک‌به‌یک نیستند، دربارهٔ تبدیل وارون نمی‌توان صحبت کرد. زیرا در این گونه تبدیل‌ها، ممکن است عضوی تصویر هیچ عضو یا چند عضو جداگانه باشد.

تعریف کلی‌تری برای تقارن‌ها. چه در ریاضیات و چه در کاربردهای آن به‌ندرت پیش می‌آید که به همهٔ تبدیل‌های مجموعهٔ مفروض نیاز داشته باشیم. موضوع این است که دربارهٔ خود مجموعه‌ها، به‌ندرت می‌توان تنها به عنوان کنار هم قرار گرفتن سادهٔ عضوها و بدون بستگی آن‌ها با یکدیگر، اندیشید. و این، امری طبیعی است، زیرا مجموعه‌هایی که در ریاضیات مطرح می‌شوند، انتزاعی از گروه‌هایی از چیزهای دنیای واقع است که عضوهای آن در جنبه‌های گوناگون خود به‌هم بستگی دارند و این بستگی، بیرون از مرزهای مجموعه‌ای است که بررسی می‌شود. در ریاضیات، بخش‌های بسیاری از این بستگی‌ها کنار گذاشته می‌شود، ولی اساسی‌ترین آن‌ها حفظ می‌شود و به‌حساب می‌آید. به‌همین جهت، باید در درجهٔ اول، به بررسی چنان تبدیل‌هایی از مجموعه‌ها پرداخت که بستگی‌های ضروری بین عضوها را به‌هم نریزد. اغلب این گونه تبدیل را ***تبدیل‌های مجاز*** یا ***تبدیل‌های خودریخت*** (یا

خودسان یا ***خوددیس*** (*automorphism*) نسبت به بستگی‌های ضروری عضوهای مجموعه می‌نامند. ازجمله، برای نقطه‌های فضا، مفهوم فاصلهٔ بین نقطه‌ها اساسی است. وجود این مفهوم، به معنی رابطهٔ بین نقطه‌ها است و به این معنی است که هر دو نقطه در فاصلهٔ معینی از یکدیگر قرار دارند. تبدیل‌هایی که این بستگی‌ها را به‌هم نریزند، چنان تبدیل‌هایی هستند که، برای آن‌ها، فاصلهٔ بین نقطه‌ها تغییر نکند. این تبدیل‌ها را «***حرکت‌های***» فضا گویند.

با استفاده از مفهوم خودریختی (اُتومورفیسم) به‌سادگی می‌توان تعریف کلی تقارن را داد. مجموعه‌ای مثل M را در نظر بگیرید، که در آن بستگی‌های معینی که بین عضوها وجود دارد به‌حساب آمده باشد و فرض کنید، P بخشی از M باشد. مجموعهٔ P را نسبت به تبدیل مجاز A از مجموعهٔ M، متقارن یا «ناوردا» و یا «پایا» گویند، به‌شرطی که تبدیل A، هر عضو مجموعهٔ P را دوباره به عضوی از مجموعهٔ P منجر کند. بنابراین، تقارن مجموعهٔ P، به این ترتیب مشخص می‌شود که مجموعهٔ تبدیل‌های مجاز از مجموعهٔ M، مجموعهٔ P را به‌خودش منجر کند. مفهوم تقارن جسم در فضا، به‌طور کامل با این تعریف سازگار است. در این‌جا، نقش مجموعهٔ M به‌عهدهٔ تمامی فضا، نقش تبدیل‌های مجاز به‌عهدهٔ «حرکت» و نقش P به‌عهدهٔ جسم مفروض است. بنابراین تقارن P با مجموعهٔ حرکت‌هایی مشخص می‌شود که به‌ازای آن‌ها، جسم P برخودش قرار گیرد.

نگاشت‌هایی که پیش از این آوردیم (انتقال‌های موازی و دوران فضا دور خط راست مفروض)، حالت‌های خاصی از حرکت‌اند. زیرا روشن است، در این تبدیل‌ها فاصلهٔ بین نقطه‌ها بی‌تغییر می‌ماند. بررسی مفصل‌تر نشان می‌دهد که هر حرکت صفحه، یا انتقال است، یا دوران دور مرکز، یا انعکاس (بازتاب) نسبت به‌خط راست، یا ترکیبی از انعکاس نسبت به‌خط راست با انتقال در طول این خط راست. به‌همین ترتیب هر حرکت فضا عبارت است از یا انتقال موازی، یا دوران دور محور، یا حرکت پیچشی (یعنی دوران دور محور، همراه با انتقال در طول همین محور)، یا انعکاس (بازتاب یا تقارن) نسبت به صفحه، که می‌تواند همراه با انتقال در طول صفحه یا دوران دور محوری عمود بر این صفحه هم باشد.

انتقال موازی، دوران و حرکت پیچشی فضا را، حرکت‌های سرهٔ آن یا حرکت‌های نوع اول فضا گویند. «حرکت‌های دیگر» (از جمله انعکاس) را حرکت‌های ناسره یا حرکت‌های نوع دوم می‌نامند. روی صفحه، حرکت‌های نوع اول عبارت‌اند از انتقال موازی و دوران؛ انعکاس نسبت به‌خط راست و انعکاس همراه با دوران یا انتقال، حرکت‌های نوع دوم‌اند.

به‌سادگی می‌توان تصور کرد که تبدیل‌ها را، وقتی با حرکت‌های نوع اول انجام شوند،

می‌توان نتیجه‌ای از حرکت پیوستهٔ فضا در درون خود یا حرکت پیوستهٔ صفحه در درون خود دانست. ولی حرکت‌های نوع دوم را به‌این ترتیب نمی‌توان به‌دست آورد، زیرا حرکت پیوسته نمی‌تواند شامل انعکاس آئینه‌ای، که جزو حرکت‌های نوع دوم است، بشود.

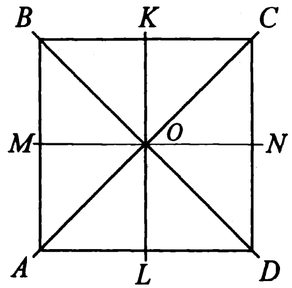
اغلب می‌گویند، صفحه در همهٔ بخش‌های خود متقارن است، یا همهٔ نقطه‌های صفحه، نسبت به‌هم، حقی برابر دارند، یا هم‌ارزند. به‌زبان دقیق تبدیل‌ها، این گزاره به‌معنای آن است که هر نقطهٔ صفحه را می‌توان به‌یاری «حرکتی» مناسب، به‌هر نقطهٔ دیگر صفحه منجر کرد.

حالت‌هایی از تقارن شکل‌ها را هم که پیش از این آوردیم، می‌توان با تعریف کلی تقارن بیان کرد. برای نمونه، جسمی که نسبت به صفحهٔ α متقارن است، در انعکاس نسبت به صفحهٔ α برخودش قرار می‌گیرد؛ جسمی که نسبت به مرکز O متقارن است، در انعکاس نسبت به مرکز O برخودش واقع می‌شود. به‌این ترتیب، درجهٔ تقارن شکل‌ها، به‌یاری مجموعه‌ای از حرکت‌های نوع اول و دوم فضا مشخص می‌شود، به‌نحوی که جسم یا شکل برخودش منطبق شود. هرچه این مجموعهٔ حرکت‌ها غنی‌تر و متنوع‌تر باشد، به‌همان اندازه درجهٔ تقارن جسم یا شکل بیشتر است. به‌ویژه، اگر این مجموعه، شامل هیچ حرکتی به‌جز تبدیل همانی نباشد، می‌توان جسم را نامتقارن نامید.

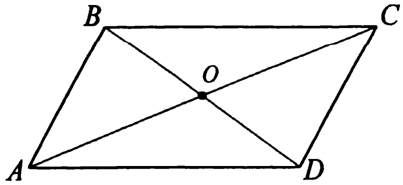
درجهٔ تقارن مربع در روی صفحه، با مجموعهٔ حرکت‌هایی از صفحه مشخص می‌شود که مربع را برخودش منطبق کند. ولی اگر مربع برخودش قرار گیرد، باید نقطهٔ برخورد قطرهای آن هم برخودش واقع شود. بنابراین، حرکت مورد جست‌وجو، مرکز مربع را بی‌حرکت نگاه می‌دارد و بنابراین، حرکت عبارت است از: یا دوران دور مرکز و یا انعکاس نسبت به‌خط راستی که از مرکز می‌گذرد. شکل ۷ نشان می‌دهد، مربع $ABCD$ نسبت به‌دوران دور نقطهٔ مرکز خود O، به‌اندازهٔ زاویه‌ای که مضربی از ۹۰ درجه باشد و همچنین در انعکاس نسبت به قطرهای AC و BD و خط‌های راست KL و MN متقارن است. این هشت حرکت، معرف ویژگی تقارنی مربع‌اند.

مجموعهٔ تقارن‌های مستطیل عبارت است از دوران دور مرکز به‌اندازهٔ ۱۸۰ درجه و انعکاس نسبت به‌خط‌های راستی که وسط‌های دو ضلع روبه‌رو را به‌هم وصل می‌کنند. مجموعهٔ تقارن‌های متوازی‌الاضلاع (شکل ۸)، تنها شامل دوران دور مرکز به‌اندازهٔ ۱۸۰ درجه، یعنی از انعکاس نسبت به مرکز و تبدیل همانی تشکیل شده است.

پیش از این، مثال‌هایی از تقارن را در جبر آوردیم. تقارن چندجمله‌ای‌های جبری نسبت



شکل ۷



شکل ۸

به چند متغیر.

ببینیم خصلت چندجمله‌ای‌های متقارن را چگونه می‌توان مشخص کرد. می‌گوییم در چندجمله‌ای $F(x_1, x_2, \dots, x_n)$ تبدیل

$$A = \begin{pmatrix} x_1, x_2, \dots, x_n \\ x_{i_1}, x_{i_2}, \dots, x_{i_n} \end{pmatrix}$$

یا

$$A = \begin{pmatrix} 1, 2, \dots, n \\ i_1, i_2, \dots, i_n \end{pmatrix}$$

انجام گرفته است وقتی که در چندجمله‌ای مفروض، همه‌جا به‌جای حرف x_1، حرف x_{i_1} و به‌جای x_2 حرف x_{i_2} و غیره را جای‌گزین کنیم. چندجمله‌ای که به‌این ترتیب به‌دست می‌آید FA می‌نامیم. در مثل اگر

$$F = x_1^2 - 2x_2 + x_3 - x_4, \quad A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

باشد، آن وقت $FA = x_3^2 - 2x_1 + x_4 - x_2$.

تقارن چندجمله‌ای، به‌یاری مجموعه‌ای از تبدیل‌های متغیرها مشخص می‌شود که با انجام آن‌ها، چندجمله‌ای تغییر نکند. برای نمونه، تقارن چندجمله‌ای $x_1^3 + 2x_2 + x_3^3 + 2x_4$، با چهار تبدیل مشخص می‌شود.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

و تقارن‌های چندجمله‌ای $x_4 + x_3^3 + 2x_2 + x_1^3$، تنها با دو تبدیل مشخص می‌شود.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

۳. گروه‌های تبدیل

ضرب تبدیل‌ها. با بررسی ویژگی‌های تبدیل‌ها، خیلی ساده می‌توان متوجه شد، برخی از این تبدیل‌ها نتیجه‌ای از چند تبدیل دیگرند. ازجمله، حرکت پیچشی، نتیجه‌ای از دوران دور محور و جابه‌جایی در طول محور است. این روند تشکیل تبدیل‌های تازه از چند تبدیل مفروض را، ضرب تبدیل‌ها می‌گویند. اگر روی عضو دل‌خواه x از مجموعهٔ M، تبدیلی مثل A انجام شود و، سپس، روی عضو جدید xA تبدیل B صورت گیرد، به عضو $(xA)B$ می‌رسیم. اگر به‌طور مستقیم تبدیل عضو نخستین x را به $(xA)B$ در نظر بگیریم، به‌آن حاصل‌ضرب تبدیل A در تبدیل B گویند و با نماد AB نشان می‌دهند. بنابراین، طبق تعریف داریم:

$$x(AB) = (xA)B$$

به‌عنوان نمونه:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$$

در واقع، تبدیل اول ۱ را به ۲ می‌رساند و در تبدیل دوم ۲ به ۴ منجر می‌شود، بنابراین، نتیجهٔ دو تبدیل به‌معنای این است که ۱ به ۴ منجر شده است و غیره. این هم چند مثال دیگر:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix};$$

$$\begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۱ & ۴ & ۳ \end{pmatrix} \begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۳ & ۱ & ۲ & ۴ \end{pmatrix} = \begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۱ & ۳ & ۴ & ۲ \end{pmatrix};$$

$$\begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۳ & ۱ & ۲ & ۴ \end{pmatrix} \begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۱ & ۴ & ۳ \end{pmatrix} = \begin{pmatrix} ۱ & ۲ & ۳ & ۴ \\ ۴ & ۲ & ۱ & ۳ \end{pmatrix}$$

دو مثال آخر نشان می‌دهند که ضرب تبدیل‌ها، عملی جابه‌جایی‌ناپذیر (یا تعویض‌ناپذیر) است، یعنی اگر جای دو عامل ضرب را باهم عوض کنیم، به‌نتیجهٔ دیگری می‌رسیم: نتیجهٔ ضرب به‌ردیف عامل‌های ضرب بستگی دارد. در ضرب حرکت‌های صفحه هم، همین مطلب تاکید می‌شود. برای مثال، فرض کنید A دوران دور مبداء O به اندازهٔ ۹۰ درجه و B انتقال در طول محور Ox به اندازهٔ واحد باشد.

ببینیم تبدیل AB و BA، نقطهٔ O را به‌کدام نقطه می‌رسانند! بنابر تعریف داریم (شکل ۹):

$$O(AB) = (OA)B = OB = M\ ,$$

$$O(BA) = (OB)A = MA = N$$

یعنی $AB \neq BA$.

برای این که طبیعت هندسی تبدیل BA را روشن‌تر کنیم، نقطهٔ P را در نظر می‌گیریم. داریم:

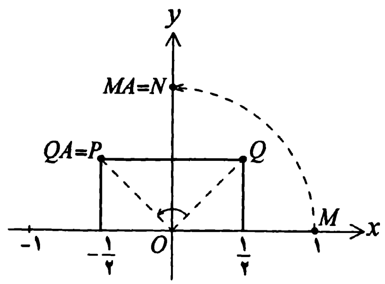
$$P(BA) = (PB)A = QA = P$$

یعنی نقطهٔ P، ضمن تبدیل BA، نقطه‌ای بی‌حرکت (ثابت) است. براین اساس، می‌توان BA را به‌طور ساده، دوران صفحه به‌اندازهٔ ۹۰ درجه دور نقطهٔ P دانست. به‌همین ترتیب

$$Q(AB) = (QA)B = PB = Q$$

یعنی AB عبارت است از دوران به‌اندازهٔ ۹۰ درجه دور نقطهٔ Q.

ضرب حرکت‌های صفحه یا فضا، در حالت کلی، از قانون بغرنج‌تری پیروی می‌کند. با وجود این، در دو حالت مهم، قانون‌های ضرب بسیار ساده‌اند. اول، اگر دوران‌های صفحه را دور یک نقطه یا دوران‌های فضا را دور یک خط راست، با زاویه‌های φ و ψ انجام دهیم، نتیجهٔ ضرب دو دوران، دورانی است به‌اندازهٔ زاویهٔ $\varphi+\psi$. دوم، اگر انتقال‌هایی را درهم ضرب کنیم که با بردارهای $\overrightarrow{MN}$ و $\overrightarrow{NP}$ مشخص می‌شوند، آن وقت حاصل‌ضرب، انتقالی



شکل ۹

است به‌اندازهٔ بردار $\overrightarrow{MP}$، یعنی مجموع بردارهای نخستین.

نظریهٔ «ضرب» تبدیل‌ها، برخی شباهت‌ها را بین ضرب عددها و ضرب تبدیل‌ها به‌یاد می‌آورد. ولی این شباهت در همه‌جا نیست. ازجمله، ضرب عددها از قانون جابه‌جایی (یا تعویض‌پذیری) پیروی می‌کند؛ در حالی که ضمن ضرب تبدیل‌ها دیدیم، این قانون ممکن است به‌هم بخورد. قانون دوم ضرب عددها، یعنی قانون شرکت‌پذیری، به‌طور کامل دربارهٔ ضرب تبدیل‌ها هم برقرار است و برای تبدیل‌های دل‌خواه A ، Bو Cاز مجموعهٔ Mداریم:

$$A(BC) = (AB)C$$

در واقع، اگر mعضو دل‌خواهی از Mباشد، آن وقت

$$m[A(BC)] = (mA)(BC) = [(mA)B]C = [m(AB)]C = m[(AB)C]$$

قانون شرکت‌پذیری اجازه می‌دهد به‌جای ضرب‌های $A(BC)$و $(AB)C$ از تبدیل‌های A ، B ، C ، تنها دربارهٔ یک ضرب

$$A(BC) = (AB)C = ABC$$

صحبت کنیم. همین قانون شرکت‌پذیری ثابت می‌کند که ضرب چهار تبدیل یا ضرب تعداد بیشتری تبدیل بستگی به‌جای پرانتزها ندارد.

بین تبدیل‌ها، تبدیلی وجود دارد که نقش عدد ۱ را در ضرب عددها به‌عهده گرفته است؛ این، تبدیل واحد یا تبدیل همانی Eاست، که هریک از عضوهای مجموعهٔ Mرا بی‌تغییر نگه می‌دارد. روشن است، تبدیل Aهرچه باشد، داریم:

$$AE = EA = A$$

به‌این حقیقت مهم هم توجه کنیم: اگر دو تبدیل را، که هرکدام از آن‌ها تناظر یک‌به‌یک هستند، درهم ضرب کنیم، به‌تبدیلی که تناظر یک‌به‌یک است، می‌رسیم. در واقع، برای پیدا کردن عضو x از مجموعهٔ M، که ضمن ضرب AB منجر به a می‌شود، کافی است عضو x_1 را که در تبدیل B به a می‌رسد پیدا کنیم، سپس عضو x_2 را جست‌وجو کنیم که با تبدیل A منجربه x_1 می‌شود چون

$$x_2(AB) = (x_2A)B = x_1A = a$$

پس x_2 همان مجهول x است.

حاصل‌ضرب تبدیل A در تبدیل وارون A^{-1}، برابر تبدیل همانی می‌شود، یعنی

$$AA^{-1} = A^{-1}A = E$$

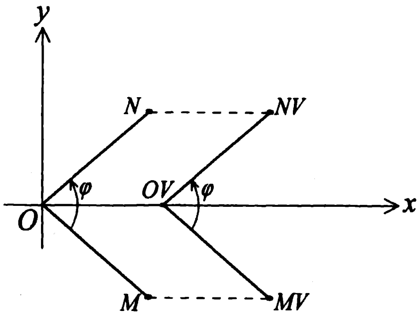
که نتیجه‌ای است از تعریف تبدیل همانی.

مثالی که دربارهٔ ضرب انتقال صفحه در دوران آورده بودیم، نشان می‌دهد، ویژگی‌های حاصل‌ضرب تبدیل‌ها را، همیشه به‌سادگی نمی‌توان از ویژگی‌های عامل‌های ضرب نتیجه گرفت. با وجود این، حاصل‌ضرب تبدیل‌هایی به‌صورت $C = B^{-1}AB$، استثنایی مهم است: در این‌جا، ویژگی‌های C، خیلی ساده به ویژگی‌های A و B بستگی دارد. اگر عضو m از مجموعهٔ M با تبدیل A منجر به n بشود، آن‌وقت انجامِ بلافاصلهٔ تبدیل B، عضو mB از تبدیل C را منجر به nB می‌کند. در واقع:

$$(mB)B^{-1}AB = mAB = nB$$

تبدیل $B^{-1}AB$ از A و از راه تبدیل آن به یاری B به‌دست می‌آید و **مزدوج** A به‌وسیلهٔ B نامیده می‌شود.

برای نمونه، دوران P_O از صفحه دور نقطهٔ O را به یاری انتقال V مزدوج می‌کنیم. با توجه به قانونی که داریم، برای پیدا کردن جای زوج نقطه‌های آغازی و پایانی در تبدیل حرکت $C = V^{-1}P_OV$، باید به‌یاری V زوج نقطه‌های متناظر را برای تبدیل P_O پیدا کرد. چون نقطهٔ O، ضمن دوران P_O بی‌حرکت می‌ماند (شکل ۱۰)، بنابراین نقطهٔ OV هم نسبت به تبدیل C ثابت است. سپس، اگر نقطهٔ M با تبدیل P به نقطهٔ N منجر می‌شود، آن وقت نقطهٔ MV با تبدیل C به نقطهٔ NV می‌رسد. به‌این ترتیب، و همان‌طور که از شکل ۱۰



شکل ۱۰

پیدا است، تبدیل C عبارت است از دوران دور نقطهٔ OV، با همان زاویهٔ φ که در دوران P وجود دارد.

به‌همین‌ترتیب ثابت می‌شود، اگر انتقال صفحه که با بردار $\overrightarrow{MN}$ مشخص شده باشد، به یاری دوران P_O به زاویهٔ φ مزدوج شود، آن‌وقت دوباره انتقال صفحه به‌دست می‌آید که با بردار دیگری مشخص می‌شود.

قانونی را که برای جست‌وجوی تبدیل $B^{-1}AB$ آورده‌ایم، می‌توان در حالتی هم که تبدیل‌ها به‌صورت جدول داده شده‌اند، با ظرافت خاصی تنظیم کرد. فرض کنید:

$$A=\begin{pmatrix} 1 & 2 & \dots & n \\ a_1 & a_2 & \dots & a_n \end{pmatrix},\ B=\begin{pmatrix} 1 & 2 & \dots & n \\ b_1 & b_2 & \dots & b_n \end{pmatrix}$$

در این صورت

$$B^{-1}AB=\begin{pmatrix} b_1 & b_2 & \dots & b_n \\ 1 & 2 & \dots & n \end{pmatrix}\begin{pmatrix} 1 & 2 & \dots & n \\ a_1 & a_2 & \dots & a_n \end{pmatrix}\begin{pmatrix} 1 & 2 & \dots & n \\ b_1 & b_2 & \dots & b_n \end{pmatrix}$$

$$=\begin{pmatrix} b_1 & b_2 & \dots & b_n \\ b_{a_1} & b_{a_2} & \dots & b_{a_n} \end{pmatrix}$$

یعنی برای این که تبدیل A را به یاری تبدیل B مزدوج کنیم، باید همهٔ عنصرهای سطرهای بالا و پایین تبدیل A را در معرض تبدیل B قرار دهیم. برای نمونه، اگر

$$A=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix},\ B=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{pmatrix}$$

آن‌وقت

$$B^{-1}AB=\begin{pmatrix} 1B & 2B & 3B & 4B & 5B \\ 3B & 5B & 4B & 1B & 2B \end{pmatrix}=\begin{pmatrix} 2 & 5 & 1 & 3 & 4 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix}$$

یادآوری می‌کنیم، دست‌کم در حالت کلی، حاصل‌ضرب دو تبدیل به ردیف عامل‌های ضرب بستگی دارد. ولی در حالت‌های خاصی، ممکن است AB و BA یکسان باشند، که در این صورت تبدیل‌های A و B را جابه‌جایی‌پذیر (یا تعویض‌پذیر) گویند. اگر $AB=BA$، آن‌وقت

$$B^{-1}AB=B^{-1}BA=A$$

به‌این ترتیب، در این حالت، تبدیل مفروض، باتوجه به‌قانون جابه‌جایی، بی‌تغییر می‌ماند.

گروه‌های تبدیل. مجموعهٔ تبدیل‌هایی که تقارن‌های یک شکل را تشکیل می‌دهند، نمی‌توانند دل‌خواه باشند و دارای این ویژگی‌ها هستند:

۱. اگر دو تبدیل متعلق به مجموعه باشند، حاصل‌ضرب آن‌ها هم متعلق به همین مجموعه است.

۲. تبدیل همانی متعلق به مجموعه است.

۳. اگر تبدیلی متعلق به مجموعه باشد، وارون آن هم، متعلق به مجموعه است.

این ویژگی‌ها برای تبدیل‌ها بسیار مهم‌اند؛ هر مجموعه از تبدیل‌های یک‌به‌یک و به روی مجموعهٔ M که دارای این سه ویژگی باشد، یک ***گروه تبدیل*** **مجموعهٔ** M نامیده می‌شود، اعم از این که این مجموعه معرف تقارن یک شکل باشد یا نباشد.

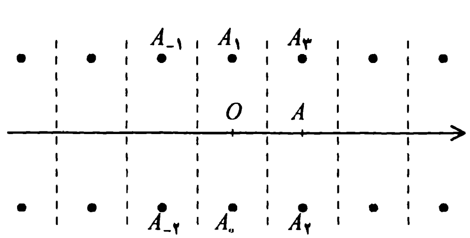
ویژگی‌های ۱ تا ۳ از دیدگاه جبر، اساسی‌اند. زیرا به ما اجازه می‌دهند، با آغاز از تبدیل‌هایی مثل A، B، C، ... که به‌مجموعهٔ مفروض تعلق دارند، تبدیل‌های مختلف تازه‌ای به‌صورت $ABAC$، $A^{-1}BCB^{-1}$ و غیره بسازیم؛ درضمن، ویژگی‌های ۱ تا ۳، تعلق همهٔ این تبدیل‌ها را به مجموعهٔ مفروض تبدیل‌ها، تضمین می‌کند.

تعداد تبدیل‌هایی که یک گروه را تشکیل می‌دهند، **مرتبهٔ گروه** نامیده می‌شود، درضمن، مرتبهٔ گروه می‌تواند محدود (متناهی) یا نامحدود (نامتناهی) باشد. به‌همین دلیل، گروه‌ها را به دو دستهٔ متناهی و نامتناهی تقسیم کرده‌اند. پیش از این دربارهٔ گروه تقارن‌های مربع در روی صفحه، صحبت کردیم. این گروه، روی‌هم، شامل هشت تبدیل است. از طرف دیگر،

مجموعهٔ بی‌پایان نقطه‌های A_i از صفحه که در شکل ۱۱ آمده است، با این حرکت‌های صفحه بر خودش قرار می‌گیرد: انتقال در طول محور OA، در این یا آن جهت، به فاصله‌ای که مضربی از فاصلهٔ OA باشد؛ انعکاس نسبت به خط‌های راست نقطه‌چین؛ و انعکاس نسبت به محور OA. از این جا دیده می‌شود، گروه تقارن‌های این شکل نامتناهی است.

مجموعهٔ تبدیل‌هایی که موضوعی را حفظ می‌کنند، یعنی تقارن آن را مشخص می‌کنند، همیشه یک گروه است. این روش دادن گروه، به‌صورت گروه تقارن‌ها، از جملهٔ مهم‌ترین روش‌هاست، و براین اساس، گروه‌های مهم بسیاری به‌دست می‌آیند. از این جمله، قبل از همه باید از گروه‌های مربوط به حرکت‌های صفحه و فضا نام برد. گروه‌های مربوط به تقارن‌های چندوجهی‌های منتظم هم بسیار جالب‌اند. می‌دانیم، در فضا تنها پنج نوع چندوجهی منتظم وجود دارد (۴، ۶، ۸، ۱۲ و ۲۰ وجهی). اگر یکی از چندوجهی‌های منتظم را انتخاب کنیم و همهٔ حرکت‌هایی از فضا را در نظر بگیریم که چندوجهی مفروض را به خودش منجر کند، یک گروه به‌دست می‌آید: گروه تقارن‌های این چندوجهی. اگر به‌جای همهٔ حرکت‌ها، تنها حرکت‌هایی از نوع اول را در نظر بگیریم، که چندوجهی را به خودش تبدیل می‌کند، آن‌وقت دوباره یک گروه به‌دست می‌آید که بخشی از گروه کامل تقارن‌های چندوجهی است. این گروه را، گروه دوران‌های چندوجهی گویند. چون ضمن قرار گرفتن چندوجهی برخودش، مرکز آن هم برخودش قرار می‌گیرد، بنابراین همهٔ حرکت‌هایی که در گروه تقارن‌های چندوجهی هستند، مرکز چندوجهی را ثابت نگه می‌دارند، در نتیجه می‌توان از دوران‌های دور محوری که از مرکز می‌گذرد، یا انعکاس‌های نسبت به صفحه‌ای که از مرکز می‌گذرد، یا سرانجام ترکیب انعکاس‌های نسبت به صفحه‌هایی که از مرکز می‌گذرند با دوران‌های دور محوری که براین صفحه‌ها عمودند، صحبت کرد.

با استفاده از این مطلب، به‌سادگی می‌توان همهٔ گروه‌های تقارن و گروه‌های دوران چندوجهی‌های منتظم را پیدا کرد. در جدول ۱، مرتبهٔ گروه‌های تقارن‌ها و گروه دوران‌های



شکل ۱۱

جدول ۱

تعداد وجه‌ها	۴	۶	۸	۱۲	۲۰
مرتبهٔ گروه‌های تقارن‌ها	۲۴	۴۸	۴۸	۱۲۰	۱۲۰
مرتبهٔ گروه‌های دوران‌ها	۱۲	۲۴	۲۴	۶۰	۶۰

چندوجهی‌های منتظم نشان داده شده است. همهٔ این گروه‌ها، متناهی‌اند.

گروه‌های جای‌گشت‌ها. از نظر تاریخی، نخستین گروه‌های تبدیلی که در ریاضیات بررسی شد، گروه‌های جای‌گشت چندجمله‌ای‌ها در متغیرهای x_1، x_2، .... ، x_n بود. بررسی این‌گونه گروه‌ها، مربوط به مسألهٔ حل رادیکالی معادله‌های از درجهٔ بالا است. روشن است، مجموعهٔ همهٔ جای‌گشت‌های متغیرها، که مقدار یک یا چند «چندجمله‌ای در این متغیرها» را تغییر ندهد، یک گروه است. چندجمله‌ای که به‌ازای همهٔ جای‌گشت‌های متغیرها تغییر نکند، چندجمله‌ای متقارن نامیده می‌شود. ازجمله، $x_1+x_2+\cdots+x_n$، چندجمله‌ای متقارن است. متناظر با آن، مجموعهٔ همهٔ جای‌گشت‌های n متغیر، **گروه متقارن** این مجموعه نامیده می‌شود.

تعداد متغیرها، درجهٔ تقارن گروه نام دارد. به‌جای جای‌گشت‌های متغیرهای x_1، .... ، x_n می‌توان به‌طور ساده، جای‌گشت‌های عددهای ۱ ، ۲ ، .... ، n را بررسی کرد. چون هر جای‌گشت از عددها را می‌توان به‌صورت

$$\begin{pmatrix} 1 & 2 & \dots & n \\ a_1 & a_2 & \dots & a_n \end{pmatrix}$$

نوشت، که در آن، a_1، a_2، ... ، a_n همان عددهای ۱ ، ۲ ، ... ، n به ردیفی دیگرند، بنابراین تعداد همهٔ جای‌گشت‌ها برابر است با تعداد ترتیب‌های n عنصر، یعنی مرتبهٔ تقارن گروه برابر است با $n! = 1\times 2\times 3\times\cdots\times n$. این مرتبه، با بزرگ شدن n، به‌سرعت رشد می‌کند؛ به‌عنوان نمونه، مرتبهٔ گروه جای‌گشت‌های ۱۰ متغیر برابر است با ۳۶۲۸۸۰۰.

این چندجمله‌ای را در نظر می‌گیریم:

$$F(x_1, x_2, \dots, x_n)=(x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)(x_3-x_2)\cdots(x_n-x_2)\cdots(x_n-x_{n-1}) \quad (1)$$

روشن است هر جای‌گشت متغیرها، یا مقدار چندجمله‌ای F را تغییر نمی‌دهد یا تنها علامت

آن را عوض می‌کند. جای‌گشت‌های نوع اول را، زوج گویند. جای‌گشت‌هایی که علامت F را عوض کنند، فرد نام دارند. مجموعهٔ جای‌گشت‌های زوج، گروه تقارن‌های چندجمله‌ای (۱) را تشکیل می‌دهند آن را گروه متناوب گویند.

حاصل‌ضرب دو جای‌گشت زوج، جای‌گشتی زوج است، زیرا جای‌گشت‌های زوج تشکیل گروه می‌دهند. حاصل‌ضرب دو جای‌گشت فرد، جای‌گشتی زوج است.

در واقع، اگر A و B، دو جای‌گشت فرد باشند، داریم:

$$FAB=(FA)B=(-F)B=-(-F)=F$$

به همین ترتیب می‌توان ثابت کرد، حاصل‌ضرب جای‌گشت زوج در جای‌گشت فرد، جای‌گشتی فرد است؛ حاصل‌ضرب جای‌گشت زوج در وارون خود جای‌گشتی زوج و حاصل‌ضرب جای‌گشتی فرد در وارون خود جای گشتی فرد است.

جای‌گشت $\begin{pmatrix} ۱ & ۲ & ۳ & \dots & n \\ ۲ & ۱ & ۳ & \dots & n \end{pmatrix} = S$ که جای عضوهای ۱ و ۲ را باهم عوض کرده است، نمونه‌ای از جای گشت فرد است.

بسط جای‌گشت به‌دُورها. ضمن مطالعهٔ گروه‌های جای‌گشت‌ها، نمایش جای‌گشت‌ها به‌صورت ضرب دورها، اهمیت زیادی دارد. بنابر تعریف، نماد $(m_۱ , m_۲ , \dots , m_k)$ به‌معنای جای‌گشتی است که در آن $m_۱$ به $m_۲$، $m_۲$ به $m_۳$، ....، $m_{k-۱}$ به m_k و m_k به $m_۱$ تبدیل شود و بقیهٔ عضوهای مجموعه به‌جای خود باقی بمانند. برای نمونه، اگر جای‌گشت عددهای ۱، ۲، ۳، ۴، ۵ در نظر باشد، آن‌وقت

$$(۱,۲,۳,۴,۵)=\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ \\ ۲ & ۳ & ۴ & ۵ & ۱ \end{pmatrix},\ (۳,۵)=\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ \\ ۱ & ۲ & ۵ & ۴ & ۳ \end{pmatrix}$$

جای‌گشت به‌صورت $(m_۱ , m_۲ , \dots , m_k)$ را **دُوری** یا یک دُور به‌طول k می‌نامند؛ $m_۱$، $m_۲$، ...، m_k عضوهای دور هستند. بنابر شرط، جای‌گشت همانی را به‌صورت $\dots=(۲)=(۱)$ می‌نویسند که طولی برابر ۱ دارد. دُور به‌طول ۲ را **ترانهش** گویند. اگر عضوهای دُور را به ردیف دوری جابه‌جا کنیم، همان جای‌گشت به‌دست می‌آید، یعنی

$$(۱,\ ۲,\ ۳)=(۲,\ ۳,\ ۱)=(۳,\ ۱,\ ۲),\ (۵,\ ۶)=(۶,\ ۵)$$

به‌سادگی می‌توان روشن کرد، دُورهای بدون عضو مشترک مثل (۳ , ۲) و (۵ , ۴ , ۱) جای‌گشت‌های جابه‌جایی‌پذیرند، بنابراین ضمن ضرب این دُورها، می‌توان توجهی به‌ردیف عامل‌های ضرب نداشت.

اهمیت دُورها در نظریهٔ کلی، براساس این قضیه است: هر جای‌گشتی را می‌توان به‌صورت ضرب دورهای بدون عضو مشترک نشان داد؛ در ضمن، این ضرب، یک‌ارزشی است، به‌شرطی که ضرب‌های با ردیف‌های مختلف عامل‌ها، مختلف به‌حساب نیاوریم.

اثبات قضیه، بلافاصله از روش چنین نمایشی دیده می‌شود. فرض کنید بخواهیم جای‌گشت

$$A=\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ & ۶ \\ ۴ & ۵ & ۶ & ۳ & ۲ & ۱ \end{pmatrix}$$

را بسط دهیم (تجزیه کنیم). A، ۱ را به ۴ ، ۴ را به۳ ، ۳ را به ۶ و ۶ را به ۱ منجر می‌کند. در نتیجه، عامل اول ضرب عبارت است از (۱,۴,۳,۶) . عدد باقی‌ماندهٔ ۲ را انتخاب می‌کنیم، می‌بینیم A، ۲ را به ۵ و ۵ را به ۲ منجر می‌کند. بنابراین عامل دوم ضرب (۲,۵) است. از آن جا که همهٔ عددها به‌حساب آمده‌اند، پس

$$\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ & ۶ \\ ۴ & ۵ & ۶ & ۳ & ۲ & ۱ \end{pmatrix} =(۱,۴,۳,۶)(۲,۵) \qquad (۲)$$

جای‌گشت را به دُورهای با عضوهای مشترک هم می‌توان تجزیه کرد، ولی این تجزیه یگانه و یک‌ارزشی نیست. مثال

$$(a_۱,a_۲,\dots,a_n)=(a_۱,a_۲)(a_۱,a_۳)\dots(a_۱,a_n)=$$

$$=(a_۲,a_۳)(a_۲,a_۴)\dots(a_۲,a_n)(a_۲,a_۱) \qquad (۳)$$

ثابت می‌کنیم، هر دور دوتایی، جای‌گشتی فرد است. پیش از این، این گزاره را دربارهٔ دور (۲ , ۱) دیده‌ایم. ولی هر دور $(i\,,\,j)$ عبارت است از $S(۲\,,\,۱)S^{-۱}$ که در آن، S هر جای‌گشتی مثل $\begin{pmatrix} ۱ & , ۲ & ,\dots \\ i & , j & ,\dots \end{pmatrix}$ که ۱ را به i و ۲ را به j می‌رساند، می‌تواند باشد. جای‌گشت $S(۲\,,\,۱)S^{-۱}$، جای‌گشتی فرد است، زیرا (۱,۲) فرد و S و $S^{-۱}$ یا هردو زوج و یا هر دو فردند.

با توجه به‌دستور (۳) ، دُور به طول $m+۱$ می‌تواند به‌صورت ضرب m جای‌گشت فرد

تجزیه شود. بنابراین، اگر $m+1$ زوج باشد، دُور به طول $m+1$ جای‌گشتی فرد است و اگر $m+1$ فرد باشد، دُور به طول $m+1$ زوج است. از این راه می‌توان زوج یا فرد بودن جای‌گشت‌هایی را که تجزیهٔ آن‌ها به دورها معلوم است، پیدا کرد. جای‌گشت $\begin{pmatrix} ۱ & ۲ & ۳ & ۴ & ۵ & ۶ \\ ۴ & ۵ & ۶ & ۳ & ۲ & ۱ \end{pmatrix}$ زوج است، زیرا بنابر دستور (۲) برابر است با حاصل‌ضرب دو جای‌گشت فرد.

زیرگروه‌ها. اگر بخشی از گروه، خود گروهی باشد، به‌آن **زیرگروهِ** گروه مفروض گویند. برای نمونه، گروه متناوب از جای‌گشت‌های متغیرهای x_1، x_2، ... ، x_n، زیرگروهی از گروه متقارن تبدیل‌های این متغیرهاست. مجموعهٔ حرکت‌های سرهٔ صفحه گروهی را تشکیل می‌دهند که زیرگروهی از گروه حرکت‌های سرهٔ و ناسرهٔ صفحه است.

از دیدگاه صوری، تبدیل همانی (یا تبدیل واحد) را هم می‌توان زیرگروه به حساب آورد. به همین ترتیب، هر گروه را می‌توان در ضمن، زیرگروهی از خودش دانست. اما به‌جز این زیرگروه‌های پیش‌پا افتاده به‌تقریب همیشه، گروه شامل تعداد زیادی زیرگروه است. شناخت همهٔ زیرگروه‌های یک گروه مفروض می‌تواند تصوری کامل دربارهٔ ساختار گروه مفروض به ما بدهد.

یکی از معمول‌ترین روش‌های تشکیل زیرگروه‌های یک گروه مفروض، معین کردن مولدهای زیرگروه است.

A_1، A_2، ... ، A_m را تبدیل‌هایی می‌گیریم که متعلق به‌گروه G باشند. مجموعهٔ H از همهٔ تبدیل‌هایی که می‌توان، از راه ضرب تعداد دل‌خواهی از تبدیل‌های مفروض و تبدیل‌های وارون آن‌ها به‌دست آورد، یک گروه است. در واقع، تبدیل همانی به‌این مجموعه تعلق دارد، زیرا آن را می‌توان به‌صورت $A_1 A_1^{-1}$ در نظر گرفت، سپس، اگر بتوان تبدیل‌های B و C را به‌صورت چنین حاصل‌ضربی در نظر گرفت، آن‌وقت، با ضرب این حاصل‌ضرب‌ها در یکدیگر، نمایش لازم را برای BC به‌دست می‌آوریم. سرانجام، اگر B را بتوان به‌صورت ضربی مثل $B=A_1^{-1}A_2A_1A_1A_2^{-1}$ نشان داد، آن‌وقت B^{-1} را هم می‌توان به‌صورت ضرب لازم نشان داد، زیرا $B^{-1}=A_2A_1^{-1}A_1^{-1}A_2^{-1}A_1$.

روشن است، گروه H زیرگروهی از گروه G است؛ H زیرگروهی است که به‌یاری تبدیل‌های A_1، .... ، A_m تولید شده است و خود تبدیل‌های A_1، .... ، A_m را مولدهای

(سازندگان) زیرگروه H می‌نامند. ممکن است H بر G منطبق باشد، در این صورت A_1، ....، A_m، خود گروه G را تولید کرده‌اند. به‌سادگی و به‌یاری مثال می‌توان قانع شد که، یک زیرگروه می‌تواند به‌کمک دستگاه‌های مختلفی از مولدها به‌دست آید.

زیرگروهی که به‌وسیلهٔ یک تبدیل A تولید می‌شود، زیرگروه دوری نامیده می‌شود. عضوهای این زیرگروه عبارت‌اند از

$$E\,,\ A\,,\ AA\,,\ AAA\,,\ \dots,\ A^{-1},\ A^{-1}A^{-1},\ A^{-1}A^{-1}A^{-1},\dots$$

و طبیعی است، آن‌ها را توان‌های تبدیل A بنامیم. در واقع

$$E=A^{0},\ A=A^{1},\ AA=A^{2},\dots,\ A^{-1}A^{-1}=A^{-2},\ A^{-1}A^{-1}A^{-1}=A^{-3},\dots$$

خیلی ساده ثابت می‌شود که همچون حساب معمولی داریم:

$$A^{m}A^{n}=A^{m+n} \quad \text{و} \quad (A^{m})^{n}=A^{mn} \tag{۴}$$

تبدیل را وقتی **متناوب** گویند که توان مثبتی از آن برابر تبدیل همانی باشد. کوچکترین نمای مثبتی که باید یک تبدیل متناوب را به آن توان رساند تا به‌تبدیل همانی رسید، مرتبهٔ این تبدیل نامیده می‌شود. طبق شرط، مرتبهٔ تبدیل نامتناوب را بی‌نهایت می‌گیرند.

چند مثال می‌آوریم. A را دوران صفحه دور نقطهٔ O و به‌اندازهٔ $\frac{360}{n}$ درجه می‌گیریم (n عددی طبیعی، مفروض و بزرگتر از واحد است). در این صورت A^{2} به معنای دوران به اندازهٔ $\frac{360}{n}\times 2$ درجه، A^{3} به معنای $\frac{360}{n}\times 3$ درجه، A^{n-1} به معنای دوران به‌اندازهٔ $\frac{360}{n}(n-1)$ درجه و A^{n} به‌معنای دوران با زاویهٔ ۳۶۰ درجه است (یعنی تبدیل همانی). بنابراین، دوران به اندازهٔ $\frac{360}{n}$ درجه، یک تبدیل متناوب از مرتبهٔ n است.

A را انتقال صفحه در طول یک خط راست می‌گیریم. در این صورت A^{2}، A^{3}، ... هم انتقال‌هایی در طول همان خط راست، به‌ترتیب به دو برابر، سه برابر،... هستند. بنابراین هیچ توان مثبتی از A، یک تبدیل همانی نیست و مرتبهٔ A بی‌نهایت است.

عضوهای گروه دوری که به‌وسیلهٔ A پدید می‌آیند، عبارت‌اند از

$$\dots,\ A^{-2},\ A^{-1},\ E,A,\ A^{2},\dots \tag{۵}$$

اگر A تبدیلی از مرتبهٔ بی‌نهایت باشد، آن‌وقت همهٔ تبدیل‌ها در دنبالهٔ (۵) باهم

متفاوت‌اند و با گروهی نامتناهی سرو کار داریم. در واقع، در غیر این صورت، باید برابری $A^k=A^l$ $(k<l)$ برقرار باشد که از آن جا به‌دست می‌آید $A^{l-k}=E$ $(l-k>0)$ که نامتناوب بودن تبدیل A را نقض می‌کند.

اکنون فرض می‌کنیم، A تبدیلی متناوب از مرتبهٔ m باشد. در این صورت

$$A^m=E\ ,\ A^{m+1}=A\ ,\ A^{m+2}=A^2\ ,\ \dots ,\ A^{m-1}=A^{-1}\ ,\ A^{m-2}=A^{-2}\ ,\ \dots$$

یعنی، دنبالهٔ (۵) شامل تبدیل‌های E، A، A^2، ...، A^{m-1} است که به‌تناوب تکرار می‌شوند؛ این تبدیل‌ها باهم فرق دارند، زیرا اگر در مثل داشته باشیم: $A^l=A^k$ $(0\le k<l<m)$، آن وقت به‌دست می‌آید $A^{l-k}=E$ $(0<l-k<m)$ که با انتخاب m در تناقض است. بنابراین، زیرگروه دوری که به‌یاری تبدیلی از مرتبهٔ m تولید می‌شود، درست شامل m تبدیل مختلف است.

گروهی که همهٔ عضوهای آن نسبت به‌هم جابه‌جایی‌پذیر (تعویض‌پذیر) باشند، **گروه جابه‌جایی‌پذیر** یا ***گروه آبلی*** نامیده می‌شود. این نام‌گذاری به‌دلیل کشف آبل ریاضی‌دان نروژی است که اهمیت این گونه گروه‌ها را در نظریهٔ معادله‌ها نشان داد.

دستور (۴) روشن می‌کند، توان‌های یک تبدیل همیشه نسبت به‌هم جابه‌جایی‌پذیرند:

$$A^mA^n=A^nA^m=A^{m+n}$$

بنابراین، زیرگروه‌های دوری، همیشه گروه‌هایی آبلی هستند.

در حساب عددها، در کنار ضرب، عمل تقسیم هم نقش مهمی دارد. در نظریهٔ گروه‌ها، به‌دلیل این که ضرب، از قانون جابه‌جایی پیروی نمی‌کند، با دو نوع تقسیم سروکار داریم: راست و چپ. در واقع، جواب معادلهٔ $Ax=B$ را، که در آن A و B تبدیل‌های مفروض و x تبدیل مجهول است، خارج قسمت راست، و جواب معادلهٔ $yA=B$ را خارج قسمت چپ از تقسیم B بر A گویند. اگر دو طرف برابری اول از سمت چپ در A^{-1} دو طرف برابری دوم را از سمت راست در A^{-1} ضرب کنیم، به‌دست می‌آید:

$$x=A^{-1}B\ ,\ y=BA^{-1}$$

بنابراین، خارج قسمت حاصل از تقسیم تبدیل B بر تبدیل A را می‌توان $A^{-1}B$ یا BA^{-1} دانست.

روی مثال‌های زیادی دیدیم که، در حالت کلی، $AB \neq BA$. «خارج قسمت» $(AB)(BA)^{-۱}$ یا $(BA)^{-۱}(AB)$ را می‌توان همچون «معیاری» برای جابه‌جایی‌ناپذیر بودن جای‌گشت‌های A و B در نظر گرفت. عبارت دوم، یعنی

$$(BA)^{-۱}(AB) = A^{-۱}B^{-۱}AB$$

را **جابه‌جاگر** (یا تعویض‌گر) A و B می‌نامند و به‌صورت $(A\ ,\ B)$ نشان می‌دهند. از دستور

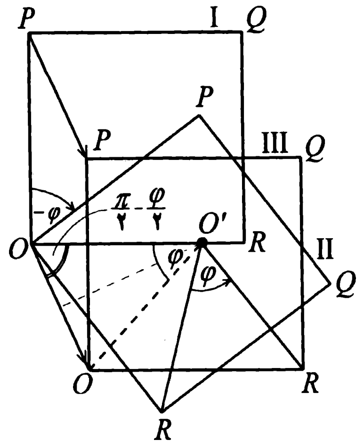
$$(A\ ,\ B) = A^{-۱}B^{-۱}AB$$

نتیجه می‌شود، خود جابه‌جاگر را می‌توان به‌عنوان خارج قسمت حاصل از «تقسیم» تبدیل مزدوج $B^{-۱}AB$ بر A به حساب آورد.

برای نمونه، اگر A انتقال صفحه باشد، آن‌وقت تبدیل مزدوج هم یک انتقال است و روشن است، خارج قسمت دو انتقال، یک انتقال است. بنابراین، جابه‌جاگر انتقال و هر حرکتی از صفحه، انتقال است. A را دورانی به اندازهٔ زاویهٔ φ دور نقطهٔ O و B را دوران یا انتقال می‌گیریم. در این صورت، تبدیل مزدوج، بازهم یک دوران به زاویهٔ φ است، ولی دور نقطهٔ جابه‌جا شدهٔ O'. بنابراین، جابه‌جاگر $(A\ ,\ B)$ در این حالت، عبارت است از حاصل‌ضرب دوران دور نقطهٔ O به زاویهٔ منفی φ و دوران دور نقطهٔ O' به زاویهٔ مثبت φ. روی شکل ۱۲ دیده می‌شود، تبدیل نهایی عبارت است از انتقال تحت زاویهٔ $\frac{\pi}{۲} - \frac{\varphi}{۲}$ نسبت به پاره‌خط راست $\overline{OO'}$ و به فاصلهٔ $۲ \times \overline{OO'} \sin \frac{\varphi}{۲}$.

به‌این ترتیب، به حقیقت جالبی می‌رسیم: برای صفحه، جابه‌جاگر هر دو حرکت نوع اول، یک انتقال موازی یا تبدیل همانی است. چون $(A,B) = E$ به‌معنای این است که $AB = BA$، بنابراین در صفحه، هر گروه جابه‌جایی‌ناپذیر حرکت‌های نوع اول، شامل بعضی انتقال‌های موازی است.

زیرگروهی که از همهٔ جابه‌جاگرهای عضوهای ممکنِ گروه G به‌وجود آید، زیرگروه جابه‌جاگر (یا تعویض‌گر یا مشتق) گروه G نامیده می‌شود. می‌توان گفت، زیرگروه جابه‌جاگر گروه G شامل آن، و تنها آن عضوهایی است که بتوان آن‌ها را به‌صورت ضرب جابه‌جاگرها نشان داد. از آن‌جا که برای صفحه، جابه‌جاگر هر دو حرکت نوع اول یک انتقال موازی است و از حاصل‌ضرب انتقال‌های موازی هم، انتقال‌های موازی به‌دست می‌آید، بنابراین می‌توان گفت که، زیرگروه جابه‌جاگر حرکت‌های صفحه از نوع اول، تنها از



شکل ۱۲

انتقال‌های موازی تشکیل شده است.

زیرگروه جابه‌جاگر گروه آبلی، شامل یک تبدیل است، زیرا از $AB=BA$ نتیجه می‌شود: $(A,B)=E$.

G را گروه متقارن همه جای‌گشت‌های عددهای ۱ ، ۲ ، ... ، n می‌گیریم. ثابت می‌کنیم، جابه‌جاگر هر دو جای‌گشت A ، B ، همیشه یک جای‌گشت زوج است. در واقع، جای‌گشت‌های AB ، BA و بنابراین $(BA)^{-۱}$ ، همیشه از لحاظ زوج یا فرد بودن یکسان‌اند؛ در این صورت جابه‌جاگر

$$(A,B)=(BA)^{-۱}(AB)$$

به‌عنوان حاصل‌ضرب دو جای‌گشتی که یا هر دو زوج و یا هر دو فردند، جای‌گشتی زوج می‌شود.

می‌بینیم زیرگروه جابه‌جاگر گروه متقارن، تنها از جای‌گشت‌های زوج تشکیل شده است. به‌سادگی می‌توان ثابت کرد که این زیرگروه بر تمامی گروه متناوب منطبق است.

زیرگروه جابه‌جاگر گروه G را، مشتق گروه هم می‌نامند و به G' نشان می‌دهند؛ زیرگروه جابه‌جاگر زیرگروه جابه‌جاگر G ، یعنی زیرگروه جابه‌جاگر دوم گروه G ، به G'' نشان داده می‌شود. با ادامهٔ این روند، می‌توانیم زیرگروه جابه‌جاگر هر مرتبه‌ای از گروه G را به‌دست آوریم.

اگر از زیرگروه‌های جابه‌جاگر گروه G ، دست‌کم یکی (و در این صورت، همهٔ

زیرگروه‌های جابه‌جاگر بعدی)، تنها شامل یک تبدیل باشد، گروه G را ***قابل حل*** (یا حل‌پذیر) می‌نامند. این نام‌گذاری از نظریهٔ معادله‌ها آمده است که در آن‌جا، قابل حل بودن گروه به معنای قابل حل بودن معادله به یاری رادیکال‌ها است. گروه حرکت‌های نوعِ اول برای صفحه قابل حل است، زیرا زیرگروه جابه‌جاگر دوم آن تنها شامل تبدیل همانی است. گروه‌های متقارن درجهٔ دوم، سوم و چهارم هم قابل حل‌اند، زیرا به ترتیب زیرگروه‌های جابه‌جاگر اول، دوم و سوم آن‌ها تنها شامل تبدیل همانی است. برعکس گروه‌های متقارن درجهٔ پنجم و یا از درجهٔ بالاتر، غیرقابل حل‌اند (حل‌ناپذیرند)، زیرا می‌توان ثابت کرد، زیرگروه جابه‌جاگر دوم آن‌ها، منطبق بر زیرگروه جابه‌جاگر اول و تنها شامل تبدیل همانی نیست.

۴. گروه‌های فِه‌دوروف

گروه‌های تقارن در شکل‌های محدود روی صفحه. پیش از این هم گفته‌ایم، تقارن شکل‌ها یا جسم‌ها به‌وسیلهٔ گروهی از حرکت‌های صفحه یا فضا مشخص می‌شود که شکل یا جسم را بر خودش منطبق کند.

ساده‌ترین این گروه‌ها را می‌توان در تقارن‌های شکل‌های محدود واقع بر صفحه پیدا کرد[1]. فرض کنید شکل محدودی در روی صفحه داده شده باشد و فرض کنید، این شکل را بتوان با حرکت A بر خودش قرار داد. در این صورت، گرانیگاه (= مرکز ثقل) O از شکل، باید ضمن حرکت A برخودش قرار گیرد، یعنی حرکت A، یا دوران دور نقطهٔ O است و یا تقارن آئینه‌ای (انعکاس) نسبت به‌خط راستی که از O می‌گذرد. بنابراین، گروه تقارن‌های هر شکل محدود واقع بر صفحه، تنها می‌تواند از دوران‌های دور گرانیگاه آن و انعکاس نسبت به خط‌های راستی که از گرانیگاه گذشته‌اند، تشکیل شده باشد.

حالت‌های مختلفی را بررسی می‌کنیم که ضمن مطالعهٔ گروه تقارن‌های شکل محدود واقع بر صفحه ممکن است پیش آید.

---

۱. محدود بودن به‌این مفهوم است که شکل در بخش محدودی از صفحه، و از جمله در درون یک دایره، واقع باشد.

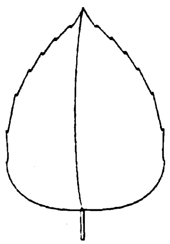
۱. گروه تقارن‌های $K_۱$ که تنها از یک تبدیل (تبدیل همانی) تشکیل شده است. این، گروه تقارن‌های هر شکل نامتقارن است (شکل ۱۳).

۲. گروه تقارن‌های $K_۲$ که از تبدیل همانی و انعکاس نسبت به‌یک خط راست تشکیل شده است (شکل ۱۴).

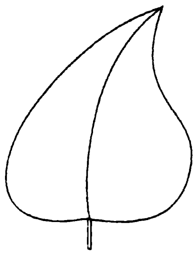
یادآوری می‌کنیم، اگر گروه K شامل انعکاس نسبت به دو خط راست (که از O گذشته‌اند) باشد و این دو خط راست، زاویهٔ φ را باهم بسازند، آن وقت حاصل‌ضرب این دو انعکاس، دورانی است به مرکز O و به اندازهٔ زاویهٔ ۲φ (شکل ۱۵). بنابراین روشن می‌شود که، گروه $K_۲$، تنها گروهی از گروه‌های تقارن‌هاست که دوران‌ها را دربر نمی‌گیرد.

۳. گروه تقارن‌های $K_۳$ که تنها شامل دوران‌هاست و در بین آن‌ها دوران‌های به‌قدر دل‌خواه کوچک وجود ندارد. در این حالت، بین دوران‌های گروه $K_۳$، دورانی با کوچکترین زاویهٔ مثبت وجود دارد. این زاویه را α درجه می‌گیریم. فرض کنیم، زاویه یک دوران دل‌خواه در $K_۳$ برابر β درجه باشد. نشان می‌دهیم که β مضربی از α است. عدد درست h را طوری پیدا می‌کنیم که داشته باشیم:

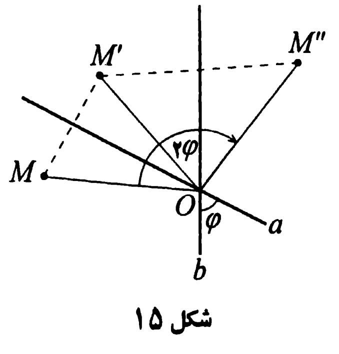
$$h\alpha^\circ \leq \beta^\circ < (h+۱)\alpha^\circ \Rightarrow ۰ \leq \beta^\circ - h\alpha^\circ < \alpha^\circ$$



شکل ۱۴



شکل ۱۳



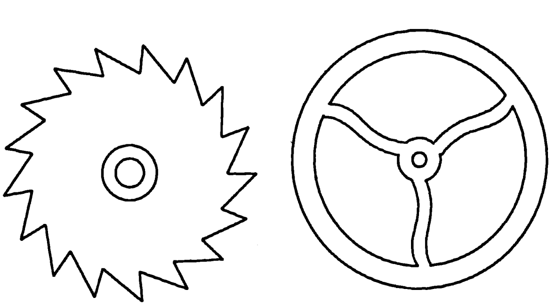
شکل ۱۵

گروه K_3 که دوران‌هایی به اندازهٔ α درجه و β درجه داشته باشد، دوران به زاویهٔ $(\beta-h\alpha)$ درجه هم خواهد داشت. اما $0 \leq \beta^\circ - h\alpha^\circ < \alpha^\circ$، و گروه شامل دوران‌هایی به زاویهٔ کمتری از α° نیست. پس $\beta^\circ - h\alpha^\circ = 0$ و در نتیجه $\beta^\circ = h\alpha^\circ$. درضمن، از آن جا که گروه K_3 شامل دوران ۳۶۰ درجه است، بنابراین برای عدد درستی مثل n باید داشته باشیم:

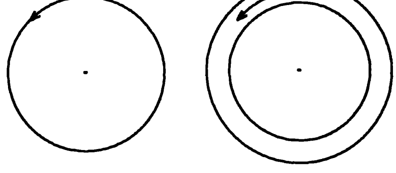
$$n\alpha^\circ = 360^\circ \Rightarrow \alpha^\circ = \frac{360^\circ}{n}$$

به‌این ترتیب، گروه K_3 شامل دوران‌هایی به اندازهٔ ۰ درجه، $\frac{360}{n}$ درجه، $2\frac{360}{n}$ درجه، ...، $(n-1)\frac{360}{n}$ درجه است. اگر n را برابر ۲ ، ۳ ، ۴ ، ... بگیریم، همهٔ گونه‌های گروه K_3 به‌دست می‌آید. نمونهٔ شکل‌هایی که گروه تقارن‌های آن‌ها تنها از دوران‌های دور نقطهٔ O یا زاویه‌هایی که مضربی از $\frac{360}{n}$ درجه هستند، به‌ازای $n = 3$ و $n = 19$ در شکل ۱۶ داده شده است.

۴. گروه تقارن‌های K_4 که تنها شامل دوران‌ها است و زاویهٔ دوران می‌تواند به‌دل‌خواه کوچک باشد. در این صورت دوران با هر زاویهٔ دل‌خواه α، را می‌توان با دقت دل‌خواه با دوران‌هایی که به‌گروه K_4 تعلق دارد، تقریب کرد. البته در این‌جا، تنها به شکل بسته توجه داریم، یعنی شکل‌هایی که نقطه‌های مرزی خود را هم شامل می‌شوند (۹ از بخش هفدهم را ببینید). به‌سادگی روشن می‌شود، برای شکل‌های بسته، گروه K_4 شامل دوران‌هایی با هر زاویهٔ دل‌خواه φ است. و این، حالتِ تقارن‌های دایره‌ای جهت‌دار است که نمونهٔ آن را می‌توان محیط دایره، نوار حلقه‌ای و غیره دانست (شکل ۱۷). در این حالت، برای همهٔ تبدیل‌ها، نه‌تنها باید شکل بر خودش منطبق شود، بلکه جهت حرکت را هم حفظ کند و روشن است که انعکاس نسبت به خط راست نمی‌تواند در بین آن‌ها باشد.



شکل ۱۶



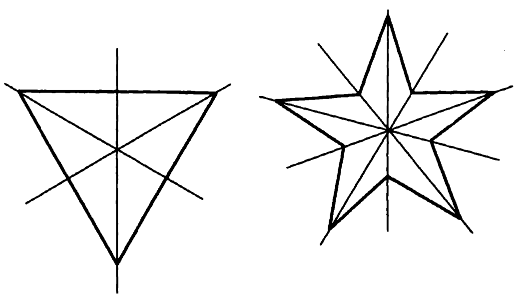
شکل ۱۷

تنها حالت باقی مانده یعنی حالتی را که، گروه تقارن‌های K، هم شامل دوران‌ها و هم شامل انعکاس‌هاست، بررسی می‌کنیم. از اثبات مطلب که بسیار ساده است، می‌گذریم و تنها به‌نتیجهٔ آن می‌پردازیم: به‌جز گروه‌های K_1 تا K_4، گروه‌های دیگری وجود دارند که تنها دو گونهٔ زیر را دربر می‌گیرند.

۵. گروه تقارن‌های K_5 شامل n انعکاس نسبت به‌خط‌های راستی که از O گذشته‌اند و صفحه را به $2n$ زاویهٔ برابر تقسیم می‌کنند، و دوران‌های با زاویه‌هایی که مضرب $\frac{360}{n}$ درجه هستند. نمونهٔ این گروه تقارن را می‌توان در nضلعی‌های منتظم دید (شکل ۱۸).

۶. گروه تقارن‌های K_6 شامل همهٔ دوران‌های دور نقطهٔ O و انعکاس نسبت به همهٔ خط‌های راست ممکنی که از O می‌گذرند. این، حالت تقارن دایره‌ای کامل است که نمونهٔ آن را می‌توان در تقارنِ بدون جهتِ محیط دایره یا حلقهٔ گرد بدون جهت دید.

گروه‌های تقارن برای شکل‌های نامحدود روی صفحه. پیدا کردن همهٔ گروه‌های ممکن تقارن‌ها برای شکل‌های نامحدود واقع بر صفحه مسأله‌ای پیچیده‌تر است. البته در عمل، اغلب ضرورتی به تمامی صفحهٔ نامحدود وجود ندارد. باوجود این، وقتی هم که بخشی از صفحه را درنظر می‌گیریم، اغلب پوشیده از چنان شکل‌های ریزی است که، خودِ این بخش صفحه، در

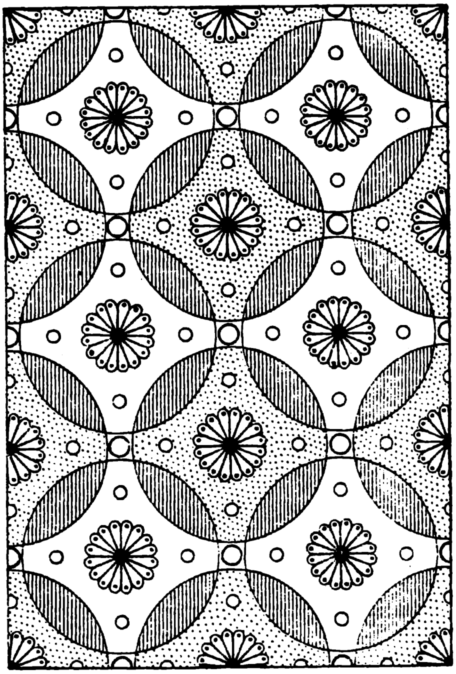


شکل ۱۸

مقایسهٔ با آن‌ها، بی‌نهایت بزرگ به‌حساب می‌آید. برای نمونه، بخشی از صفحه که به‌طور هموار هاشور خورده است، شامل نقش‌هایی با اندازه‌های میکروسکوپی است. چنین نقش‌هایی در ساختار همگن و درونیِ فلز هم وجود دارد.

همچنین می‌توان از نقش و نگارهای دیواری و یا بافته‌هایی از قالی یا پارچه نام برد که در آن‌ها شکل‌هایی به‌تکرار آمده است. شیوهٔ این نقش و نگارها بسیار گوناگون است و به صورت هنرمندانه‌ای در میان بسیاری از ملت‌ها از زمان‌های بسیار دور تا زمان ما وجود دارد و تکامل یافته است. در شکل ۱۹ ، نمونه‌ای از نقش و نگار مصری که در میانه‌های هزارهٔ سوم پیش از میلاد بر سقف نقش بسته است، دیده می‌شود.

برای گروه تقارن‌های شکل‌های محدود، به‌طور جداگانه حالت‌های ۱ ، ۲ ، ۳ و ۵ را بررسی کردیم، وقتی که گروه تقارن‌ها شامل دوران‌هایی با زاویهٔ دل‌خواه کوچک نیست؛ همچنین به‌حالت‌های ۴ و ۶ پرداختیم که چنین دوران‌هایی در گروه وجود دارد. برای بررسی گروه‌های تقارن برای شکل‌های نامحدود، به‌ویژه در حالت فضایی، این تقسیم‌بندی به‌گروه‌های گسسته و گروه‌های همراه با تبدیل‌های به‌اندازهٔ کافی کوچک، اهمیت زیادی



شکل ۱۹

دارد. به‌همین دلیل، در آغاز تلاش می‌کنیم، بین این دو حالت، مرزبندی لازم را پیدا کنیم.

گروه حرکت‌های صفحه را گسسته گویند، وقتی که هر نقطهٔ صفحه را بتوان در چنان دایره‌ای قرار داد که، هر حرکت گروه، یا نقطهٔ مفروض را بی‌حرکت نگاه دارد و یا آن را یکباره از آن دایره بیرون ببرد.

در این‌جا هم می‌توان همهٔ گروه‌های گسستهٔ حرکت‌های صفحه را پیدا کرد. همهٔ این گروه‌ها، گروه‌های تقارن‌های شکل‌های روی صفحه‌اند. طبیعی است، باید سه نوع گروه گسستهٔ تقارن‌ها را از هم جدا کرد:

I. روی صفحه، نقطه‌ای وجود دارد که، به‌ازای همهٔ تبدیل‌های تقارنی، بی‌حرکت می‌ماند. این نوع شامل گروه‌های K_1، K_2، K_3 و K_5 که پیش از این نام برده‌ایم، می‌شود.

II. روی صفحه، نقطهٔ بی‌حرکت وجود ندارد، ولی خط راستی وجود دارد که به‌ازای همهٔ تبدیل‌های گروه، برخودش قرار می‌گیرد. این خط راست را محور گروه می‌نامند. گروه‌های تقارن این نوع مربوط به آرایش و نقش و نگاری هستند که به‌صورت نواری (حاشیه‌ای) بی‌انتها کشیده شده است. از این گونه گروه‌ها، روی هم هفت تا وجود دارد:

۱. گروه تقارن‌های L_1 که تنها از انتقال‌هایی به فاصلهٔ مضرب‌هایی از یک پاره‌خط راست a تشکیل شده است.

۲. گروه L_2 که از گروه L_1 با اضافه کردن دورانی به‌اندازهٔ ۱۸۰ درجه نسبت به‌یکی از نقطه‌های محور گروه به‌دست می‌آید.

۳. گروه L_3 که از گروه L_1 با اضافه کردن انعکاس نسبت به‌خط راستی عمود بر محور گروه به‌دست می‌آید.

۴. گروه L_4 که از گروه L_1 با اضافه کردن انعکاس نسبت به محور به‌دست می‌آید.

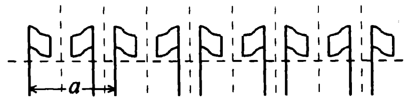
۵. گروه L_5 که از گروه L_1 با اضافه کردن انتقال به اندازهٔ پاره‌خط راست $\frac{a}{2}$ و انعکاس نسبت به محور به‌دست می‌آید.

۶. گروه L_6 که از گروه L_4 با اضافه کردن انعکاس نسبت به خط راستی عمود بر محور گروه به‌دست می‌آید.

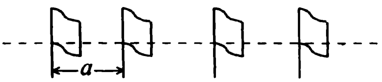
۷. گروه L_7 که از گروه L_5 با اضافه کردن انعکاس نسبت به خط راستی عمود بر محور گروه به‌دست می‌آید.

درجدول ۲ ،طرحی از «حاشیه‌ها»، متناظر با هریک از گروه‌های L_1 تا L_7 داده شده است.

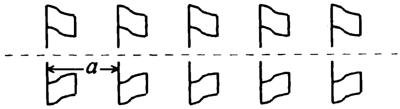
جدول ۲



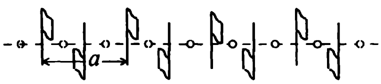
L_3



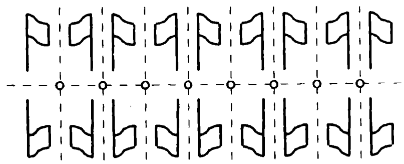
L_1



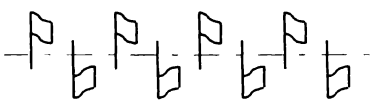
L_4



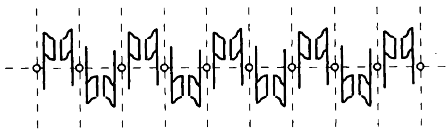
L_2



L_6



L_5



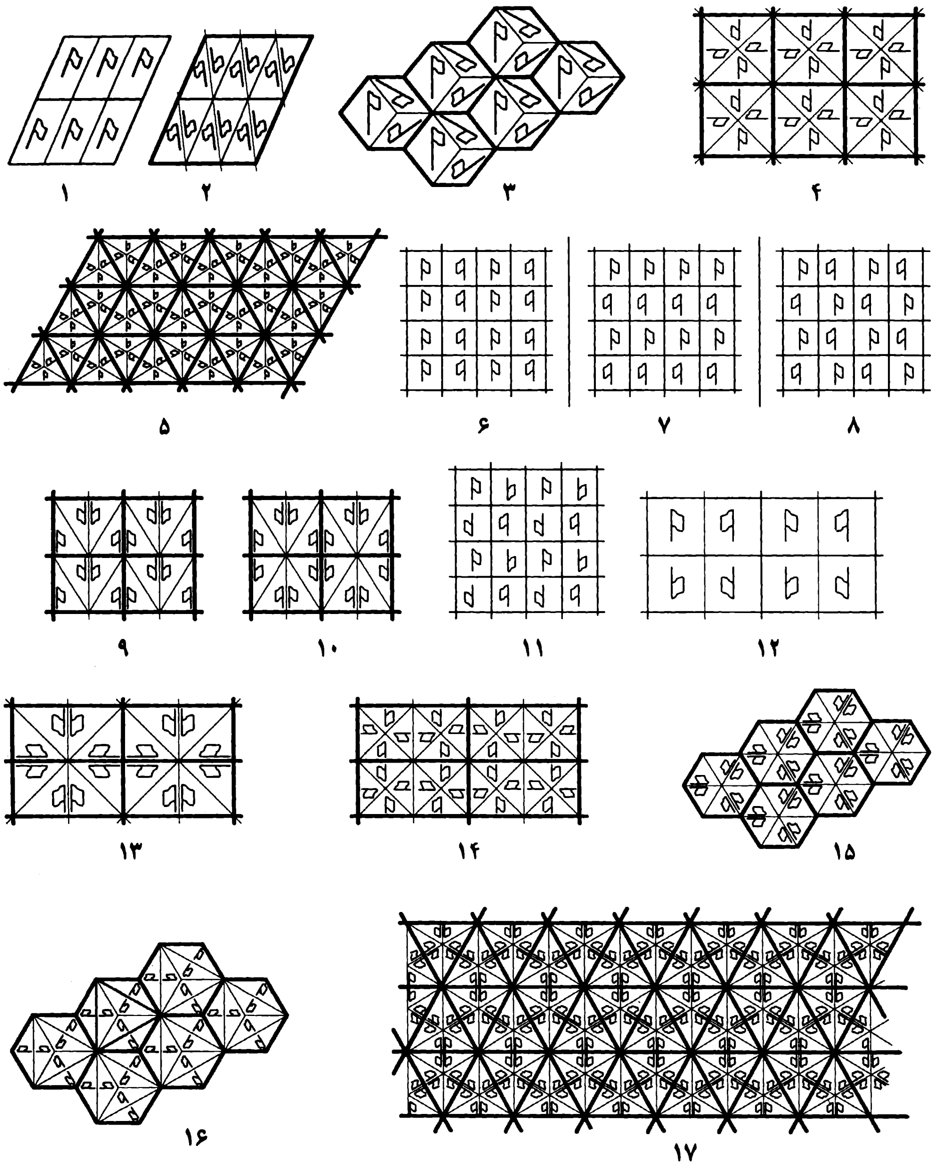
L_7

III. روی صفحه، نقطه یا خط راستی وجود ندارد که به‌ازای همهٔ تبدیل‌های گروه بر خودش قرار گیرد. گروه‌های این نوع را، ***گروه‌های فِه‌دُورُوف در روی صفحه*** گویند. این‌ها گروه‌های تقارن نقش و نگارهای بی‌پایان در روی صفحه‌اند. از این نوع گروه‌ها، روی هم هفده تا وجود دارد که پنج تا از آن‌ها تنها از حرکت‌های نوع اول، و دوازده تا از حرکت‌های نوع اول و نوع دوم تشکیل شده‌اند.

در جدول ۳ ، طرحی از نقش و نگارهای متناظر با هریک از ۱۷ حالت گروه‌های فِه‌دُورُوف در روی صفحه داده شده است؛ درضمن، هر گروه تنها از همان حرکت‌هایی تشکیل شده است که هر پرچم در روی شکل منجر به هر پرچم دیگری از همان شکل می‌شود.

جالب است یادآوری کنیم، هنرمندان نقش‌های آرایشی، در عمل همهٔ گروه‌های ممکن تقارن را کشف کرده‌اند. برای نظریه گروه‌ها، تنها این می‌ماند که ثابت شود نوع دیگری وجود ندارد.

جدول ۳



گروه‌های مربوط به‌بلورشناسی (کریستالوگرافیک). در سال ۱۸۹۰ ، یِ.س.فِه‌دُوُروف، هندسه‌دان و بلورشناس روس، توانست با روش‌های نظری ـ گروهی، یکی از مسأله‌های اساسی بلورشناسی را حل کند: مسألهٔ رده‌بندی دستگاه فضایی منتظم نقطه‌ها. این، نخستین حالت از کاربرد مستقیم نظریهٔ گروه‌ها برای حل مسأله‌های

مربوط به دانش‌های طبیعی بود که در ضمن، تأثیری جدی بر پیشرفت نظریهٔ گروه‌ها داشت.

جسم‌های بلوری شده (کریستالیزه) دارای این ویژگی هستند که اتم‌های تشکیل دهندهٔ آن‌ها، به‌مفهومی، یک دستگاه منتظم را در فضا می‌سازند. حرکت‌هایی از فضا را در نظر می‌گیریم که نقطه‌های دستگاه را دوباره به‌نقطه‌های دستگاه منجر کنند. این حرکت‌ها، گروهی را تشکیل می‌دهند که با بررسی ویژگی‌های آن می‌توان، خودِ مفهوم دستگاه منتظم را، دقیق‌تر تنظیم کرد.

دستگاهی از نقطه‌های فضا را ***دستگاه منتظم از نقطه‌های فضا*** گوییم، وقتی که

۱) هر نقطهٔ دستگاه بتواند به وسیلهٔ حرکتی که دستگاه را روی خودش قرار می‌دهد، به هر نقطهٔ دیگر دستگاه منجر شود؛

۲) هیچ کره‌ای با شعاع محدود، شامل بی‌نهایت نقطه از دستگاه نباشد؛

۳) عدد مثبتی مثل r وجود داشته باشد که هر کرهٔ به شعاع r، دست‌کم شامل یکی از نقطه‌های دستگاه باشد.

مسألهٔ مطالعه ساختمان جسم‌های بلوری شده، با نظم بخشیدن به دستگاه منتظم فضایی از نقطه‌ها، بستگی تنگاتنگی دارد که، به‌نوبهٔ خود، به تنظیم گروه حرکت‌های گسستهٔ فضا بستگی دارد. شبیه حالت صفحه، گروه حرکت‌های H از فضا را گسسته گوییم، وقتی که بتوان نزدیک هر نقطهٔ A از فضا کره‌ای با شعاع مثبت r و به مرکز A چنان در نظر گرفت که هر حرکت موجود در H، یا نقطهٔ A را در جای خود نگه دارد و یا آن را از مرزهای کره خارج کند.

می‌توان ثابت کرد، مجموعهٔ حرکت‌هایی از فضا که دستگاه منتظم فضایی نقطه‌ها را بر خودش قرار می‌دهند، بی‌تردید یک گروه گسسته است؛ در ضمن همهٔ نقطه‌های دستگاه را می‌توان از هر نقطهٔ دستگاه، به‌یاری تبدیل‌های این گروه به‌دست آورد. برعکس، اگر H گروه گسسته‌ای باشد، آن وقت با انتخاب نقطهٔ دل‌خواه A از فضا و حرکت دادن آن به‌کمک همهٔ حرکت‌های ممکنی که در H وجود دارد، به دستگاهی از نقطه‌ها می‌رسیم که دارای ویژگی‌های ۱ و ۲ هستند. از راه شرط‌های اضافی نه‌چندان دشوار، می‌توان از گروه‌های گسسته، گروه‌هایی را جدا کرد که با انتخاب درست نقطه A، دستگاه فضایی منتظم نقطه‌هایی را، که دارای هر سه ویژگی ۱ و ۲ و ۳ باشد، تشکیل دهد. این‌گونه گروه‌های گسسته را، ***گروه‌های فه‌دُروف*** یا ***گروه‌های مربوط به***

بلورشناسی می‌گویند. از آن چه گفتیم، روشن می‌شود که نخستین گام در ضمن اساسی‌ترین گام، در بررسی دستگاه‌های فضایی منتظم از نقطه‌ها، پیدا کردن گروه‌های **فه‌دوروف** است. به‌منظور نیازهای دانش‌های طبیعی، باید نه‌تنها به بررسی گروه‌هایی پرداخت که تنها شامل حرکت‌های سره‌اند، بلکه گروه‌هایی را هم باید بررسی کرد که هم حرکت سره و هم حرکت‌های ناسره را دربر می‌گیرند (یعنی حرکت‌هایی که شامل انعکاس هم هستند). تعداد گروه‌های **فه‌دوروف** که تنها از حرکت سره تشکیل شده‌اند، بسیار کمتر از تعداد گروه‌های **فه‌دوروف** است که هم حرکت‌های سره و هم حرکت‌های ناسره را شامل می‌شوند؛ و این حالت کلی‌تر برای به‌دست‌آوردن همهٔ دستگاه‌های منتظم فضایی نقطه‌ها، که در طبیعت ساختمان جسم‌های بلوری شده وجود دارد، لازم است.

یادآوری این مطلب لازم است که تنها به یاری نظریهٔ گروه‌ها بود که راهی برای پی‌بردن به‌غنای استثنایی حالت فضایی، که با حالت مسطحهٔ آن قابل مقایسه نیست، گشوده شد.

دشواری مسألهٔ فضایی را، در مقایسه با حالت روی صفحه، می‌توان با توجه به جدول زیر فهمید:

تعداد گروه‌های فِه‌دُوروف در حالت فضایی

گروه‌هایی که تنها شامل حرکت‌های نوع اول‌اند ...	۶۵
گروه‌هایی که حرکت‌های نوع دوم را هم شامل می‌شوند ...	۱۶۵
روی‌هم...	۲۳۰

بحث مفصل دربارهٔ گروه‌های **فه‌دوروف** در فضا و یاد کردن از همهٔ گونه‌های آن، ده‌ها صفحه از کتاب را دربر می‌گیرد. به‌همین دلیل، در این‌جا به‌همین اندازه اکتفا می‌کنیم و خوانندهٔ علاقه‌مند را به‌کتاب اختصاصی مراجعه می‌دهیم.

پیشرفت دانش بلورشناسی، ضرورت ادامهٔ کار را دربارهٔ تکامل مفهوم تقارن جلو ریاضی‌دان گذاشته است. امکان‌ها و راه‌های تازهٔ این بحث، تا اندازه‌ای در کتاب «تقارن‌ها و نامتقارن‌ها در شکل‌های محدود» **نوشتهٔ *آ.و.* شوب نیکوف** بلورشناس نشان داده شده است.

۵. گروه‌های گالوا

نتیجه‌هایی که در صفحه‌های پیشین دربارهٔ آن‌ها صحبت کردیم، تا اندازه‌ای، تصوری دربارهٔ نقش نظریهٔ گروه‌ها در شکل دادن به‌مسأله‌های مربوط به فرم‌های بلوری شده به ما می‌دهد، ولی بلورشناسی تنها دلیل پیدایی نظریهٔ گروه‌ها نبود. در نیمهٔ اول سدهٔ نوزدهم، ***لاگرانژ*** ریاضی‌دان فرانسوی متوجه بستگی بین ویژگی‌های تقارنی ریشه‌های معادلهٔ جبری و امکان حل معادله‌ها به یاری رادیکال‌ها شد. در سه دههٔ نخست سدهٔ نوزدهم در کارهای *آبل* و ***گالوا***، این بستگی با عمق بیشتری بررسی و منجر به حل مسألهٔ مهم و «بغرنج» مربوط به شرط‌های لازم برای قابل حل بودن معادلهٔ جبری به یاری رادیکال‌ها شد. این راه‌حل، به‌طور کامل بر بررسی ظریفی از ویژگی‌های گروه‌های جای‌گشتی متکی بود و در واقع باید آن را، آغاز پدید آمدن نظریهٔ گروه‌ها دانست.

مطالعهٔ بستگی بین ویژگی‌های معادله‌های جبری و ویژگی‌های گروه‌ها، موضوع نظریهٔ گسترده‌ای است که نام نظریهٔ ***گالوا*** را برخود دارد.

مفهوم و اهمیت تاریخی نظریهٔ گالوا در بخش چهارم (جلد اول) شرح داده شده است. ولی از آن‌جا که نظریهٔ ***گالوا*** نقشی قطعی و سرنوشت‌ساز در تکامل نظریهٔ گروه‌ها داشته است، اساسی‌ترین موضوع‌های آن را دوباره در این جا مطرح می‌کنیم که در واقع برای تسلط بیشتر بر خود نظریهٔ گروه‌ها است. اثبات دقیق گزاره‌های مربوط به نظریهٔ ***گالوا*** نیاز به مفهوم‌های کمکی زیادی دارد که در این جا از آن‌ها گذشته‌ایم.

گروه معادله‌های جبری. این معادلهٔ درجهٔ n را در نظر می‌گیریم:

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0 \qquad (۶)$$

ضریب‌های این معادله را مقدارهایی مفروض، مثل عددهایی مختلط به حساب می‌آوریم. مجموعهٔ مقدارهایی را که می‌توان از ضریب‌های معادله به‌کمک تعداد محدودی از عمل‌های جمع، تفریق، ضرب و تقسیم به‌دست آورد، ***هیات ضریب‌های*** (یا میدان ضریب‌های) معادله گویند.

ازجمله، اگر ضریب‌های معادله عددهای گویا باشند، هیات ضریب‌ها شامل همهٔ عددهای گویا است؛ اگر هم معادله به صورت $x^2 + \sqrt{2}x + 1 = 0$ باشد، هیات ضریب‌ها از

همهٔ عددهای به‌صورت $a+b\sqrt{۲}$ تشکیل شده است که در آن، a و b عددهایی گویا هستند.

اکنون ریشه‌های معادلهٔ مفروض را $\xi_۱$ ، $\xi_۲$ ، .... ، ξ_n می‌گیریم. مجموعهٔ مقدارهایی را که می‌توان به‌کمک تعداد محدودی از عمل‌های جمع، تفریق، ضرب و تقسیم به‌وسیلهٔ ریشه‌های $\xi_۱$ ، ... ، ξ_n به‌دست آورد، ***هیات شکافنده (یا تجزیه)*** معادله می‌نامیم. برای نمونه، هیات شکافنده معادلهٔ $x^۲+۱=۰$ عبارت است از مجموعهٔ عددهای مختلط به صورت $a+bi$ (با عددهای گویای a و b)، و هیات شکافندهٔ معادلهٔ $x^۲+\sqrt{۲}x+۱=۰$ عبارت است از مجموعه عددهای به‌صورت $a+bi+c\sqrt{۲}+di\sqrt{۲}$ که در آن a ، b ، c و d عددهایی گویا هستند.

ضریب‌های معادله، باتوجه به دستورهای ویت، به‌کمک عمل‌های جمع و ضرب، از ریشه‌های آن به‌دست می‌آیند، بنابراین هیات شکافندهٔ معادله همیشه شامل هیات ضریب‌های آن است. گاهی این میدان‌ها برهم منطبق‌اند.

نگاشت یک‌به‌یک A، هیات شکافنده به روی خودش، خودریختی یا خودسانی (*automorphism*) هیات شکافنده نسبت به هیات ضریب‌ها نامیده می‌شود، به‌شرطی که برای هر دو عضو هیات شکافنده، مجموع به‌مجموع، و حاصل‌ضرب به‌حاصل‌ضرب تبدیل؛ و هر عضو هیات ضریب‌ها تبدیل به‌خودش شود. این ویژگی‌ها را می‌توان به‌صورت دستورهایی نوشت:

$$(a+b)A=aA+bA\,,(ab)A=aA\,.\,bA\,,\alpha A=\alpha(a\,,b\in K,a\in P) \qquad (۷)$$

که در آن، aA عضوی است که عضو a ضمن نگاشت A به‌آن تبدیل می‌شود؛ P هیات ضریب‌ها؛ K هیات شکافنده است.

پیش از این (در بند ۲ ، زیر عنوان «تعریف کلی‌تری برای تقارن‌ها») گفتیم، مجموعهٔ همهٔ خودریختی‌های هیات شکافنده نسبت به هیات ضریب‌ها، یک گروه است، همین گروه را، ***گروه گالوا*** از معادلهٔ مفروض گویند.

برای این که تصور مشخص‌تری از گروه ***گالوا*** داشته باشیم، قبل از هر چیز یادآوری می‌کنیم که خودریختی‌های گروه ***گالوا***، ریشه‌های معادلهٔ مفروض را، دوباره به ریشه‌های همین معادله منجر می‌کند. در واقع، اگر x ریشهٔ معادلهٔ (۶) باشد، آن وقت با اِعمال خودریختی A در دو طرف معادله، و استفاده از ویژگی‌های (۷) به‌دست می‌آید:

$$(xA)^n+a_۱A\,(xA)^{n-۱}+\cdots+a_nA=۰\cdot A\,;$$

چون $A = 0$. ۰ و $a_iA=a_i$، پس

$$(xA)^n + a_1 (xA)^{n-1} + \cdots + a_n = 0$$

و این، همان نتیجه‌ای است که می‌خواستیم. بنابراین، هر خودریختی A، جای‌گشت معینی از ریشه‌های معادله است. از طرف دیگر، با آگاهی از این جای‌گشت، خودریختی را می‌شناسیم، زیرا همهٔ عضوهای هیات شکافنده، تنها به‌یاری عمل‌های خودریختی، از ریشه‌ها به‌دست می‌آیند. این مطلب روشن می‌کند که، به‌جای گروه خودریختی‌ها، می‌توان گروه جای‌گشت‌های ریشه‌های معادله را که متناظر با آن است، در نظر گرفت. از این‌جا، به‌ویژه این نتیجه به‌دست می‌آید که همهٔ گروه‌های ***گالوا***، متناهی‌اند.

پیدا کردن گروه ***گالوای*** یک معادله، مسأله‌ای ساده نیست و تنها در بعضی حالت‌ها، گروه ***گالوا***، به‌طور نسبی ساده است. برای نمونه، معادلهٔ (۶) را با ضریب‌های حرفی a_1، ...، a_n درنظر می‌گیریم. هیات ضریب‌های این معادله، شامل کسرهای گویایی از ضریب‌ها، یعنی کسرهایی است که صورت و مخرج آن‌ها، چندجمله‌ای‌هایی از a_1، ...، a_n باشد. هیات شکافنده، شامل کسرهای گویایی از ریشه‌های معادله، یعنی ξ_1، ....، ξ_n است که با ضریب‌های معادله به‌وسیلهٔ دستورهای **ویت** به‌هم مربوط‌اند:

$$\begin{aligned} -a_1 &= \xi_1 + \xi_2 + \cdots + \xi_n \\ a_2 &= \xi_1 \xi_2 + \xi_1 \xi_3 + \cdots + \xi_{n-1} \xi_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ (-1)^n a_n &= \xi_1 \xi_2 \cdots \xi_n \end{aligned} \tag{8}$$

چون معادله (۶) «کلی است»، می‌توانیم ریشه‌های آن را متغیرهای مستقلی به‌حساب آوریم. در این صورت، هر جای‌گشت این ریشه‌ها، خودریختی هیات شکافنده را به‌وجود می‌آورد. دستورهای (۸) نشان می‌دهند، هر خودریختی ضریب‌ها از این گونه، به خودش منجر می‌شود و به همراه آن‌ها، همهٔ کسرهای گویایی که از آن‌ها به‌دست می‌آیند، به خودشان منجر می‌شوند. به‌این ترتیب، گروه ***گالوا*** از معادلهٔ کلی درجه n، در واقع تقریباً گروه متقارن از همهٔ جای‌گشت‌های n حرف است.

می‌توانیم معادلهٔ با ضریب‌های عددی هم نشان دهیم که گروه ***گالوای*** آن گروه متقارن است. برای نمونه ثابت می‌شود که گروه ***گالوا*** برای معادلهٔ

$$1 - \frac{n}{1}x + \frac{n(n-1)}{1\times 2}\cdot\frac{1}{1\times 2}x^2 - \frac{n(n-1)(n-2)}{1\times 2\times 3}\cdot\frac{1}{1\times 2\times 3}x^3 + \cdots +$$

$$+(-1)^n \frac{1}{1\times 2\times\cdots\times n}x^n = 0 \qquad (9)$$

برای هر مقدار n، گروه متقارن از درجهٔ n است.

به‌طور کلی، روش‌هایی وجود دارد که بتوان معادله‌هایی را با هر گروه از پیش داده شده‌ای، به‌عنوان گروه **گالوا** ساخت، ولی با این شرط که ضریب‌ها را بتوان به‌دل‌خواه انتخاب کرد. اگر هم بخواهیم معادله‌ای بسازیم که ضریب‌های گویا داشته باشد، چنین ساختمانی تنها برای گونه‌های جداگانه‌ای از گروه‌ها شناخته شده است. بیشترین موفقیت را در این زمینه، ***ای.پ.شافاره‌ویچ*** ریاضی‌دان شوروی به‌دست آورد، زیرا روش‌هایی برای ساختن معادله‌های با ضریب‌های گویا پیدا کرد که گروه **گالوای** آن‌ها، هر گروه مفروض حل‌پذیر باشد.

حل‌پذیر بودن معادله به‌یاری رادیکال‌ها. همان‌طور که از تعریف گروه **گالوای** معادله برمی‌آید، این گروه، با تقارن درونی ریشه‌های معادله، مشخص می‌شود. تمامی مسأله امکان تبدیل حل یک معادلهٔ مفروض، به حل معادله‌های از درجهٔ پایین‌تر؛ وهم بسیاری از مسأله‌های دیگر را ممکن است بتوان منجر به‌مسألهٔ ساختن گروه **گالوا** کرد. ولی گروه **گالوا** از هر معادلهٔ درجهٔ n، گروهی از جای‌گشت‌های درجهٔ n است که همهٔ بستگی‌های آن را، دست‌کم به‌صورت نظری، می‌توان با آزمایش پیدا کرد.

بررسی گروه **گالوا**، روش ارزنده‌ای برای حل مسألهٔ مربوط به معادله‌های جبری از درجهٔ بالا است. ازجمله می‌توان ثابت کرد، معادله‌ای به‌یاری رادیکال‌ها حل‌پذیر است که گروه **گالوای** مربوط به آن حل‌پذیر باشد، و البته تنها در همین حالت (برای حل‌پذیر بودن گروه، بند ۳ را ببینید). پیش از این گفته‌ایم، گروه‌های متقارنِ درجهٔ ۲ ، ۳ و ۴ حل‌پذیراند و این، متناظر با قابل حل بودن معادله‌های درجهٔ ۲ ، ۳ و ۴ به یاری رادیکال‌ها است. گروه **گالوا** از معادله‌های «کلی» درجهٔ ۵ ، ۶ ، ... ، گروه‌های متقارنی از همین درجه‌ها هستند و این گروه‌ها حل‌پذیر نیستند. از این جا نتیجه می‌شود، معادله‌های جبری بالاتر از درجهٔ چهارم را نمی‌توان به یاری رادیکال‌ها حل کرد.

ازجملهٔ معادله‌هایی که با رادیکال‌ها حل نمی‌شوند، می‌توان از معادلهٔ (۹) ، برای $n>4$،

نام برد، زیرا گروه‌های *گالوای* مربوط به آن نیز، گروه‌های متقارن‌اند.

۶. مفهوم‌های اصلی در نظریهٔ کلی گروه‌ها

نظریهٔ گروه‌ها، در سدهٔ نوزدهم، بیشتر به‌عنوان نظریهٔ گروه تبدیل‌ها پیش رفت. ولی در جریان زمان، هر روز بیش از پیش روشن شد که اغلب نتیجه‌گیری‌ها، تنها مربوط به‌این است که تبدیل‌ها را می‌توان ضرب کرد و در ضمن، این عمل ویژگی‌های خاص خود دارد. از طرف دیگر موضوع‌هایی پیدا شد که هیچ ربطی به‌تبدیل‌ها نداشتند، ولی می‌شد دربارهٔ آن‌ها عملی انجام داد (که آن را، شبیه قبل، ضرب می‌نامیم) که دارای همان ویژگی‌های گروه‌های تبدیل‌اند و قضیه‌های اصلی گروه‌های تبدیل هم دربارهٔ آن‌ها صادق است. به‌این مناسبت از سال‌های پایانی سدهٔ نوزدهم، مفهوم گروه، نه‌تنها برای دستگاه تبدیل‌ها، بلکه برای هر دستگاهی از عضوهای دل‌خواه به‌کار رفت.

تعریف کلی گروه. این تعریف را برای گروه پذیرفته‌اند: فرض کنید که هر زوج مرتب a و b از عضوهای مجموعهٔ G، متناظر با عضوی معین مثل c از همان مجموعه باشد. در این صورت گویند، در مجموعهٔ G، عملی تعریف شده است. برای چنین عملی نام‌های ویژه‌ای گذاشته‌اند: جمع، ضرب، ترکیب (*composition*). عضوی از مجموعهٔ G که پاسخ‌گوی a و b ضمن این عمل است، در این حالت به‌ترتیب، مجموع، حاصل‌ضرب، نتیجهٔ ترکیب عضوهای a و b نامیده می‌شود و به‌ترتیب به‌صورت نمادین $a+b$، ab، $a * b$ معرفی می‌شود. اصطلاح‌های «جمع» و «ضرب» در چنان حالت‌هایی هم که عمل، هیچ ربطی به‌عمل‌های معمولی جمع و ضرب عددها ندارد، به‌کار می‌رود.

مجموعهٔ G همراه با عمل $*$ که برای آن تعریف شده است، نسبت به‌این عمل گروه نامیده می‌شود، به شرطی که با این اصل موضوع‌ها سازگار باشد:

۱. برای هر سه عضو x، y، z از G داشته باشیم:

$$x*(y*z)=(x*y)*z \qquad \text{(قانون شرکت‌پذیری)}$$

۲. بین عضوهای مجموعه G، عضوی مثل e وجود داشته باشد که برای هر عضو x از G

$$x*e=e*x=x$$

۳. برای هر عضو a از G، عضوی مثل a^{-1} در G وجود داشته باشد، به‌نحوی که

$$a * a^{-1} = a^{-1} * a = e$$

عضو e را، که در اصل موضوع ۲ داشتیم، ***عضو همانی*** (یا ***عضو خنثای***) گروه و عضو a^{-1} را، که در اصل موضوع ۳ از آن صحبت کردیم، ***وارون عضو*** a گویند. اگر به عمل درونی گروه، نام جمع یا ضرب داده باشیم، آن وقت عضو همانی را، به‌ترتیب، عضو صفر یا عضو واحد می‌نامیم. در این حالت‌ها، اصل موضوع‌های گروه، به این صورت نشان داده می‌شوند:

۱) $x(yz) = (xy)z$,	۱) $x + (y+z) = (x+y) + z$,
۲) $xe = ex = x$,	۲) $x + 0 = 0 + x = x$,
۳) $x x^{-1} = x^{-1} x = e$	۳) $x + (-x) = (-x) + x = 0$,

در بندهای پیشین، مثال‌های زیادی از گروه دیدیم. در این گروه‌ها، هر عضو یک تبدیل، و عمل درونِ گروه، ضرب تبدیل‌ها بود. مجموعهٔ عددهای حقیقی ۰ ، ۱± ، ۲± ، ... ، نسبت به‌عمل جمع یک گروه تشکیل می‌دهند، *زیرا* مجموع عددهای درست، عددی درست است؛ جمع عددهای درست دارای ویژگی شرکت‌پذیری است؛ عضو همانی، عدد درست ۰ است؛ و در بین این عددها، برای هر عدد a ، عدد وارون (در این جا قرینهٔ) $-a$ وجود دارد. مثال دیگر، مجموعهٔ عددهای حقیقی (به‌جز عدد صفر) است که نسبت به ضرب یک گروه تشکیل می‌دهد. در واقع، حاصل‌ضرب هر دو عدد حقیقی غیرصفر، یک عدد حقیقی مخالف صفر است؛ عمل ضرب عددهای حقیقی شرکت‌پذیر است؛ عضو همانی وجود دارد و برابر است با ۱ ؛ هر عدد غیر صفر و حقیقی وارونی دارد $a^{-1} = \frac{1}{a}$. از این مثال‌ها تا جایی که بخواهیم وجود دارد.

گرچه عمل درونی گروه را می‌توان با نام‌های مختلفی یاد کرد، ولی ما از این به بعد، کم‌وبیش همه‌جا، آن را ضرب می‌نامیم. مفهوم‌های زیرگروه، درجهٔ عضو گروه، گروه دُوری و مرتبهٔ عضو گروه را به همان معنایی می‌گیریم که برای گروه‌های تبدیل داشتیم و در این‌جا، تعریف آن‌ها را تکرار نمی‌کنیم (بند ۳ را ببینید). تنها یادآوری می‌کنیم، عضو a از گروه G را مزدوج عضو b می‌نامیم، وقتی که در G عضوی مثل x وجود داشته باشد، به‌نحوی که داشته باشیم: $b = x^{-1}ax$. چون $a = a^{-1}aa$ ، پس هر عضو گروه با خودش مزدوج است. سپس، از $b = x^{-1}ax$ نتیجه می‌شود: $xbx^{-1} = a$ یا $a = (x^{-1})^{-1}bx^{-1}$ ، یعنی اگر عضو a با b مزدوج

باشد، آن وقت b هم مزدوج با a است. سرانجام، اگر $b=x^{-1}ax$ و $c=y^{-1}by$، آن‌وقت

$$c=y^{-1}x^{-1}axy=(xy)^{-1}a(xy)$$

بنابراین، دو عضوی که هر کدام با عضو سومی مزدوج باشند، خود مزدوج یکدیگرند. این ویژگی نشان می‌دهد، همهٔ عضوهای گروه، به رده‌های مختلفی از عضوهای مزدوج با یکدیگر، تقسیم می‌شوند. ولی، اگر گروه جابه‌جایی‌پذیر باشد، یعنی برای هر x و y داشته باشیم $xy=yx$، آن‌وقت عضوهای مزدوج برهم منطبق می‌شوند و هر رده از عضوهای مزدوج تنها شامل یک عضو است.

یکریختی. (یکسانی یا یک‌دیسگی .*isomorphism*). دو مطلب را در مفهوم گروه می‌توان دید. برای این که یک گروه مفروض باشد، باید: ۱) مشخص کنیم چه موضوع‌هایی عضو آن هستند، و ۲) قانون ضرب عضوها را مشخص کنیم. به این مطلب و به‌بررسی ویژگی‌های گروه، می‌توان از دیدگاه‌های مختلفی توجه کرد. می‌توان بستگی بین ویژگی‌های جداگانهٔ عضوهای گروه و مجموعهٔ این ویژگی‌ها را در رابطه با عمل درونی گروه بررسی کرد. از این دیدگاه، اغلب برای بررسی گروه‌های مشخص جداگانه و از جمله، برای بررسی گروه حرکت‌های فضا یا صفحه استفاده می‌کنند. ولی می‌توان به بررسی ویژگی‌هایی از گروه پرداخت که به‌طور کامل از طریق ویژگی‌های مربوط به عمل درونی گروه سرچشمه گرفته‌اند. این دیدگاه، بیشتر دربارهٔ گروه‌های *انتزاعی* یا ***نظریهٔ کلی گروه‌ها***، کاربرد دارد. اگر بخواهیم روشن‌تر بیان کنیم، این دیدگاه براساس مفهوم یکریختی یا یک دیسگی (ایزومورفیسم) قرار دارد.

دو گروه را یکریخت (یا یک دیسه یا یکسان) گویند، وقتی بتوان عضوهای یکی از آن‌ها را در تناظر عضوهای گروه دوم چنان قرار داد که حاصل‌ضرب عضوهای دل‌خواه گروه اول، پاسخ‌گوی حاصل‌ضرب عضوهای متناظر آن‌ها در گروه دوم باشد. وجود تناظر یک‌به‌یک بین عضوهای دو گروه، وقتی ویژگی‌های یاد شده را داشته باشند، **یکریختی** نامیده می‌شود.

به‌سادگی دیده می‌شود، وقتی یکریختی دو گروه، یک عضو از گروه اول را متناظر با عضوی از گروه دوم قرار می‌دهد / آنگاه این دو عضو ویژگی‌های مشابهی نسبت به عمل درونی گروه دارند. به‌این ترتیب، در یکریختی، عضو همانی، عضوهای وارون یکدیگر، عضوهای مرتبهٔ n و زیرگروه‌ها، در یکی از گروه‌ها، با عضو همانی، عضوهای وارون

یکدیگر، عضوهای مرتبهٔ n و زیرگروه‌های گروه دیگر، متناظر می‌شوند. بنابراین، می‌توان گفت، نظریهٔ انتزاعی گروه‌ها، تنها ویژگی‌هایی از گروه‌ها را بررسی می‌کند که ضمن نگاشت‌های یکریخت، تغییر نمی‌کنند. برای نمونه، از دیدگاه نظریهٔ انتزاعی گروه‌ها، همه جای‌گشت‌های چهار عنصر و گروه حرکت‌های سره و ناسره که چهاروجهی منتظم را به‌خودش منجر می‌کند، ویژگی‌های مشابهی دارند، زیرا با یکدیگر یکریخت‌اند. در واقع، این حرکت‌ها، راس‌های چهاروجهی را دوباره به راس‌های آن تبدیل می‌کنند. تعداد این حرکت‌ها برابر است با ۲۴. اگر هر حرکت را جای‌گشتی از راس‌ها بدانیم، تناظری یک‌به‌یک بین عضوهای دو گروه به‌دست می‌آید که همان یکریختی مورد جست‌وجوی ماست.

نظریهٔ لگاریتم‌ها، نمونهٔ جالبی از یکریختی را به ما می‌دهد. اگر هر عدد مثبت حقیقی را متناظر با لگاریتم آن قرار دهیم، تناظری یک‌به‌یک، بین مجموعهٔ همهٔ عددهای حقیقی مثبت با مجموعهٔ همهٔ عددهای‌حقیقی مثبت ومنفی به‌دست‌می‌آید. رابطهٔ $\log(xy) = \log x + \log y$ نشان می‌دهد که نگاشت مفروض عبارت است از نگاشت گروه عددهای حقیقی مثبت نسبت به ضرب، به روی گروه همهٔ عددهای حقیقی نسبت به جمع. اهمیت و ارزش عملی این یکریختی، برای همگان روشن است.

به‌عنوان نمونه‌هایی از گروه‌های نایکریخت، می‌توان از گروه‌های متناهی‌ای یاد کرد که مرتبه‌های مختلف داشته باشند.

همان‌طور که گفتیم، گروه‌های انتزاعی با قانون ضرب عضوها، بدون توجه به‌طبیعت آن‌ها، تعریف می‌شوند. به‌نحوی که گروه‌های مفروض مشخص یکریخت را، همچون مُدل یک گروه انتزاعی مشخص می‌توان در نظر گرفت.

گروه‌های انتزاعی را با روش‌های گوناگونی می‌توان داد، که طبیعی‌ترین آن‌ها، دست‌کم برای گروه‌های مشخص، عبارت است از عرضهٔ مستقیم «جدول‌های ضرب».

برای گروه‌های مرتبهٔ n، که عضوهای آن به ردیفی نوشته شده‌اند، این جدول‌ضرب به‌صورت مربعی شامل n سطر و n ستون درمی‌آید. در خانهٔ iامین سطر و jامین ستون این جدول، عضوی نوشته شده است که نمایندهٔ حاصل‌ضرب عضو با شمارهٔ i در عضو با شمارهٔ j است. چنین جدول‌ضربی برای گروه‌های متناهی، گاهی *مربع کیلی* (Cayley) نامیده می‌شود.

باوجود این، دادن گروه‌ها به‌وسیلهٔ جدول‌ضرب، در عمل و به دلیل حجم بزرگ جدول، اغلب قابل استفاده نیست.

راه‌های دیگری هم برای دادن گروه‌های انتزاعی وجود دارد. با یکی از این روش‌ها، یعنی دادن گروه به‌وسیلهٔ عضوهای مولد و تعیین رابطه‌های معرف، آشنا هستیم. ولی اغلب گروه جبری را، به‌وسیلهٔ گروه مشخص همریخت آن، و به‌ویژه گروه‌های تبدیل، مشخص می‌کنند.

به‌طور طبیعی در برابر این پرسش قرار می‌گیریم: آیا هر گروه جبری را می‌توان به‌عنوان یک گروه تبدیل بررسی کرد؟ پاسخ به این پرسش را، این قضیه می‌دهد: هر گروه G، با یک گروه تبدیل از مجموعهٔ عضوهای آن، یکریخت است.

در واقع، g را عضو ثابتی از G می‌گیریم و A_g را تبدیلی از مجموعهٔ عضوهای G فرض می‌کنیم که، به‌ازای آن، عضو x از G، به عضو xg تبدیل می‌شود. تبدیل A_g تناظر یک‌به‌یک است، زیرا معادلهٔ

$$xA_g = xg = a$$

برای هر عضو مفروض a، دارای جواب منحصر $x = ag^{-1}$ است. از طرف دیگر، حاصل‌ضرب gh از عضوهای گروه، متناظر تبدیل $A_g A_h$ است، زیرا

$$xA_{gh} = x(gh) = (xg)h = (xA_g)A_h = x(A_g A_h)$$

عضو همانی e از گروه G، متناظر تبدیل همانی و عضو وارون g^{-1} متناظر تبدیل وارون است. بنابراین، مجموعهٔ Γ همهٔ تبدیل‌هایی، که متناظر عضوهای گروه G هستند، گروه تبدیلی را تشکیل می‌دهند که با G یکریخت است. به‌سادگی می‌توان ثابت کرد، اگر تعداد عضوهای G بیشتر از ۲ باشد، آن وقت گروه Γ شامل همهٔ تبدیل‌های مجموعهٔ G نمی‌شود و تنها زیرگروهی از گروه «متقارن» همهٔ تبدیل‌های این مجموعه است.

زیرگروه‌های نرمال (یا ناوردا) و گروه بَهری (یاگروه خارج‌قسمت). P و Q را مجموعه‌های دل‌خواهی از عضوهای گروه G می‌گیریم. حاصل‌ضرب مجموعهٔ P در مجموعهٔ Q، که به‌صورت نمادین PQ نشان می‌دهیم، به مجموعه‌ای از عضوهای گروه G گفته می‌شود که بتوان آن‌ها را به‌صورت حاصل‌ضرب عضوی از P در عضوی از Q تصور کرد. در حالت خاص، حاصل‌ضرب gP که در آن g عضوی از گروه G است ـ عبارت است از مجموعهٔ حاصل‌ضرب‌های عضو g در هریک از عضوهای مجموعهٔ P.

زیرگروه H از گروه G را ***زیرگروه نرمال*** یا ***زیرگروه ناوردا*** از گروه G گویند، وقتی که برای

هر g از G داشته باشیم: $gH=Hg$. مجموعه‌های به‌صورت gH و Hg را که در آن‌ها H زیرگروهی دل‌خواه است ـ به‌ترتیب همردهٔ **راست** و همردهٔ **چپ** گروه G نسبت به زیرگروه H شامل عضو g گویند (همرده = هم‌مجموعه). به‌این ترتیب، می‌توان گفت، زیرگروه‌های نرمال، آن زیرگروه‌هایی هستند که همرده‌های راست و چپ آن‌ها که شامل یک عضو مشخص هستند، برهم منطبق‌اند.

اگر H، زیرگروهی نرمال باشد، آن وقت می‌توان ثابت کرد، حاصل‌ضرب دو همرده، بازهم یک همرده است، یعنی

$$aH.bH=ab.H$$

زیرگروه H، به‌خودی خود، همرده عضو همانی یا هر عضو h خود است، زیرا $hH=H$. حاصل‌ضرب همرده‌ها شرکت‌پذیر است:

$$(aH.bH)cH=(ab.c)H=(a.bc)H=aH(bH.cH)$$

زیرگروه H، در این ضرب، نقش عضو همانی را به‌عهده دارد:

$$H.aH=eH.aH=(ea)H=aH$$

و شبیه آن $aH.H=aH$. همردهٔ $a^{-1}H$، وارونِ aH است، زیرا

$$aH.a^{-1}H=aa^{-1}H=H$$

بنابراین، اگر هر همرده نسبت یک به زیرگروه نرمال به عنوان عضوی از یک مجموعهٔ تازه در نظر بگیریم، این مجموعه نسبت به‌عمل ضرب همرده‌ها، یک گروه است. این گروه را **گروه بهری** (گروه خارج‌قسمت = *factor group*) گروه G نسبت به‌زیرگروه نرمال H گویند و با نماد G/H نشان می‌دهند.

به‌سادگی روشن می‌شود، برای گروه‌های متناهی، هر همرده نسبت به هر زیرگروه H، به همان تعداد عضو مختلف دارد که در خود زیرگروه H موجود است؛ درضمن همرده‌های مختلف، دارای عضو مشترک نیستند. از این جا نتیجه می‌شود، تعداد همرده‌های گروه متناهی G، برای هر زیرگروه H از آن، برابر است با خارج قسمت مرتبهٔ G بر مرتبهٔ H که از آن‌جا قضیهٔ مهم **لاگرانژ** نتیجه می‌شود: مرتبهٔ هر زیرگروه یک گروه متناهی، بخشیابی از مرتبهٔ گروه است.

از تعریف زیرگروه نرمال دیده می‌شود، در گروه‌های آبلی، هر زیرگروه، یک زیرگروه نرمال است. حالت مرزی دیگر را در به اصطلاح گروه‌های ***ساده*** می‌توان دید که هر زیرگروهی از آن، به‌جز گروه یکانی و خودِ گروه، زیرگروه نرمال نیست. در کنار گروه‌های آبلی و گروه‌های ساده، گروه‌های حل‌پذیر ـ که تعریف آن‌ها را در بند ۳ دادیم، اهمیت زیادی دارند. می‌توان ثابت کرد، گروه‌های حل‌پذیر، زنجیرهٔ محدودی از زیرگروه‌های نرمال دارند: G ، G_1 ، G_2 ، ... ، G_k ، به نحوی که نخستین این زیرگروه‌ها، خودِ گروه G است؛ و درضمن همهٔ گروه‌های بهری G/G_1 ، G_1/G_2 ، ... ، G_{k-1}/G_k ، گروه‌های آبلی‌اند.

همریختی. (=همسانی = هم‌دیسگی *homomorphism*). مفهوم گروه بهری، با مفهوم نگاشت همریخت ـ که برای تمامی نظریهٔ گروه‌ها مفهومی اساسی است ـ بستگی تنگاتنگ دارد.

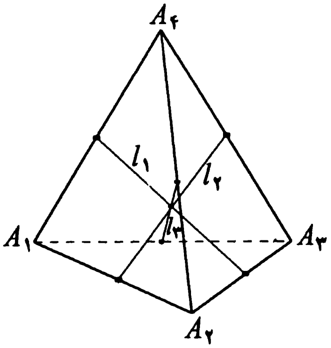
نگاشت یک‌ارزشی مجموعهٔ عضوهای گروه G را به روی مجموعهٔ عضوهای گروه H ، **همریختی** یا **نگاشت همریخت** گویند، به‌شرطی که حاصل‌ضرب هر دو عضو گروه اول، متناظر حاصل‌ضرب عضوهای متناظر در گروه دوم باشد.

به‌این ترتیب، اگر برای هر عضو x از گروه G ، عضو متناظر آن را در گروه H با x' نشان دهیم، نگاشت همریخت به‌این معنا است که

$$(x_1x_2)' = x_1' x_2'$$

از مقایسهٔ تعریف همریختی با تعریف یکریختی دیده می‌شود، نگاشت یکریخت تناظر یک‌به‌یک است، در حالی که نگاشت همریخت، تنها در یک‌سو یک‌ارزشی است: هر عضو گروه G متناظر تنها یک عضو گروه H است، ولی ممکن است عضوهای مختلفی از گروه G ، با یکی از عضوهای گروه H متناظر باشند. به‌مفهوم معینی، می‌توان گفت، در نگاشت یکریخت، گروه H ، نمونهٔ دقیقی از گروه G است، درحالی که در نگاشت همریخت، ضمن عبور از G به H ، اختلاف برخی از عضوهای G ازبین می‌رود، چراکه همهٔ آن‌ها، تنها به‌یک عضو H منجر می‌شوند. ولی این «ناهنجاری» نگاشت همریخت را نباید «کمبودی» برای آن دانست، بلکه برعکس، بسیار سودمند است و امکان کاربرد نگاشت همریخت را به‌عنوان ابزار نیرومندی برای بررسی ویژگی‌های گروه‌ها فراهم می‌آورد.

در بسیاری از موقعیت‌های مربوط به تبدیل‌ها، با نگاشت‌های همریخت روبه‌رو می‌شویم. برای نمونه، گروه تقارن‌های چهاروجهی منتظم را در نظر می‌گیریم (شکل ۲۰). این گروه یکریخت با گروه متقارن جای‌گشت‌های چهار عضو است، زیرا یک و تنها یک



شکل ۲۰

حرکت (از نوع اول یا از نوع دوم) وجود دارد که راس‌های A_1، A_2، A_3، A_4، را به‌هر ترتیب مفروض دیگری منجر می‌کند.

اکنون خط‌های راست l_1، l_2، l_3، را در نظر می‌گیریم که نقطه‌های وسط یال‌های روبه‌رو را به‌هم وصل می‌کنند. هر حرکتی که چهاروجهی را بر خودش منطبق کند، جای‌گشتی از l_1، l_2، l_3، را به‌وجود می‌آورد و هر جای‌گشتی از این سه خط راست، با تقارنی از چهاروجهی همراه است. روشن است، حاصل‌ضرب تبدیل‌های چهاروجهی، متناظر است با حاصل‌ضرب جای‌گشت‌های خط‌های راست l_1، l_2، l_3. روی شکل ۲۰ به‌سادگی می‌توان دنبال کرد، چگونه به‌صورتی طبیعی، در این‌جا نگاشت همریخت گروه متقارنِ جای‌گشت‌های چهار عضو A_1، A_2، A_3، A_4به روی گروه متقارن جای‌گشت‌های سه عضو l_1، l_2، l_3، وجود دارد. بدون دشواری می‌توان عضوهایی از گروه «بزرگتر» را پیدا کرد که در این همریختی «روی هم قرار می‌گیرند».

چند مثال دیگر هم می‌آوریم. مجموعهٔ جای‌گشت‌های n عنصر، برای $n>2$، یک گروه غیرآبلی (جابه‌جایی‌ناپذیر) است. از طرف دیگر، عددهای ۱+ و ۱− مربوط به ضرب‌ها هم، گروهی را تشکیل می‌دهند. اکنون هر جای‌گشت زوج از n عنصر را با عدد ۱+ و هر جای‌گشت فرد را با عدد ۱− متناظر می‌کنیم. یک نگاشت همریخت از گروه متقارن جای‌گشت‌های n عنصر، به روی گروه {۱− و ۱+} به‌دست می‌آید، زیرا در بند ۳ دیدیم، حاصل‌ضرب دو جای‌گشتی که هر دو زوج یا هر دو فرد باشند، یک جای‌گشت زوج، و حاصل‌ضرب دو جای‌گشتی که یکی زوج و دیگری فرد باشد، جای‌گشتی فرد است.

مثالی دیگر: اگر هر عدد حقیقی $x\neq 0$ را متناظر با قدرمطلق آن $|x|$ قرار دهیم، آن وقت

نگاشت گروه عددهای حقیقی مثبت و منفی نسبت به‌ضرب (صفر را کنار گذاشته‌ایم)، به روی گروه عددهای حقیقی مثبت نسبت به‌ضرب، یک نگاشت همریخت است، زیرا $|xy| = |x| \cdot |y|$.

پیش از این دیدیم، هر حرکت A از نوع اول در روی صفحه را می‌توان به‌عنوان حاصل‌ضرب دوران V_A دور نقطهٔ ثابت O و یک انتقال موازی D_A در نظر گرفت. دوران‌های دور نقطهٔ O، تشکیل یک گروه می‌دهند. بنابراین، تناظر $A \to V_A$، نگاشت یک‌ارزشی گروه حرکت‌های نوع اول روی صفحه، به روی گروه دوران‌های صفحه دور نقطهٔ O است. ثابت می‌کنیم این، نگاشتی همریخت است. از تجزیه‌های

$$A = V_A D_A\,,\ \ B = V_B D_B$$

نتیجه می‌شود:

$$AB = V_A D_A V_B D_B = (V_A V_B)(V_B^{-1} D_A V_B D_B)$$

پرانتز اول، دوران دور نقطهٔ O و پرانتز دوم، حاصل‌ضرب انتقال $V_B^{-1} D_A V_B$ در انتقال D_B و بنابراین، یک انتقال است. این نشان می‌دهد، حرکت AB، متناظر دوران $V_A V_B$ است، یعنی با نگاشتی همریخت سروکار داریم.

سرانجام، ثابت می‌کنیم گروه بهری G/N از یک گروه دل‌خواه G نسبت به‌هر زیرگروه نرمال N از آن، نگارهٔ همریخت گروه G است.

در واقع، اگر هر عضو g از گروه G را متناظر با همردهٔ gN که شامل g است قرار دهیم، نگاشت همریخت G روی G/N به‌دست می‌آید، زیرا حاصل‌ضرب gh، متناظر همردهٔ ghN است، یعنی حاصل‌ضرب همرده‌های gN و hN.

با مراجعه به ویژگی کلی نگاشت همریخت، ثابت می‌کنیم، عضو همانی ضمن هر همریختی، تبدیل به عضو همانی می‌شود؛ درضمن، عضوهای وارون یکدیگر، بازهم تبدیل به عضوهای وارون می‌شوند.

در واقع، اگر e عضو همانی گروه G و e' نگارهٔ آن در H باشد، آن وقت از $ee = e$ نتیجه می‌شود $e'e' = e'$؛ از این‌جا، اگر ε را عضو همانی گروه H فرض کنیم، به‌دست می‌آید: $\varepsilon = e'e'^{-1} = e'$، حکم اول ثابت شد. اکنون x و y را عضوهای وارون در G، و x' و y' را نگاره‌های آن در H می‌گیریم. از $xy = e$ نتیجه می‌شود $x'y' = e' = \varepsilon$، یعنی x' و y' عضوهای

وارون در H هستند، یعنی

$$(x^{-1})' = x'^{-1}$$

این گزاره‌ها به ما امکان می‌دهند، نگارهٔ هر حاصل‌ضربی از عضوهای G را پیدا کنیم. برای مثال

$$(ab^{-1}c^{-1}dh^{-1})' = a'(b^{-1})'(c^{-1})'d'(h^{-1})' = a'b'^{-1}c'^{-1}d'h'^{-1}$$

قضیه‌ای را که در این‌جا می‌آوریم، اساس تمامی نظریهٔ نگاشت‌های همریخت را تشکیل می‌دهد.

در نگاشت همریخت گروه دل‌خواه G به روی گروه H، مجموعهٔ N از عضوهای گروه G که بر عضو همانی e' از گروه H نگاشته می‌شود، عبارت است از زیرگروهی نرمال در گروه G؛ مجموعه‌ای از عضوهای G که بر عضو ثابت دل‌خواهی از گروه H نگاشته می‌شود، عبارت است از همردهٔ G نسبت به N، به‌این ترتیب تناظر یک‌به‌یک بین همرده‌های G نسبت به N با عضوهای گروه H عبارت است از نگاشتی یکریخت بین H و گروه بهری G/N.

این قضیه را ثابت می‌کنیم. a و b را عضوهای دل‌خواهی از N می‌گیریم. این، به‌آن معنی است که $a' = b' = e'$، که در آن، پریم به‌معنای نگاره‌های عضوهای G در H است. ولی در این صورت

$$(ab)' = a'b' = e'e' = e' \ ,$$

$$(a^{-1})' = a'^{-1} = e'^{-1} = e'$$

یعنی ab و عضوهای وارون a^{-1} و b^{-1} به N تعلق دارند و بنابراین، مجموعهٔ N، یک گروه است. سپس، برای عضو دل‌خواه g از G

$$(g^{-1}ag)' = g'^{-1}a'g' = g'^{-1}e'g' = g'^{-1}g' = e'$$

یعنی $g^{-1}ag$ به‌ازای هر g از G و هر a از N، در N است، ولی از این‌جا نتیجه می‌شود، N زیرگروهی نرمال است. گزارهٔ اول قضیه ثابت شد.

برای اثبات گزارهٔ دوم، در گروه G، عضو دل‌خواه g را انتخاب می‌کنیم و مجموعهٔ U از همهٔ عضوهای u از G که نگارهٔ u' آن‌ها بر نگارهٔ g' از عضو g منطبق است، در نظر

می‌گیریم. فرض کنیم $u \in gN$، یعنی $u = gn$ $(n \in N)$، در این صورت

$$u' = g'n' = g'e' = g'$$

بنابراین $gN \subset U$. برعکس، اگر $u' = g'$، آن‌وقت

$$(g^{-1}u)' = g'^{-1}u' = g'^{-1}g' = e'$$

یعنی $g^{-1}u = n$ $(n \in N)$. از این‌جا $u = gn$ و بنابراین $U \subset gN$. از $gN \subset U$ و $U \subset gN$ نتیجه می‌شود: $U = gN$.

سرانجام، گزارهٔ سوم قضیه روشن است: همرده‌های gN و hN از گروه بهری G/N، متناظر با عضوهای g' و h' در H هستند، و حاصل‌ضرب این همرده‌ها، بنابر دستور

$$gN \cdot hN = ghN$$

متناظر با $(gh)' = g'h'$ است و این همان است که باید ثابت می‌کردیم.

قضیهٔ مربوط به همریختی نشان می‌دهد، هر نگارهٔ همریخت H از گروه G، یکریخت با گروه بهری G/N است. بنابراین، همهٔ نگاره‌های همریخت از گروه G، با دقت تا مرز یکریختی، گروه‌های بهری G هستند.

۷. گروه‌های پیوسته

گروه‌های لی. گروه‌های تبدیل پیوسته. موفقیتی که نظریهٔ گروه‌ها، برای مسألهٔ مربوط به‌حل معادله‌های جبری از درجه‌های بالا پیدا کرد، ریاضی‌دانان سدهٔ نوزدهم را برآن داشت تلاش خود را در زمینهٔ حل معادله‌های از نوع‌های دیگر، و در درجهٔ اول معادله‌های دیفرانسیلی (که نقش زیادی در کاربردهای ریاضیات دارند) آغاز کنند. این تلاش، پایان موفقیت‌آمیزی داشت. گرچه نظریهٔ گروه‌ها، در معادله‌های دیفرانسیلی، نقش به‌کلی متفاوتی در مقایسه با معادله‌های جبری داشت، بررسی‌هایی که برای بهره‌گیری از نظریهٔ گروه‌ها در معادله‌های دیفرانسیلی انجام گرفت، منجر به‌گسترش بی‌اندازهٔ خود مفهوم گروه و پیدایی نظریه‌های تازه‌ای به‌نام گروه‌های پیوسته و گروه‌های *لی* شد که در جهت تکامل

دادن شاخه‌های گوناگون ریاضیات، نقشی مهم و اساسی داشتند.

اگر گروه معادله‌های جبری تنها شامل تعداد محدودی تبدیل بود، گروه معادله‌های دیفرانسیلی، نامتناهی از آب در آمد. به‌جز این معلوم شد تبدیل‌های مربوط به گروه معادلهٔ دیفرانسیلی را می‌توان به‌وسیلهٔ دستگاه محدودی از پارامترها داد که، با تغییر مقدارهای عددی آن‌ها بتوان همهٔ تبدیل‌های گروه را به‌دست آورد. فرض کنید همهٔ تبدیل‌های گروه با مقدار پارامترهای a_1، a_2، ... ، a_r معین شوند. اگر به‌این پارامترها، مقدارهای x_1، x_2، ... ، x_r را بدهیم، تبدیلی مثل X به‌دست می‌آید؛ و اگر مقدارهای تازه y_1، y_2، ... ، y_r را به پارامترها بدهیم، تبدیلی مثل Y حاصل می‌شود. طبق شرط، حاصل‌ضرب این تبدیل‌ها $Z=XY$ در گروهی قرار دارد، یعنی به‌این مقدارهای جدید و معین z_1، z_2، ... ، z_r از پارامترها به‌دست می‌آید. مقدار z_i به x_1، x_2، ... ، x_r، y_1، y_2، ... ، y_r، بستگی دارد، یعنی تابع‌هایی بُرداری از آن‌هاست:

$$z_1=\varphi_1(x_1\,,\,x_2\,,\,\dots\,,\,x_r\,,\,y_1\,,\,y_2\,,\,\dots\,,\,y_r)\;,$$

$$z_2=\varphi_2(x_1\,,\,x_2\,,\,\dots\,,\,x_r\,,\,y_1\,,\,y_2\,,\,\dots\,,\,y_r)\;,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$z_r=\varphi_r(x_1\,,\,x_2\,,\,\dots\,,\,x_r\,,\,y_1\,,\,y_2\,,\,\dots\,,\,y_r)\;.$$

گروه‌هایی که عضوهای آن‌ها بستگی پیوسته‌ای با مقدارهای دستگاه محدودی از پارامترها داشته باشند، و قانون ضرب آن‌ها به‌وسیلهٔ تابع‌های با مشتق دوم φ_1، ... ، φ_r بیان می‌شود، گروه‌های *لی* نام دارند، به‌نام سوفیس *لی* ریاضی‌دان نروژی که نخستین‌بار این گروه‌ها را بررسی کرد.

در نیمهٔ نخست سدهٔ نوزدهم، در کارهای *لُباچوسکی* دستگاه هندسی تازه‌ای طرح شد که به‌نام او معروف شده است. به‌تقریب در همان زمان، دستگاه هندسی مستقلی پدید آمد که هندسهٔ تصویری نام گرفت؛ اندکی دیرتر هندسهٔ ریمانی ساخته شد. در نتیجه، در نیمهٔ دوم سدهٔ نوزدهم، دستگاه‌های هندسی مستقلی وجود داشت که با دیدگاه‌های گوناگون «شکل‌های فضایی دنیای واقع» را بررسی می‌کردند. پیوستن همهٔ این دستگاه‌های مختلف هندسی، با حفظ تفاوت‌های کیفی آن‌ها، تنها به یاری نظریهٔ گروه‌ها ممکن شد.

تبدیل‌های یک‌به‌یک مجموعهٔ نقطه‌های یک فضای هندسی را در نظر می‌گیریم، که بستگی‌های اساسی بین شکل‌هایی را که در این هندسه بررسی می‌شود، تغییر ندهد.

مجموعهٔ این تبدیل‌ها، گروهی را تشکیل می‌دهد که آن را گروه حرکت‌ها یا **خودریختی** (=خوددیسگی *automorphism*) هندسهٔ مفروض گویند. گروه حرکت‌ها به‌طور کامل با هندسهٔ مفروض مشخص می‌شود، زیرا اگر گروه حرکت‌ها معلوم باشد، آن وقت هندسهٔ متناظر آن را می‌توان به‌عنوان دانشی در نظر گرفت که کارش بررسی چنان ویژگی‌هایی از مجموعهٔ نقطه‌هاست که ضمن تبدیل‌های گروه مفروض، تغییر نمی‌کنند. روش رده‌بندی دستگاه‌های مختلف هندسی به‌وسیلهٔ گروه حرکت‌های آن‌ها، در نیمهٔ دوم سدهٔ نوزدهم و در کارهای **ف.کلاین** پدید آمد. دربارهٔ این روش و دستگاه‌های مختلف هندسی در بخش مربوط به فضاهای انتزاعی صحبت کرده‌ایم. در این جا تنها یادآوری می‌کنیم، گروه‌هایی از حرکت‌های دستگاه‌های هندسی که در سدهٔ گذشته مطالعه شد، همان گروه‌های *لی* هستند. با توجه به این مسأله بود که مطالعهٔ گروه‌های *لی* اهمیت خاصی پیدا کرد.

به‌دلیل بستگی‌های فراوانی که گروه‌های *لی* با شاخه‌های گوناگون ریاضیات و مکانیک داشت، این نظریه با نیروی زیادی از همان آغاز پیدایی خود تا زمان ما پیش رفته است. در ضمن معلوم شد، برخی از مسأله‌هایی که برای گروه‌های محدود قابل حل نبود، به‌طور نسبی به‌سرعت برای گروه‌های *لی* قابل حل است. ازجمله، مسألهٔ مربوط به رده‌بندی گروه‌های متناهی ساده (یعنی گروه‌های متناهی‌ای که دارای زیرگروه نرمال مایه‌دار نیستند به‌نتیجه زیادی نرسیده است، در حالی که رده‌بندی گروه‌های سادهٔ *لی*، در پایان سده نوزدهم و به‌وسیلهٔ ***کی‌لینگ*** و ***کارتان*** پیدا شده بود[1]. ***و.و.مُورُوزوف***، ***آ.ای.مالتِسف*** و ***ب.ب.دین‌کین*** ریاضی‌دانان شوروی توانستند راه‌حل کامل مسألهٔ رده‌بندی ***زیرگروه‌های*** سادهٔ گروه‌های *لی* را، که همه‌جا در انتظار حل آن بودند، پیدا کنند. ***ای.م.هلفوند*** و ***م.آ.نای‌مارک*** ریاضی‌دانان دیگر شوروی به پژوهش در جهت دیگری پرداختند و نمایش‌های پیوسته گروه‌های سادهٔ *لی* به‌وسیلهٔ تبدیل‌های یکانی فضای هیلبرتی را جست‌وجو کردند، مسأله‌ای که برای آنالیز و فیزیک اهمیت دارد.

بررسی گروه‌های *لی* با روش خاصی که مبتنی‌بر «گروه‌های بی‌نهایت کوچک» (=گروه‌های اینفینی تزیمال) یا جبر *لی* است، تحقق می‌یابد. بحث مفصل‌تر در این باره را در بند ۱۳ خواهید دید.

---

۱. رده‌بندی گروه‌های متناهی ساده در سال ۱۹۸۰ و با کار صدها ریاضیدان به پایان رسید (ویراستار).

گروه‌های توپولوژیک. همراه با پیشرفت گستردهٔ نظریهٔ سنتی گروه‌های *لی* در اتحاد شوروی، نظریهٔ کلی‌تر گروه‌های توپولوژیک یا گروه‌های پیوسته، به‌موفقیت‌های استثنایی رسید. برخلاف مفهوم گروه *لی* که در آن باید عضوهای گروه با دستگاه محدودی از پارامترها معین شوند و در ضمن قانون ضرب به یاری تابع‌های مشتق‌پذیر بیان شود، مفهوم گروه توپولوژیک ساده‌تر و عام‌تر است. به‌گروهی توپولوژیک گفته می‌شود که برای عضوهای آن، به‌جز عمل‌های معمولی گروه، مفهوم همسایگی و نزدیکی هم تعریف شده باشد، درضمن از همسایگی عضوهای گروه، همسایگی حاصل‌ضرب‌های آن‌ها و همسایگی عضوهای وارون نتیجه شود.

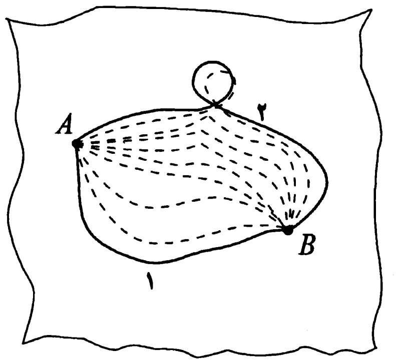
در آغاز، مفهوم گروه توپولوژیک، به‌خاطر مرتب کردن و به‌نظم درآوردن بسیاری از مفهوم‌های نظریهٔ گروه‌های *لی* پدید آمد. ولی به‌تدریج معلوم شد که این مفهوم، تا چه‌اندازه برای شاخه‌های دیگر ریاضیات هم اهمیت دارد. نخستین کارها در زمینهٔ نظریهٔ گروه‌های توپولوژیک کلی، در آغاز سال‌های ۲۰ سدهٔ بیستم انجام گرفت، ولی نتیجه‌گیری‌های عمیقی که این نظریه را به‌صورت شاخهٔ تازه‌ای معرفی کرد، مربوط به پایان سال‌های ۲۰ و آغاز سال‌های ۳۰ سدهٔ بیستم است. بخش عمده‌ای از این موفقیت‌ها به *ل.س.پون تریاگین* ریاضی‌دان شوروی تعلق دارد که باید او را یکی از بنیانگذاران نظریهٔ امروزی گروه‌های پیوسته دانست. کتاب او به‌نام «گروه‌های پیوسته»، که برای نخستین بار در ادب ریاضی، آگاهی‌های جامعی دربارهٔ نظریهٔ گروه‌های پیوسته در اختیار می‌گذارد، تا چند ده سال پس از انتشار، تنها کتاب راهنما در این زمینه بود.

۸. گروه‌های بنیادی (یا اصلی)

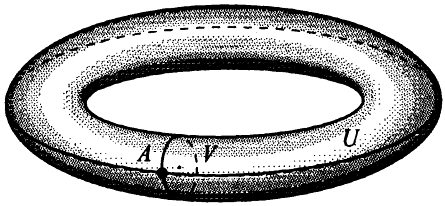
گروه‌های مشخصی راکه در بندهای پیشین بررسی کردیم، اغلب گروه‌های تبدیل این یا آن مجموعه بودند. استثنایی که در این باره وجود داشت، گروه عددها نسبت به جمع و ضرب بود. اکنون به مثالی می‌پردازیم که در نظریهٔ گروه‌ها اهمیت زیادی دارد و از همان آغاز، به‌جای استفاده از تبدیل‌ها، گروه را به‌عنوان یک دستگاه جبری همراه با یک عمل تعریف می‌کنیم.

گروه بنیادی (یا گروه اصلی). سطحی مثل S و نقطهٔ متحرک M را در روی آن در نظر می‌گیریم. فرض می‌کنیم، نقطهٔ M باید روی سطح منحنی پیوسته‌ای را رسم کند که از A آغاز می‌شود و در B پایان می‌یابد. در این صورت مسیر مشخصی از A تا B به‌دست می‌آید. این مسیر می‌تواند خودش را چندبار قطع کند یا در برخی تکه‌ها روی خودش قرار گیرد. برای نشان دادن مسیر، تنها کافی نیست منحنی را که نقطهٔ M روی آن حرکت کرده است، بدهند. باید تکه‌هایی از منحنی را هم مشخص کرد که نقطهٔ M چندبار از آن‌ها گذشته است؛ درضمن جهت حرکت نقطه هم باید در دست باشد. به‌عنوان مثال، نقطه می‌تواند محیط دایره‌ای را چندبار و در جهت‌های مختلف بپیماید؛ درضمن همهٔ این مسیرهای دایره‌ای مختلف به‌حساب می‌آیند. دو مسیر را، وقتی که آغاز و پایان آن‌ها یکی است، هم‌ارز گویند، به شرطی که یکی از آن‌ها را بتوان با تغییری پیوسته، تبدیل به دیگری کرد. روی صفحه یا روی سطح کره، هر دو مسیری که نقطه‌های A و B را به‌هم می‌پیوندند، هم‌ارزند (شکل ۲۱). ولی روی سطح چنبره، از جمله مسیرهای بستهٔ U و V (شکل ۲۲)، که از نقطهٔ A آغاز و در همان نقطهٔ A پایان یافته‌اند، هم‌ارز نیستند.

اگر به‌جای چنبره، سطح استوانه‌ای دوار را (استوانه دواری که سطح جانبی آن از دو طرف ادامه یافته باشد) در نظر بگیریم و روی آن مسیر X را انتخاب کنیم (شکل ۲۳)، آن‌وقت به‌سادگی معلوم می‌شود که هر مسیر بسته‌ای که از نقطهٔ A آغاز شده باشد، با مسیر X^n هم‌ارز است (که در آن n می‌تواند برابر ۰، ۱±، ۲±، ...، باشد). منظور از X^n $(n>0)$، تکرار n بار مسیر X است؛ X^0 مسیر صفر ـ تنها شامل نقطهٔ A است؛ مسیر X^{-n} همان مسیر X^n در جهت عکس است؛ مثال: $Z\sim X^{-1}$، $Y\sim X^2$، $U\sim X^0$ (شکل ۲۳). این مثال، اهمیت



شکل ۲۱

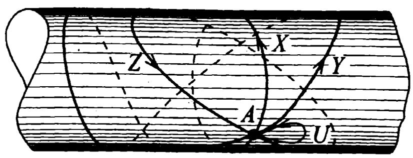


شکل ۲۲

مفهوم همه‌ارزی مسیرها را نشان می‌دهد: درعین حال که مسیرهای بستهٔ روی سطح استوانه‌ای، مجموعه‌ای بس بزرگ را تشکیل می‌دهد، از دیدگاه هم‌ارزی، همهٔ این مسیرها تنها منجر به دورزدن دایرهٔ X به تعداد دفعات مشخصی می‌شوند. در حالت $m \neq n$، مسیرهای X^m و X^n هم‌ارز نیستند.

به بررسی سطحی دل‌خواه برمی‌گردیم و فرض می‌کنیم دو مسیر در روی آن داده شده باشد: مسیر U که از نقطهٔ B می‌رود و مسیر V که از نقطهٔ B به‌نقطهٔ C می‌رسد. اگر شرط کنیم، نقطهٔ مفروض، ابتدا مسیر AB و سپس BC را بپیماید، مسیر AC به‌دست می‌آید که، به‌طور طبیعی و طبق معمول، آن را حاصل‌ضرب مسیر $U=AB$ در مسیر $V=BC$ می‌نامیم و با نماد UV نشان می‌دهیم. اگر مسیرهای U و V، به‌ترتیب، با مسیرهای U_1 و V_1 هم‌ارز باشند، آن‌وقت حاصل‌ضرب آن‌ها UV و U_1V_1 هم، هم‌ارزند. ضرب مسیرها، به‌این مفهوم از قانون شرکت‌پذیری پیروی می‌کنند که، یعنی اگر یکی از حاصل‌ضرب‌های $U(VW)$ یا $(UV)W$ معین باشد، آن‌وقت دیگری هم معین است و دو حاصل‌ضرب، مسیرهای هم‌ارز را معرفی می‌کنند. اگر نقطهٔ M، مسیر $U=AB$ را در جهت عکس بپیماید، آن‌وقت مسیر وارونِ $U^{-1}=BA$، که از نقطهٔ B به‌نقطهٔ A می‌رود، به‌دست می‌آید. حاصل‌ضرب مسیر AB در مسیر معکوس آن BA، مسیر بسته‌ای هم‌ارز مسیر صفر است که تنها شامل نقطهٔ A است.

بنابر تعریف، هر دو مسیری را نمی‌توان درهم ضرب کرد، بلکه تنها مسیرهایی درهم ضرب می‌شود که در آن‌ها، نقطهٔ پایان مسیر اول بر نقطهٔ آغاز مسیر دوم منطبق باشد. این



شکل ۲۳

کمبود را وقتی می‌توان ازبین برد که تنها مسیرهای بسته را در نظر بگیریم، مسیرهای بسته‌ای که نقطهٔ آغاز آن‌ها یکی، یعنی نقطهٔ A باشد. هر دو مسیر از این‌گونه را می‌توان درهم ضرب کرد که در نتیجه، دوباره مسیری بسته با آغاز A به‌دست می‌آید. به‌جز این، برای هر مسیر بسته که از A آغاز شده باشد، مسیر وارون هم همین ویژگی‌ها را دارد.

اکنون شرط می‌کنیم، مسیرهای هم‌ارز را نمایندگان مختلف یک «مسیر» به‌حساب آوریم، «مسیرهایی» که با روش‌های مختلفی روی سطح داده شده باشند؛ و مسیرهای ناهم‌ارز را نمایندگان «مسیرهایی» بدانیم که به‌واقع باهم اختلاف دارند. نکته‌هایی را که در بالا گفتیم، نشان می‌دهد، در چنین حالتی، مجموعهٔ همهٔ مسیرهای بسته (گیومه‌ها را کنار گذاشته‌ایم) که از نقطه‌ای مثل A از سطح آغاز می‌شوند، نسبت به‌عمل ضرب مسیرها، یک گروه تشکیل می‌دهند. عضو همانی (یا عضو خنثا) در این گروه، عبارت است از مسیر صفر و عضو وارون برای هر مسیر مفروض، همان مسیر تنها در جهت عکس است.

گروه مسیرها، در حالت کلی، تنها به شکل سطح بستگی ندارد، بلکه به انتخاب نقطهٔ آغاز A هم مربوط است. ولی اگر سطح به تکه‌های جداگانه تقسیم نشده باشد، یعنی هر دو نقطهٔ دل‌خواه آن را بتوان با مسیری پیوسته و روی سطح به‌هم پیوست، آن‌وقت گروه‌های مسیرها، که پاسخ‌گوی نقطه‌های مختلف سطح هستند، گروه‌هایی یکریخت (ایزومورفیسم) می‌شوند و، در این حالت، می‌توان به‌طور ساده دربارهٔ گروه مسیرهای سطح S صحبت کرد، بدون این که نامی از نقطهٔ A ببریم. همین گروه مسیرهای سطح است که به آن ***گروه بنیادی*** (یا گروه اصلی) گویند.

اگر سطح S، صفحه یا سطح کره باشد، آن‌وقت گروه مسیرها تنها شامل یک عضو است، زیرا روی صفحه یا سطح کره، هر مسیری به یک نقطه منقبض می‌شود، ولی همان‌طور که دیدیم روی سطح نامحدود استوانهٔ دوار، مسیرهای بسته‌ای وجود دارد که به یک نقطه منقبض نمی‌شوند. چون هر مسیر بسته‌ای که روی سطح استوانه‌ای از نقطهٔ A آغاز شود، هم‌ارز درجه‌ای از مسیر X است (شکل ۲۳)، درضمن درجه‌های مختلف X، هم‌ارز باهم نیستند. بنابراین گروه مسیرهای سطح استوانه‌ای، گروهی دوری ولی نامتناهی است. می‌توان ثابت کرد، گروه مسیرهای چنبره (شکل ۲۲)، از مسیرهای به‌صورت $U^m V^n$ تشکیل شده است (m و n می‌توانند هریک از عددهای ۰ ، ۱± ، ۲± ، ... را اختیار کنند)؛ درضمن $UV = VU$ و برابری $U^m V^n = U^{m_1} V^{n_1}$ تنها وقتی برقرار است که داشته باشیم $m = m_1$ و $n = n_1$ (به‌یاد بیاوریم، در بررسی گروه مسیرها، برابری مسیرها

به معنای هم‌ارزی آن‌ها است).

اهمیت گروه مسیرها را می‌توان با ویژگی آن که در این‌جا می‌آید، روشن کرد. فرض می‌کنیم، به‌جز سطح S، سطح دیگری مثل S_1 هم داده شده باشد، به‌نحوی که بتوان بین نقطه‌های دو سطح S و S_1، تناظر یک‌به‌یک و پیوسته برقرار کرد. برای نمونه، می‌توان این تناظر را به این ترتیب به‌دست آورد که سطح S_1 از سطح S، به یاری تغییر شکل پیوستهٔ آن، بدون ایجاد پارگی و بدون این که نقطه‌های آن روی هم قرار بگیرند، به‌دست آمده باشد. هر مسیر واقع بر سطح S، متناظر با مسیری در سطح S_1 است. درضمن، مسیرهای هم‌ارز، متناظر با مسیرهای هم‌ارزند، حاصل دو مسیر در S متناظر با حاصل‌ضرب دو مسیر هم‌ارز آن‌ها در S_1 است، به‌نحوی که گروه مسیرهای روی S_1، با گروه مسیرهای S یکریخت است. به‌زبان دیگر، اگر گروه مسیرها را از دیدگاه انتزاعی، یعنی با دقت تا یکریختی گروه‌ها درنظر بگیریم، برای همهٔ تبدیل‌های ممکنِ سطح، به‌شرطی که پیوسته و یک‌به‌یک باشند، گروهی ناورداست. اگر گروه مسیرها، برای دو سطح مختلف باشند، به‌این معنی است که این دو سطح را نمی‌توان، به‌طور پیوسته، به‌یکدیگر تبدیل کرد. به‌عنوان مثال، صفحه را نمی‌توان بدون پاره کردن و روی هم گذاشتن، بر سطح استوانه‌ای قرار داد، زیرا گروه مسیرهای صفحه شامل یک عضو و گروه مسیرهای سطح استوانه‌ای بی‌نهایت عضو دارد.

ویژگی‌هایی از شکل که ضمن تبدیل‌های پیوسته و یک‌به‌یک متقابل (همسانریخت) بی‌تغییر می‌ماند، در شاخهٔ خاصی از ریاضیات که توپولوژی نام دارد، بررسی می‌شود و اندیشه‌های بنیادی آن را در بخش هجدهم دیده‌اید. ناورداهای تبدیل‌های همسانریخت، ***ناورداهای توپولوژیک*** نامیده می‌شود. گروه مسیرها، یکی از نمونه‌های اساسی ناورداهای توپولوژیک است. روشن است که گروه مسیرها را، به‌جز دربارهٔ سطح‌ها، بلکه برای هر مجموعه‌ای از نقطه‌ها می‌توان تعریف کرد؛ تنها به این شرط که در این مجموعه‌ها بتوان دربارهٔ مسیرها و تغییر شکل آن‌ها صحبت کرد.

رابطه‌های معرف. روش‌های محاسبهٔ گروه‌های مسیرها، به تفصیل در توپولوژی مطالعه می‌شود. در عین حال، این گروه‌ها را با روش ویژه‌ای تعریف می‌کنند که در نظریهٔ گروه‌ها و برای گروه‌های انتزاعی، و البته نه‌تنها برای گروه‌های بنیادی در توپولوژی، به‌کار می‌رود. این روش را می‌آوریم.

G را یک گروه در نظر می‌گیریم. عضوهای g_1، g_2، ... ، g_n، وقتی عضوهای مولد گروه

G نامیده می‌شوند که هر عضو g بتواند به‌صورت

$$g = g_{i_1}^{\alpha_1} g_{i_2}^{\alpha_2} \dots g_{i_k}^{\alpha_k}$$

نشان داده شود، که در آن، هریک از i_1، i_2، ...، i_k، عددی از عددهای ۱، ۲، ...، n هستند؛ اندیس‌هایی که پشت‌سر هم نباشند می‌توانند برابر باشند؛ تعداد عامل‌های k دل‌خواه است؛ نماهای α_1، α_2، ...، α_k، عددهایی درست و مخالف صفر (مثبت یا منفی) هستند.

برای این که گروه G را بشناسیم، کافی است به‌جز عضوهای مولد آن، بدانیم کدام حاصل‌ضرب‌ها نمایندهٔ یک عضو گروه هستند و کدام حاصل‌ضرب‌ها مُعرف عضوهای مختلف‌اند. بنابراین، برای مفروض بودن گروه، باید همهٔ برابری‌های به‌صورت

$$g_{i_1}^{\alpha_1} g_{i_2}^{\alpha_2} \dots g_{i_k}^{\alpha_k} = g_{j_1}^{\beta_1} g_{j_2}^{\beta_2} \dots g_{j_l}^{\beta_l}$$

را که در گروه G وجود دارند، شمرد. از آن‌جا که چنین برابری‌هایی، همیشه یک مجموعهٔ بی‌پایان را تشکیل می‌دهند، به‌جای شمردن همهٔ آن‌ها، تنها چنان برابری‌هایی را می‌دهند که بقیه را بتوان به‌یاری اصل موضوع‌های گروه از آن‌ها نتیجه گرفت. همین برابری‌ها هستند که ***رابطه‌های معرف*** نامیده می‌شوند.

روشن است که یک گروه را می‌توان به‌صورت‌های متفاوتی به‌یاری رابطه‌های معرف مشخص کرد.

برای نمونه، گروه H را با مولدهای a، b و رابطه‌های

$$a^2 = b^3 \,,\; ab = ba \qquad (10)$$

در نظر می‌گیریم. فرض می‌کنیم $c = ab^{-1}$، خواهیم داشت:

$$a = bc \,,\; a^2 = b^2c^2 \,,\; b^3 = b^2c^2 \,,\; b = c^2 \,,\; a = c^3$$

می‌بینیم، همهٔ عضوهای گروه H را می‌توان برحسب یک عضو c بیان کرد، در ضمن

$$a = c^3 \,,\; b = c^2$$

چون رابطه‌های (۱۰)، بلافاصله از این برابری‌ها به‌دست می‌آیند، بنابراین برای c، رابطهٔ نابدیهی وجود ندارد. بنابراین، گروه H، گروهی دُوری و نامتناهی با عضو مولد c است.

اگر بتوان چنان مولدهایی را در گروه پیدا کرد که به‌وسیلهٔ هیچ‌گونه رابطهٔ نابدیهی به‌هم بستگی نداشته باشند، آن‌وقت گروه را *آزاد* و این مولدها را ***مولدهای آزاد*** می‌نامند. در مثل، اگر گروه دارای مولدهای آزاد a و b باشد، آن‌وقت هر عضو آن، به‌طور منحصربه‌فرد و به‌صورت

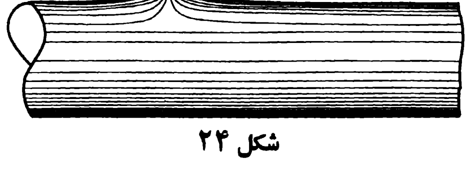
$$a^{\alpha_0} b^{\beta_1} a^{\alpha_1} b^{\beta_2} a^{\alpha_2} \dots b^{\beta_k} a^{\alpha_k}$$

نوشته می‌شود که در آن: $k = 0, 1, 2, \dots, n$، نماهای α_0، β_1، α_1، ...، β_k، α_k عددهای درستی (مثبت یا منفی) و مخالف صفرند، به‌جز «نماهای مرزی» α_0 و α_k که می‌توانند مقدارهای صفر را بپذیرند. همین استدلال دربارهٔ گروه‌های آزادِ با تعداد بیشتری مولد هم درست است.

اگر مولدها و رابطه‌های معرف را برای دو گروهی که عضوهای مشترکی ندارند، بنویسیم، آن‌وقت از اجتماع این رابطه‌ها، گروه تازه‌ای به‌دست می‌آید که گروه آزاد حاصل‌ضرب این دو گروه نامیده می‌شود.

نظریهٔ گروه‌های آزاد و، کلی‌تر از آن، نظریهٔ حاصل‌ضرب‌های آزاد، جای مهمی در نظریهٔ گروه‌ها دارد. اگر با دیدگاه هندسی به‌این موضوع بنگریم، حاصل‌ضرب‌های آزاد گروه‌های H_1 و H_2، گروه مسیرهای چنان شکلی است که بتوان آن را به‌صورت مجموع دو شکل بسته که تنها در یک نقطه به‌هم رسیده‌اند و گروه‌های مسیر آن‌ها H_1 و H_2 هستند، در نظر گرفت. می‌دانیم، گروه مسیرهای سطح استوانه‌ای گروه آزادی با یک مولد است. از اشاره‌ای که در این جا داشتیم، از جمله نتیجه می‌شود، گروه مسیرهای سطحی که در شکل ۲۴ نشان داده شده است، گروه آزادی با دو مولد است.

شبیه آن‌چه برای گروه بنیادی سطح تعریف کردیم، می‌توان برای گروه بنیادی جسم‌های



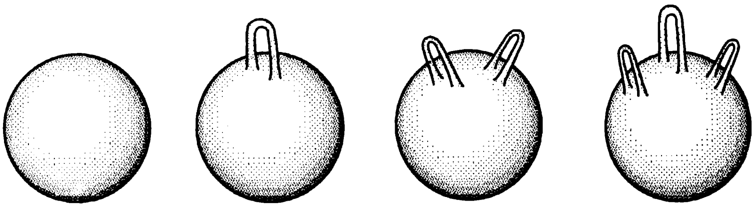
شکل ۲۴

فضایی، صورت محدود یا نامحدود، هم درنظر گرفت.

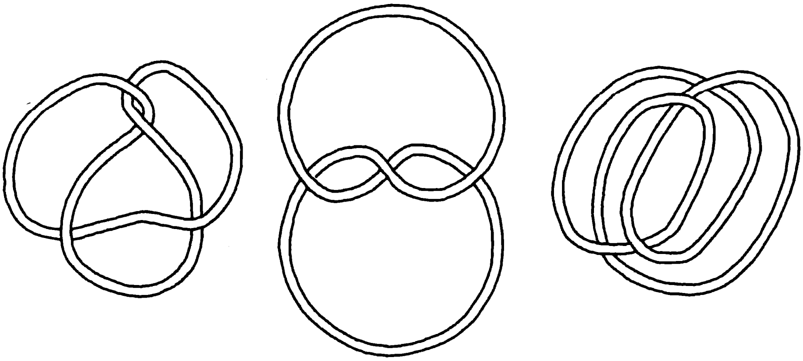
گره‌ها و گروه گره‌ها. همان‌طور که گفتیم، دو سطح را وقتی از نظر توپولوژیک، یکی به‌حساب می‌آورند که بتوان یکی از آن‌ها را با تبدیل پیوسته متقابل (همسانریخت) و به‌صورت یک‌به‌یک به دیگری منجر کرد. مسألهٔ به نظم درآوردنِ همهٔ سطح‌های بسته، حل شده است. روشن شده است هر سطح بسته‌ای که در فضای معمولی ما واقع باشد، از دیدگاه توپولوژی، یا هم‌ارز با کره است و یا هم‌ارز با کره‌ای که چند دسته دارد (شکل ۲۵). برای نمونه، سطح چنبره را، که در شکل ۲۲ داده‌ایم، می‌توان با تغییر شکل همسانریخت به‌کره‌ای با یک دسته و سطح مکعب را به سطح کره و غیره منجر کرد. بنابراین مطالعهٔ گروه‌های بنیادی برای سطح‌های بسته، چندان جالب نیست، زیرا سطح‌های بسته را، بدون این گروه‌ها هم، به‌نظم درآورده‌اند. با وجود این، مسأله‌های بسیار ساده‌ای وجود دارند، که بدون وارد کردن گروه‌های بنیادی، نتوانسته‌اند آن‌ها را حل کنند. ازجملهٔ این مسأله‌ها، می‌توان مسألهٔ مهم و مشهور مربوط به گره‌ها را نام برد.

گره به منحنی بسته‌ای گفته می‌شود که در فضای معمولی سه بُعدی واقع باشد. وضع قرارگرفتن گره در فضا می‌تواند بسیار مختلف باشد (در شکل ۲۶ ، نمونه‌هایی آمده است). دو گره را هم‌ارز گویند وقتی که بتوان یکی از آن‌ها را با تغییر شکل دادن با روندی پیوسته، بدون پاره‌کردن و بدون این که با خودش گیر کند، به دیگری تبدیل کرد. از همین جا، بلافاصله دو مسأله پدید می‌آید: ۱) اگر تصویرهای دو گره را در روی صفحه داشته باشیم، چگونه می‌توان هم‌ارزی یا ناهم‌ارزی آن‌ها را تشخیص داد؟ ۲) گره‌های ناهم‌ارز را چگونه می‌توان رده‌بندی کرد؟

هردویِ این مسأله‌ها تاکنون بی‌حل باقی مانده‌اند و برخی راه‌حل‌های جزیی که در این



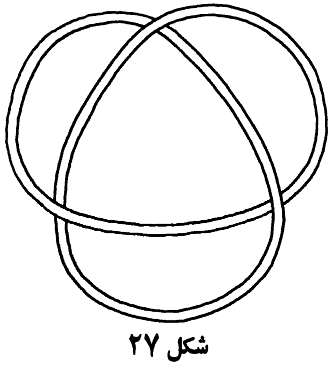
شکل ۲۵



شکل ۲۶

زمینه پیدا شده است، به نظریهٔ گروه‌ها مربوط می‌شود[1].

نقطه‌هایی از فضا را که متعلق به‌گره مفروض‌اند، درمی‌آوریم و گروه بنیادی مجموعهٔ نقطه‌های باقی‌مانده را درنظر می‌گیریم. همین گروه را گروه گره می‌نامند. بلافاصله روشن می‌شود، اگر گره‌ها هم‌ارز باشند، آن‌وقت گروه آن‌ها یکریخت (ایزومورف) هستند. بنابراین، از یکریخت نبودن گروه‌های گره‌ها، می‌توان دربارهٔ ناهم‌ارزی خود گره‌ها نتیجه گرفت. برای مثال، گروه گرهی که منجر به محیط دایره می‌شوند، یک گروه دُوری است، ولی گروه گرهی که به‌شکل «سه‌برگی» (شکل ۲۷) است، گروهی پیچیده‌تر است. گروه اخیر جابه‌جایی‌ناپذیر است و بنابراین با گروه دایره، یکریخت یا ایزومورف نیست. در نتیجه می‌توان گفت، گره سه‌برگی را، بدون پاره کردن آن نمی‌توان به دایره تبدیل کرد، حقیقتی که



شکل ۲۷

۱. در سال‌های اخیر پیشرفت‌های زیادی در جهت حل این مسأله‌ها شده است (ویراستار).

به‌صورت آزمایشی روشن است، ولی برای اثبات آن، به ردیف استدلال‌های دقیقی از ریاضیات نیاز داریم.

با کمال تأسف، با استفاده از گروه گره‌ها هم، مسأله‌های دشوار مربوط به گره‌ها بدون حل باقی مانده‌اند. در توپولوژی، روش‌های بسیار ساده‌ای وجود دارد که با مفروض بودن شکل گره، می‌توان مولدها و رابطه‌های معرف گروه آن را پیدا کرد. ولی برای این که از گروه‌ها استفاده کنیم و به مقایسهٔ گره‌های مختلف بپردازیم، باید بتوانیم مسألهٔ یکریختی (ایزومورفیسم) یا نایکریختی گروه‌ها را با معلوم بودن مولدها و رابطه‌های معرف آن‌ها حل کنیم، مسأله‌ای که تاکنون حل نشده است. از این گذشته، پ.س.*نُوی‌کُوف* ریاضی‌دان شوروی قضیهٔ مهمی را ثابت کرد که بنابر آن، هیچ روش منظم واحدی (دقیق‌تر، هیچ آلگوریتم نرمالی) پیدا نمی‌شود که به‌کمک آن بتوان همیشه روشن کرد: آیا دو دستگاه مفروض رابطه‌های معرف برای مولدهایی از یک گروه، می‌تواند وجود داشته باشد یا نه؟ این قضیه، ما را به تردید می‌اندازد که: آیا به‌واقع می‌تواند روشی کلی برای روشن کردن هم‌ارزی گره‌ها، با مفروض بودن شکل آن در روی صفحه، وجود داشته باشد؟

۹. نمایش‌های گروه و سرشت‌های آن‌ها

نظریهٔ کلی گروه به دلیل روش‌های خود، تا اندازه‌ای هندسهٔ مقدماتی را به‌یاد می‌آورد: در هر دو حالت، پایهٔ کار بر دستگاه معینی از اصل موضوع‌ها گذاشته شده است که با آغاز از آن تمامی نظریه ساخته می‌شود. ولی نمونهٔ هندسهٔ تحلیلی نشان می‌دهد، برای بررسی مسأله‌های هندسی، تا چه اندازه روش‌های تحلیلی و عددی لازم است.

به‌کار گرفتن روش‌های آنالیز و جبر سنتی در نظریهٔ گروه‌ها، نظریهٔ *نمایش‌های گروه* نام دارد. همان‌طور که هندسهٔ تحلیلی روش‌هایی را به‌دست می‌دهد که از یک طرف، مسأله‌های هندسی به یاری آنالیز حل می‌شوند و، از طرف دیگر، بر بسیاری از مسأله‌های دشوار آنالیز پرتوی هندسی می‌اندازد، همان‌طور هم، نظریهٔ نمایش‌ها نه تنها به‌عنوان ابزاری کمکی به بررسی ویژگی‌های گروه خدمت می‌کند، بلکه با به‌هم پیوستن مسأله‌ها و مفهوم‌های ژرف آنالیز و نظریهٔ گروه‌ها، امکان می‌دهد برای حقیقت‌های مربوط به گروه، رابطه‌های عددی به‌دست آید، و برای بستگی‌های تحلیلی تفسیرهایی از نظریهٔ گروه‌ها پیدا

شود. به‌ویژه امروز، بخش بزرگی از کاربردهای نظریهٔ گروه‌ها در فیزیک، مربوط به نظریهٔ نمایش‌ها است.

نمایش گروه به‌یاری ماتریس‌ها. در جبر خطی (بخش شانزدهم را ببینید)، عمل ضرب برای ماتریس‌ها بررسی می‌شود. این عمل شرکت‌پذیر است ولی، در حالت کلی، جابه‌جایی‌پذیر نیست. ماتریس‌های مربعی ناتکین (وارون‌پذیر) از مرتبهٔ معینی، نسبت به عمل ضرب، تشکیل یک گروه می‌دهند. در واقع، نتیجهٔ ضرب دو ماتریس ناتکین ماتریسی ناتکین است، نقش عضو همانی به‌عهدهٔ ماتریس همانی است و برای هر ماتریس ناتکین، ماتریس وارونی وجود دارد که باز هم ناتکین است.

گروه G را در نظر می‌گیریم و هر عضو g از آن را متناظر با ماتریس A_g، که ماتریس ناتکین معینی از مرتبهٔ n و شامل عددهای مختلط است، قرار می‌دهیم؛ درضمن این تناظر به‌نحوی باشد که هر ضرب عضوهای گروه، متناظر با ضرب ماتریس‌های متناظر باشد. $A_{gh}=A_g.A_h$. در این صورت می‌گویند، نمایشی از گروه G با ماتریس‌های مرتبهٔ n داده شده است. اغلب واژهٔ «ماتریس» را کنار می‌گذارند و به‌طور ساده می‌گویند: نمایشِ مرتبهٔ n گروه G. نمایش مرتبهٔ n گروه مفروض G عبارت است از نگاشت همریخت (هُومُومُورف) گروه G در گروه ماتریس‌های ناتکین مرتبهٔ n. از ویژگی‌های کلی نگاشت‌های همریخت نتیجه می‌شود، برای هر نمایش، عضو همانی گروه G تبدیل به ماتریس همانی می‌شود، عضوهای وارونِ یکدیگر در G، تبدیل به ماتریس‌های وارونِ یکدیگر می‌شود.

ماتریس‌های مرتبهٔ اول، عددهای مختلط جداگانه‌اند. بنابراین، یک نمایش درجهٔ اول گروه G به معنی آن است که برای هر عضو گروه G، عدد مختلطی داده باشیم، و درضمن حاصل‌ضرب عضوهای گروه پاسخ‌گوی حاصل‌ضرب عددهای مختلط متناظر است. برای نمونه، نگاشتی که به‌ازای آن، جای‌گشت زوج گروه متقارن، متناظر عدد ۱ و جای‌گشت فرد متناظر ۱ − است، نمایشی درجه اول است.

اگر هر عضو گروه G را متناظر با ماتریس همانی (یکه) E و از درجهٔ n قرار دهیم، نمایشی از گروه G به‌دست می‌آید که آن را نمایش همانی (یکه) درجهٔ n می‌نامیم. اگر گروه G متناهی و شامل عضوهای غیرواحد باشد، آن‌وقت به‌جز نمایش‌های همانی با درجه‌های مختلف، گروه G دارای بی‌نهایت نمایش دیگر هم هست. روش‌های پیداکردن این نمایش‌ها را در این جا شرح می‌دهیم.

اگر یکی از نمایش‌های گروه G را بدانیم، می‌توان مجموعهٔ بی‌پایانی از نمایش‌های دیگر را به‌دست آورد. فرض می‌کنیم $g \to A_g$ نمایش مفروضی از گروه G به‌وسیلهٔ ماتریس‌های مرتبهٔ n باشد. ماتریس دل‌خواه ناتکین P را با همان درجهٔ n انتخاب و فرض می‌کنیم $B_g = P^{-1}A_gP$. تناظر $g \to B_g$ دوباره نمایشی از گروه G است، زیرا

$$P_{gh}=P^{-1}A_{gh}P=P^{-1}A_gA_hP=P^{-1}A_gPP^{-1}A_hP=B_gB_h$$

نمایشی که با این روش از نمایش مفروض و از راه انتخاب ماتریس‌های مختلف P به‌دست می‌آید، هم‌ارز نمایش مفروض نامیده می‌شود. در نظریهٔ نمایش‌ها، نمایش‌های هم‌ارز اختلاف اساسی باهم ندارند و معمول است که همهٔ آن‌ها، با تقریب هم‌ارزی، در نظر گرفته می‌شوند.

روش دیگری که برای یافتن نمایش‌های تازه به‌کار می‌رود، عبارت است از جمع مستقیم نمایش‌ها، به‌این ترتیب: فرض کنید $g \to A_g$، $g \to B_g$، دو نمایش از گروه G به‌وسیلهٔ ماتریس‌های از مرتبه‌های m و n باشند. این نگاشت را در نظر می‌گیریم:

$$g \to \begin{pmatrix} A_g & 0 \\ 0 & B_g \end{pmatrix}$$

با درنظر گرفتن قاعدهٔ ضرب ماتریس‌ها (بخش شانزدهم را ببینید) داریم:

$$gh \to \begin{pmatrix} A_{gh} & 0 \\ 0 & B_{gh} \end{pmatrix} = \begin{pmatrix} A_gA_h & 0 \\ 0 & B_gB_h \end{pmatrix} = \begin{pmatrix} A_g & 0 \\ 0 & B_g \end{pmatrix} \begin{pmatrix} A_h & 0 \\ 0 & B_h \end{pmatrix}$$

یعنی این نگاشت، دوباره نمایشی از گروه G است که آن را مجموع دو نمایش مفروض می‌گویند و به‌صورت $A_g + B_g$ نشان می‌دهند. اگر جمله‌های جمع را جابه‌جا کنیم، نمایش دیگری به‌دست می‌آید:

$$g \to \begin{pmatrix} B_g & 0 \\ 0 & A_g \end{pmatrix}$$

که البته با نمایش قبلی هم‌ارز است. بنابراین، اگر نمایش‌های هم‌ارز قابل تمیز نباشند، جمع نمایش‌ها از قانون جابه‌جایی پیروی می‌کند. به‌سادگی روشن می‌شود، با همان شرط، جمع نمایش‌ها از قانون شرکت‌پذیری هم پیروی می‌کند. اگر نمایش‌های A_g، B_g، C_g، ... از گروه G را در اختیار داشته باشیم، می‌توان از مجموع آن‌ها، نمایش‌هایی از درجه‌های بالاتر به‌دست آورد:

$$A_g+B_g+C_g\,,\ A_g+A_g+A_g+A_g\,,\ \ldots$$

برای نمونه، عددهای ۱ ، ۱− ، i ، $-i$ نسبت به‌ضرب، تشکیل یک گروه می‌دهند. اگر هر عدد این گروه را متناظر با خود این عدد قرار دهیم، به نمایش درجهٔ اول می‌رسیم. به‌عنوان نمایش دوم می‌توان نگاشت ۱→۱ ، ۱−→۱− ، $i\to -i$ ، $-i\to i$ را درنظر گرفت. مجموع‌های این نمایش‌ها، ما را به این نگاشت می‌رساند:

$$1\to\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\ -1\to\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix},\ i\to\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix},\ -i\to\begin{pmatrix}-i & 0\\ 0 & i\end{pmatrix}$$

که اگر به‌کمک ماتریس $P=\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}$ آن را تبدیل کنیم، نمایش هم‌ارز به‌دست می‌آید:

$$1\to\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\ -1\to\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix},\ i\to\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\ -i\to\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

یادآوری این مطلب جالب است که همهٔ ماتریس‌های این نمایش، حقیقی‌اند.

فرض کنیم، همهٔ ماتریس‌های یک نمایش از درجهٔ n برای گروه G، به‌این صورت باشند:

$$g\to A_g=\begin{pmatrix}B_g & C_g\\ 0 & D_g\end{pmatrix}$$

که در آن، B_g و D_g ماتریس‌های مربعی‌اند و گوشهٔ پایین و چپ A_g، به‌طور کامل با صفرها پر شده است. از ضرب ماتریس‌های A_g و A_h به‌دست می‌آید:

$$A_{gh}=A_gA_h=\begin{pmatrix} B_gB_h & B_gC_h+C_gD_h \\ \cdot & D_gD_h \end{pmatrix}$$

یعنی $B_{gh}=B_gB_h$ ، $D_{gh}=D_gD_h$. این، نشان می‌دهد که نگاشت‌های $g \to B_g$ و $g \to D_g$ هم، نمایش‌هایی از گروه G ، ولی از درجهٔ پایین‌تری هستند. نمایش A_g را، و هر نمایش هم‌ارز آن را، **ساده‌شدنی** (یا تحویل‌پذیر) می‌نامند. نمایشی را هم که با هیچ یک از این نوع نمایش‌ها هم‌ارز نباشد، **ساده‌نشدنی** (یا تحویل‌ناپذیر) گویند.

اگر در همهٔ ماتریس‌های A_g، به‌جز گوشهٔ چپ پایین، گوشه راست بالای C_g هم با صفر پر شده باشد، آن‌وقت هر نمایش هم‌ارز آن را **تجزیه‌پذیر** به مجموع نمایش‌های B_g و D_g گویند. نمایشی که هم‌ارز مجموع نمایش‌های ساده‌نشدنی باشد، کاملاً ساده‌شدنی (یا ***کاملاً ساده‌شدنی*** (یا ***تحویل‌پذیر***) نام دارد.

در نظریهٔ گروه‌ها ثابت می‌شود، هر نمایشی از گروه متناهی، کاملاً ساده شدنی است[1]. از این‌جا نتیجه می‌شود، برای یافتن همهٔ نمایش‌های گروه متناهی، کافی است نمایش‌های ساده‌نشدنی آن‌ها را در اختیار داشته باشیم، زیرا بقیهٔ نمایش‌ها، از مجموع‌های مختلف ساده‌نشدنی‌ها به‌دست می‌آیند.

در عمل، محاسبهٔ نمایش‌های ساده‌نشدنی یک گروه متناهی، اغلب مسأله‌ای دشوار است که تنها برای خانواده‌های جداگانه‌ای از گروه‌های محدود، به‌صورتی روشن، قابل حل است مثل گروه‌های جابه‌جایی‌پذیر؛ برای گروه‌های متقارن و بعضی گروه‌های دیگر، ولو اینکه از دیدگاه نظری، ویژگی‌های نمایش گروه‌های متناهی، به‌تفصیل بررسی شده است.

برای هر گروه متناهی، نمایش خاصِ «منظمی» (یا نمایش عادی) درنظر گرفته شده است که به‌صورت زیر ساخته می‌شود. g_1 ، g_2 ، ... ، g_n ، را عضوهای گروه مفروض G می‌گیریم، آن‌ها را به ردیف دل‌خواهی شماره‌گذاری و فرض می‌کنیم:

$$g_ig_k=g_{i_k}, \quad (i\,,\ k=1\,,\ 2\,,\ \dots\,,\ n)$$

با انتخاب مقدار ثابتی برای k، ماتریسی درجه n می‌سازیم که در آن، در سطر iام عدد ۱

---

۱. به‌یاد داشته باشیم، در این‌جا نمایش گروه به‌وسیلهٔ ماتریس‌ها را در نظر داریم، ماتریس‌هایی با عضوهای عددهای مختلط دل‌خواه.

را در جای i_kام و در دیگر جاها صفر گذاشته باشیم (i را برابر ۱، ۲، ...، n بگیرید) و آن را R_{g_k} می‌نامیم. تناظر $g_k \to R_{g_k}$ (k برابر ۱، ۲، ...، n) ***نمایش عادی*** (یا منظم) گروه G نامیده می‌شود.

این‌که در واقع، در این‌جا با یک نمایش سروکار داریم، بامحاسبه‌های ساده‌ای ثابت می‌شود. همچنین می‌توان ثابت کرد، با تغییر شماره‌گذاری عضوهای گروه، به‌نمایشی هم‌ارز می‌رسیم؛ بنابراین با دقت تا هم‌ارزی، هر گروه متناهی تنها یک نمایش عادی دارد.

قضیه‌های اصلی نظریهٔ نمایش گروه‌های متناهی را به‌صورتی کوتاه منظم می‌کنیم. تعداد نمایش‌های ساده‌نشدنی مختلف (ناهم‌ارز) گروه متناهی، مقداری محدود و برابر است با تعداد خانواده‌های مزدوج (بند ۶ را ببینید) عضوهای این گروه. درجهٔ نمایش‌های ساده‌نشدنی بخشیابی از مرتبهٔ گروه است؛ درضمن نمایش عادی هم‌ارز است با مجموع همهٔ نمایش‌های ناهم‌ارز و ساده‌نشدنی که در آن، هر نمایش ساده‌نشدنی به‌تعداد درجهٔ آن تکرار می‌شود.

از این‌جا بستگی جالبی بین مرتبهٔ گروه محدود و درجه‌های نمایش‌های ساده‌نشدنی نتیجه می‌شود که در این‌جا می‌آوریم.

تعداد عضوهای گروه G را n، تعداد خانواده‌های مزدوج عضوها را k و درجهٔ نمایش‌های ساده‌نشدنی را، به‌ترتیب، n_1، n_2، ...، n_k، می‌گیریم. باتوجه به‌ساختمان نمایش عادی دیده می‌شود که درجهٔ آن برابر است با n. چون به‌جز این، نمایش عادی هم‌ارز مجموع n_1 نمایش، که هم‌ارزند با نخستین نمایش ساده‌نشدنی، به‌اضافهٔ n_2 نمایش از هم‌ارزهای نمایش دوم و غیره؛ درضمن برای جمع نمایش‌ها، درجه‌های آن‌ها به‌هم اضافه می‌شود، می‌توان به این برابری رسید:

$$n = n_1^2 + n_2^2 + \cdots + n_k^2 \qquad (11)$$

اگر هر عضو گروه را متناظر با ۱ قرار دهیم، به‌نمایش ساده‌نشدنی ولی بدیهی درجه ۱ می‌رسیم که هر گروهی آن را دارد. اگر n_1 را در دستور (۱۱)، همین نمایش واحد به‌حساب آوریم، می‌توان دستور (۱۱) را به‌این صورت نوشت:

$$n = 1 + n_2^2 + \cdots + n_k^2$$

که در آن n_2، ...، n_k نمایندهٔ درجه‌های نمایش‌های ساده‌نشدنی و نابدیهی هستند.

از آن‌جا که $n_۲$، ... ، n_k بخشیاب‌های عدد n هستند، با معلوم بودن k، اغلب می‌توان از همین برابری (۱۱) مقدارهای $n_۲$، ... ، n_k را پیدا کرد. برای نمونه، گروه متقارن $S_۳$ از جای‌گشت‌های سه عضو دارای سه خانوادهٔ مزدوج جای‌گشت است: (۱)؛ (۱۲)، (۱۳)، (۲۳)؛ (۱۲۳)، (۱۳۲). به‌ازای ۶ = n و ۳ = k، برای (۱۱) تنها یک دستگاه جواب دارد:

$$۶ = ۱^۲ + ۱^۲ + ۲^۲$$

بنابراین، $S_۳$ دارای دو نمایش مختلف درجه ۱ و یک نمایش ساده‌نشدنی درجه ۲ است.

مثال دیگر را می‌توان گروه آبلی متناهی دانست. در این‌جا هر عضو خانوادهٔ جداگانه‌ای را تشکیل می‌دهد. بنابراین $k = n$، و از دستور (۱۱) نتیجه می‌شود:

$$n_۱ = n_۲ = \cdots = n_k = ۱$$

یعنی همهٔ نمایش‌های ساده‌نشدنی این گروه از درجهٔ اول‌اند و تعداد آن‌ها برابر است با مرتبهٔ گروه.

نمایش‌های ساده‌نشدنی گروه‌های آبلی را، سرشت‌های (یا مشخصه‌های) آن‌ها می‌نامند. برای یک نمایش گروه غیرآبلی، مجموعهٔ اثرهای (یعنی مجموع عضوهای قطری) ماتریس‌هایی که نمایش را تشکیل می‌دهند، سرشت (یا مشخصه) آن می‌نامند. سرشت‌های گروه‌های متناهی ویژگی‌ها و بستگی‌های جالبی دارند. بررسی نمایش‌ها و سرشت‌های گروه‌ها، نظریهٔ گروه‌ها را با نتیجه‌گیری‌های کلی و جالبی غنی کرده است که کاربرد گستردهٔ خود را در فیزیک نظری هم به‌دست آورده‌اند.

۱۰. نظریهٔ کلی گروه‌ها

پیش از این هم گفته‌ایم، در جریان به تقریب تمامی سدهٔ نوزدهم، نظریهٔ گروه بیشتر همچون نظریهٔ گروه تبدیل‌ها تکامل یافت. ولی به‌تدریج روشن شد، مطالعهٔ گروه، به‌عنوان موضوعی که مطالعهٔ گروه‌های تبدیل‌ها بتواند منجر به مطالعهٔ گروه‌های انتزاعی و زیرگروههای آن‌ها شود، جنبهٔ اساسی‌تری دارد.

عبور از نظریهٔ گروه‌های تبدیل به نظریهٔ گروه‌های انتزاعی، در آغاز در نظریهٔ گروه‌های

متناهی تحقق یافت. ولی پیشرفت تند نظریهٔ گروه‌های *لی*، همچنین نفوذ نظریهٔ گروه‌ها در توپولوژی، ضرورت ساختن نظریهٔ کلی گروه‌ها را مطرح کرد که در آن، گروه‌های متناهی تنها حالت خاصی از آن را دربر می‌گیرند.

نخستین نوشته دربارهٔ نظریهٔ گروه‌ها که در آن، به‌روشنی، این دیدگاه مطرح شد، کتاب ***اُ.یو.شمیت*** بود که در سال ۱۹۱۶ در کیف منتشر شد. **شمیت** در سال‌های ۲۰ سدهٔ بیستم هم، قضیهٔ مهمی را دربارهٔ نظریهٔ گروه‌های نامتناهی به‌دست آورد که نقطهٔ آغازی برای دیگر ریاضی‌دانان شوروی برای بررسی‌های بعدی بود. با فعالیت‌های **شمیت** و ***پ.س.آلکساندروف*** که برای ساده‌کردن اندیشه‌های جبر امروزی انجام دادند، در مسکو مکتب بزرگ نظریهٔ گروه‌ها به‌وجود آمد که یکی از شاگردان آن، ***آ.گ.کورُش*** رهبری آن را به‌عهده داشت. به‌ویژه قضیه‌ای که او ثابت کرد بسیار مشهور است. او ثابت کرد، هر زیرگروه حاصل‌ضرب آزاد، خود حاصل‌ضرب آزادی است از زیرگروه‌هایی که هر یک با زیرگروهی از یک عامل حاصل‌ضرب یکریخت هستند به اضافهٔ شاید یک زیرگروه آزاد جداگانه. بعدها نوشته‌ای دربارهٔ نظریهٔ گروه‌ها چاپ کرد که شامل همهٔ موضوع‌هایی بود که تا آن زمان دربارهٔ نظریهٔ گروه‌ها به‌دست آمده بود. این اثر در تمامی جهان شهرت پیدا کرد و به‌صورت گسترده‌ای در همه‌جا منتشر شد.

به‌دنبال مکتب جبری ریاضی‌دانان مسکو، مکتب جبری لنین‌گراد و سایر شهرهای شوروی، به‌نظریهٔ گروه‌ها پرداختند و تا اندازهٔ زیادی آن را تکامل دادند. این تلاش همچنان ادامه پیدا کرد، به‌نحوی که ریاضی‌دانان شوروی توانستند در همهٔ جنبه‌های نظریهٔ گروه‌ها، موفقیت‌های جالبی به‌دست آورند.

۱۱. عددهای فرامختلط

برای حل مسأله‌های عملی با روش جبری، اغلب به‌حالت‌های ساده‌ای از یک یا چند معادله برمی‌خوریم که از آن‌ها می‌توان مقدارهای مجهول‌ها را به‌دست آورد. مجهول‌ها عبارت‌اند از خصلت‌های کمیتی موضوع‌هایی که بررسی می‌کنیم؛ معادله‌ها هم، بستگی‌های واقعی بین این موضوع‌ها را بیان می‌کنند.

در حالت‌هایی که با مقدارهای ساده‌ای مثل جرم، حجم یا فاصله سروکار داریم و برای

مشخص‌کردن رفتار و خصلت آن‌ها، وجود عددی کافی است، به‌همین روش متوسل می‌شویم. ولی در عمل با موضوع‌هایی هم برخورد می‌کنیم که خصلت و موقعیت آن‌ها، تنها با یک عدد مشخص نمی‌شود. با پیشرفت صنعت و فن، موضوع‌ها طبیعتِ پیچیده‌تری پیدا می‌کنند که برای توضیح آن‌ها، به چند و گاهی به بی‌نهایت عدد نیاز داریم. نمونهٔ این گونه موضوع‌ها را در کمیت‌های فیزیکی، مثل نیرو، سرعت و شتاب مشاهده کرد که با جهت خود مشخص می‌شوند و برای تعیین مقدار آن، باید سه عدد را در اختیار داشته باشیم. همچنین می‌دانیم، برای مشخص بودن موضع نقطه در فضا به سه عدد، برای صفحه به سه عدد، برای خط راست به چهار عدد و برای تعیین موضع یک جسم صُلب به شش عدد نیاز داریم. به‌این ترتیب، وقتی بخواهیم به کمک جبر مسأله‌ای را حل کنیم که به موضوع‌های بغرنج مربوط می‌شود، با معادله‌هایی سروکار پیدا می‌کنیم که تعداد زیادی مجهول دارند و اغلب کار کردن با آن‌ها دشوارتر از زمانی است که مسأله را به‌طور مستقیم، و با یاری گرفتن از ویژگی‌های هندسی یا فیزیکی آن‌ها، حل کنیم. از این‌جا به‌طور طبیعی به این اندیشه رسیدند که برای مشخص کردن رفتار موضوع‌های پیچیده‌تر، به‌جای دستگاه معمولی عددها، از عددهای بغرنج‌تر و عام‌تری استفاده کنند، عددهایی که بتوان همان عمل‌های حسابی مربوط به عددهای معمولی را، دربارهٔ آن‌ها انجام داد. این طرح برای تعمیم مفهوم عدد، به همان اندازه طبیعی بود که تاریخ دانش دربارهٔ پیشرفت تدریجی مفهوم عدد، از عدد طبیعی به عدد کسری، سپس به نسبت دو عدد حقیقی (گویا یا گنگ) و سرانجام به عدد مختلط گواهی می‌دهد.

عددهای مختلط. در بخش چهارم (جلد اول) با ویژگی‌های بنیادی عددهای مختلط و ساده‌ترین کاربردهای آن‌ها آشنا شده‌ایم. در این‌جا تنها به بنیان‌گذاری مفهوم عددهای مختلط توجه می‌کنیم. به‌یاد می‌آوریم، خودِ مفهوم عدد مختلط چگونه تعریف می‌شود. در آغاز تنها عددهای حقیقی معمولی را بررسی می‌کنند و اشاره می‌کنند ریشهٔ دوم یک عدد منفی معنایی ندارد، زیرا توان دوم هر عدد حقیقی، یا مثبت است و یا صفر. سپس یادآوری می‌کنند، نیازهای دانش ریاضی‌دانان را واداشت عبارت‌های به‌صورت $a+b\sqrt{-1}$ را هم، به‌عنوان گونهٔ ویژه‌ای از عدد در بررسی‌های خود وارد کنند که به‌خاطر اختلافی که با عددهای حقیقی معمولی داشتند، نام عددهای موهومی را بر آن‌ها گذاشتند. اگر توجه کنیم، این عددهای موهومی، از همان قانون‌های عمل‌های مربوط به عددهای حقیقی پیروی

می‌کنند، آن‌وقت ریشهٔ دوم هر عدد منفی را می‌توان برحسب $i = \sqrt{-۱}$ نوشت و نتیجهٔ هر عمل حسابی را روی عددهای حقیقی و عددهای موهومی (به‌شرطی که تعداد عمل‌ها محدود باشد)، همیشه به‌صورت $a + bi$ نشان داد که در آن، a و b عددهایی حقیقی باشند.

روشن است، چنین تعریفی از عددهای موهومی تا اندازهٔ زیادی با عقل سلیم ناسازگار است. زیرا در آغاز گفته می‌شود، عبارت‌هایی مثل $\sqrt{-۱}$ ، $\sqrt{-۲}$ و غیره معنایی ندارند و سپس، این عبارت‌های بی‌معنی را عددهای موهومی می‌نامند. همین امر موجب تردیدهایی بین ریاضی‌دانان سده‌های هفدهم و هجدهم دربارهٔ به‌رسمیت شناختن عددهای مختلط و استفادهٔ از آن‌ها شد. ولی این تردیدها در آغاز سدهٔ نوزدهم ازبین رفت، زیرا نمایش هندسی عددهای مختلط به‌وسیلهٔ نقطه‌های واقع بر صفحه پیدا شد. بنیان‌گذاری خالص حسابیِ عددهای مختلط، اندکی دیرتر و به‌وسیلهٔ ***بایای*** ریاضی‌دان مجارستانی و ***هامیلتون*** ریاضی‌دان انگلیسی داده شد. طرح این بنیان‌گذاری را در این جا می‌آوریم.

به‌جای عدد $a + bi$ ، به‌طور ساده دربارهٔ زوج عددهای حقیقی (a , b) صحبت می‌کنیم. دو زوج عدد را وقتی برابر می‌نامیم که جمله‌های اول آن‌ها باهم و جمله‌های دوم آن‌ها باهم برابر باشد، یعنی $(a , b) = (c , d)$، وقتی و تنها وقتی که داشته باشیم: $a = c$ و $b = d$. جمع و ضرب زوج‌ها با دستورهای

$$(a , b) + (c , d) = (a + c , b + d);$$
$$(a , b) \cdot (c , d) = (ac - bd , ad + bc)$$

تعریف می‌شود. برای نمونه داریم:

$$(۲ , ۳) + (۱ , -۲) = (۳ , ۱) , (۲ , ۳)(۱ , -۲) = (۸ , -۱) ;$$
$$(۳ , ۰) + (۲ , ۰) = (۵ , ۰) , (۳ , ۰)(۲ , ۰) = (۶ , ۰)$$

این مثال‌ها نشان می‌دهند، در حالت خاص، وقتی با عمل‌های حسابی روی زوج‌هایی سروکار داریم که جملهٔ دوم آن‌ها برابر صفر باشد، مثل این است که عمل‌ها را تنها روی جملهٔ اول آن‌ها انجام داده باشیم. اکنون اگر نماد i را برای زوج $(۰ , ۱)$ به‌کار ببریم، به‌دست می‌آید:

$$(a , b) = a(۱ , ۰) + b(۰ , ۱) = a + bi ,$$
$$i^۲ = (۰ , ۱)(۰ , ۱) = (-۱ , ۰) = -۱$$

یعنی به نمادگذاری معمولی عددهای مختلط می‌رسیم.

به‌این ترتیب، از این دیدگاه، عددهای مختلط عبارت‌اند از زوج‌هایی از عددهای حقیقی معمولی؛ و قاعدهٔ عمل با عددهای مختلط، تنها گونهٔ ویژه‌ای از عمل روی عددهای حقیقی است.

عددهای فرامختلط. کاربرد گسترده و موفقیت‌آمیز عددهای مختلط، ریاضی‌دانان دههٔ نخست سدهٔ نوزدهم را برآن داشت تا دربارهٔ مسألهٔ عام‌تری بیندیشند: همان‌طور که عددهای مختلط با زوج‌هایی از عددهای حقیقی ساخته می‌شوند، آیا نمی‌توان، شبیه آن، عددهای مختلطی را از درجه‌های بالاتر ساخت که با سه عدد حقیقی یا بیشتر از آن، قابل بیان باشند؟ در میانه‌های سدهٔ نوزدهم، دستگاه‌های خاص مختلفی از این‌گونه عددهای مختلط درجهٔ بالا یا عددهای فرامختلط ضمن بررسی‌های ریاضی‌دانان پدید آمد، ولی در پایان سدهٔ نوزدهم و نیمهٔ اول سدهٔ بیستم بود که نظریهٔ عمومی عددهای فرامختلط تنظیم شد که در ضمن، کاربردهای مهمی در ریاضیات و فیزیک پیدا کرد.

به‌این ترتیب، عدد فرامختلط را از مرتبهٔ n می‌نامیم، که با n عدد حقیقی $(a_1, a_2, \dots, a_n)$ مشخص شود؛ درضمن، این عددهای حقیقی را مختصات عدد فرامختلط می‌خوانیم. دو عدد فرامختلط

$$(a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n)$$

را برابر می‌گیریم، وقتی که مختصات متناظر آن‌ها، باهم برابر باشند، یعنی وقتی که داشته باشیم:

$$a_1=b_1, \quad a_2=b_2, \dots, a_n=b_n$$

عمل جمع روی عددهای فرامختلط را، با این دستور طبیعی تعریف می‌کنیم:

$$(a_1, a_2, \dots, a_n)+(b_1, b_2, \dots, b_n)=(a_1+b_1, a_2+b_2, \dots, a_n+b_n)$$

که شبیه عمل جمع برای عددهای مختلط است.

به‌همین ترتیب، برای ضرب یک عدد فرامختلط در یک عدد حقیقی، بنابر تعریف می‌پذیرند:

$$a(a_1, a_2, \dots, a_n)=(aa_1, aa_2, \dots, aa_n)$$

به‌جز این، باید عمل ضرب دو عدد فرامختلط را هم تعریف کرد؛ در ضمن حاصل‌ضرب باید عددی فرامختلط باشد.

نسخه‌برداری از ضرب دو عدد مختلط، برای تعریف ضرب دو عدد فرامختلط ممکن نیست. ضرب را می‌توان با روش‌های مختلفی تعریف کرد که درضمن، دستگاه‌های مختلفی از عددهای فرامختلط به‌دست می‌آید. بنابراین، قبل از همه، باید روشن کرد از این تعریف چه هدفی دنبال می‌شود. بی‌تردید بهتر آن است که تعریف عمل‌ها روی عددهای فرامختلط، با ویژگی‌هایی که عمل‌های معمولی روی عددهای حقیقی دارند، شباهت داشته باشد. ویژگی‌های این عمل‌های معمولی روی عددهای حقیقی کدام‌اند؟

با بررسی دقیق ویژگی‌های عددها و عمل‌های روی آن‌ها، ویژگی‌هایی که اغلب در جبر استفاده می‌شوند، به‌سادگی به‌این ویژگی‌ها پی می‌بریم:

۱. برای هر دو عدد، مجموع آن‌ها منحصر به فرد (یکتا) است.

۲. برای هر دو عدد، حاصل‌ضرب آن‌ها منحصر به فرد است.

۳. عدد صفر با ویژگی $a + 0 = a$ برای هر عدد a وجود دارد.

۴. برای هر عدد a، عدد متقابل (یا قرینهٔ) x وجود دارد، به‌نحوی که در برابری $a + x = 0$ صدق می‌کند.

۵. جمع از قانون جابه‌جایی پیروی می‌کند:

$$a + b = b + a$$

۶. جمع از قانون شرکت‌پذیری پیروی می‌کند:

$$(a + b) + c = a + (b + c)$$

۷. ضرب از قانون جابه‌جایی پیروی می‌کند:

$$ab = ba$$

۸. ضرب از قانون شرکت‌پذیری پیروی می‌کند:

$$(ab)c = a(bc)$$

۹. ضرب از قانون پخشی نسبت به‌جمع پیروی می‌کند:

$$a(b + c) = ab + ac\ ,\ \ (b + c)a = ba + ca$$

۱۰. برای هر a و هر $b \neq 0$، عدد x وجود دارد که در برابری $bx = a$ صدق می‌کند.

ویژگی‌های ۱ تا ۱۰ در نتیجهٔ تجزیه و تحلیلی دقیق جدا شده‌اند و پیشرفت ریاضیات در سدهٔ اخیر، اهمیت آن‌ها را نشان داد. در زمان ما ، هر دستگاهی از مقدارها که با ویژگی‌های

۱ تا ۱۰ سازگار باشد، **هیأت** نامیده می‌شود. ازجمله، مجموعهٔ همهٔ عددهای گویا، مجموعهٔ همهٔ عددهای حقیقی، مجموعهٔ همهٔ عددهای مختلط، هرکدام یک هیأت‌اند، زیرا عددها را در هر کدام از این مجموعه‌ها می‌توان باهم جمع و درهم ضرب کرد؛ درضمن این عمل‌ها با ویژگی‌های ۱ تا ۱۰ سازگارند. به‌جز این سه هیأت مهم‌تر، می‌توان از تعداد بی‌شماری هیأت‌های دیگر نام برد که از عددها تشکیل شده باشند. ولی در کنار هیأت‌های عددی، هیأت‌های دیگری هم وجود دارند که شامل مقدارهای با طبیعت دیگری هستند. در دبیرستان با کسرهای جبری آشنا شده‌ایم، یعنی کسرهایی که صورت و مخرج هریک از آن‌ها، یک چندجمله‌ای نسبت به برخی حرف‌ها باشد. کسرهای جبری را می‌توان باهم جمع، از هم کم، درهم ضرب و برهم تقسیم کرد؛ درضمن این عمل‌ها، از ویژگی‌های ۱ تا ۱۰ پیروی می‌کنند. بنابراین کسرهای جبری، دستگاهی از موضوع‌ها را تشکیل می‌دهند که یک هیأت است. می‌توان مثال‌های دیگری برای هیأت آورد که شامل کمیت‌هایی با طبیعت‌هایی بغرنج و بغرنج‌تر باشند. با توجه به اهمیتی که ویژگی‌های ۱ تا ۱۰ برای تعیین هیأت دارند، باید مسالهٔ عمل ضرب را دربارهٔ عددهای فرامختلط به‌نحوی طرح ریخت که عددهای فرامختلط تشکیل یک هیأت بدهند. در چنین صورتی، می‌توانیم به عددهای مختلط کلی‌تر و تازه‌ای دست یابیم. ولی در همان آغاز سدهٔ نوزدهم معلوم شد، این امکان، تنها برای عددهای فرامختلط مرتبهٔ ۲ وجود دارد که در واقع همان عددهای مختلط عادی‌اند. این نتیجه‌گیری نشان داد، عددهای مختلط، مقام ویژه‌ای دارند و گسترش دستگاه عددی به بیرون از مرز عددهای مختلط، به‌شرطی که بخواهیم همهٔ ویژگی‌های ۱ تا ۱۰ برقرار باشند، ممکن نیست. بنابراین، برای تلاش‌های بعدی در ساختن دستگاه‌های عددی از مرتبهٔ بالاتر، باید از یک یا چند ویژگی ۱ تا ۱۰ صرف‌نظر کرد.

چهارگان‌ها (*quaternions*). از نظر تاریخی، نخستین دستگاه فرامختلطی که وارد ریاضیات شد، دستگاه چهارگانی، یعنی «چهارعددی» بود که به‌وسیلهٔ ***هامیلتون***، ریاضی‌دان و مکانیک‌شناس انگلیسی در میانهٔ سدهٔ نوزدهم مطرح شد. این دستگاه با همهٔ ویژگی‌های ۱ تا ۱۰ ، به‌جز ویژگی ۷ (قانون جابه‌جایی در ضرب) سازگار است.

چهارگان‌ها به‌این صورت شرح داده می‌شوند. برای چهارتایی‌های (۰ , ۰ , ۰ , ۱) ، (۰ , ۱ , ۰ , ۰) ، (۰ , ۰ , ۱ , ۰) و (۱ , ۰ , ۰ , ۰) نمادهای ۱ ، i ، j ، k را در نظر می‌گیریم.

در این صورت با توجه به‌برابری

$$(a , b , c , d) = a(1 , 0 , 0 , 0) + b(0 , 1 , 0 , 0) + c(0 , 0 , 1 , 0) + d(0 , 0 , 0 , 1)$$

هر چهارگان را می‌توان، به‌طور یک‌ارزشی، به‌صورت

$$(a , b , c , d) = a \cdot 1 + b \cdot i + c \cdot j + d \cdot k$$

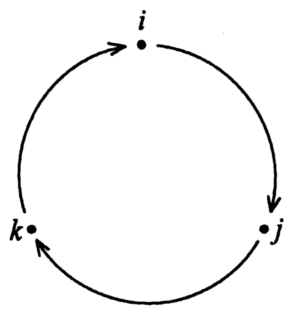
نشان داد. چهارگان ۱ را واحد دستگاه مقدارها می‌گیریم، یعنی برای هر چهارگان α فرض می‌کنیم:

$$1 \cdot \alpha = \alpha \cdot 1 = \alpha$$

سپس، بنابر تعریف فرض می‌کنیم:

$$\begin{aligned} i^2 = j^2 = k^2 &= -1 , \\ ij = -ji &= k , \\ ik = -ki &= -j , \\ jk = -kj &= i \end{aligned}$$

این «جدول‌ضرب» چهارگان‌ها به‌سادگی و به یاری شکل ۲۸ قابل درک است. در این شکل نقطه‌های i ، j ، k روی محیط دایره معرف چهارگان متوالی i ، j ، k هستند. حاصل‌ضرب دو جملهٔ مجاور چهارگان برابر است با سومی به شرطی که حرکت از جملهٔ اول به جملهٔ دوم،



شکل ۲۸

طبق شکل، در جهت حرکت عقربه‌های ساعت باشد و برابر سومی با علامت منفی است، به شرطی که جهت حرکت برخلاف حرکت عقربه‌های ساعت باشد. با آگاهی از این «جدول‌ضرب»، می‌توان حاصل‌ضرب هر دو چهارگان دل‌خواه را با استفاده از اصل موضوع پخشی ۹ به‌دست آورد. به‌این ترتیب.

$$
\begin{aligned}
&(a\cdot 1 + b\cdot i + c\cdot j + d\cdot k)\,(a_1\cdot 1 + b_1\cdot i + c_1\cdot j + d_1\cdot k) = \\
&= aa_1\cdot 1 + ab_1\cdot i + ac_1\cdot j + ad_1\cdot k + ba_1\cdot i + bb_1\cdot ii + bc_1\cdot ij + \\
&+ bd_1\cdot ik + ca_1\cdot j + cb_1\cdot ji + cc_1\cdot jj + cd_1\cdot jk + da_1\cdot k + db_1\cdot ki + \\
&+ dc_1\cdot kj + dd_1\cdot kk = \\
&= (aa_1 - bb_1 - cc_1 - dd_1)\cdot 1 + (ab_1 + ba_1 + cd_1 - dc_1)\, i + \\
&+ (ac_1 + ca_1 - bd_1 + db_1)\cdot j + (ad_1 + da_1 + bc_1 - cb_1)\cdot k
\end{aligned}
$$

عامل ۱ را در نخستین جملهٔ چهارگان، به‌طور معمول، حذف می‌کنند و به‌جای «$a.1$» می‌نویسند «a». برابری‌های $ij = -ji$، $ik = -ki$، و $jk = -kj$ نشان می‌دهند که ضرب چهارگان‌ها، از قانون جابه‌جایی پیروی نمی‌کند. در این‌جا، عامل‌های ضرب هم‌ارز و دارای یک نقش نیستند. بنابراین، وقتی با ضرب چهارگان‌ها سروکار داریم، باید ردیف عامل‌های ضرب را رعایت کرد. دربارهٔ عمل‌های دیگر چهارگان‌ها، دشواری خاصی وجود ندارد. قانون شرکت‌پذیری ۸ در ضرب چهارگان‌ها برقرار است و به‌سادگی می‌توان آن را دربارهٔ چهارگان‌های پایهٔ ۱، i، j، k و به یاری «جدول‌ضرب» آزمایش کرد؛ عبور به‌حالت کلی هم روشن است.

در چهارگان $a + bi + cj + dk$، عدد a را بخش حقیقی یا بخش اسکالر آن و مجموع $bi + cj + dk$ را بخش بُرداری آن گویند. دو چهارگان

$$a + bi + cj + dk \text{ و } a - bi - cj - dk$$

را که تنها در علامت بخش برداری باهم فرق دارند، مزدوج هم گویند. روشن است که مجموع دو چهارگان مزدوج یکدیگر، برابر است با عددی حقیقی. از این گذشته، برای حاصل‌ضرب دو چهارگان مزدوج، اگر بنابر دستور ضرب عمل کنیم، چنین است:

$$(a+bi+cj+dk)(a-bi-cj-dk) = a^2+b^2+c^2+d^2 \qquad (12)$$

یعنی حاصل‌ضرب دو چهارگان مزدوج هم، عددی حقیقی است.

مجموع توان‌های دوم ضریب‌های چهارگان $a+bi+cj+dk$، یعنی $a^۲+b^۲+c^۲+d^۲$ را هنج (یا نُرم *norm*) چهارگان می‌نامند. از آن‌جا که توان دوم هر عدد حقیقی، عددی نامنفی است، بنابراین هنج هر چهارگان هم، عددی نامنفی است و تنها برای چهارگان صفر، برابر صفر می‌شود.

دستور (۱۲) نشان می‌دهد، حاصل‌ضرب هر چهارگان در چهارگان مزدوج خود، برابر هنج چهارگان مفروض است.

اگر هر چهارگان مزدوج با چهارگان مفروض را با علامت ستاره مشخص کنیم، آن‌وقت با ضرب مستقیم، درستی دستور

$$(\alpha\beta)^*=\beta^*\alpha^*$$

ثابت می‌شود. از این‌جا به نتیجهٔ جالبی می‌رسیم: هنج حاصل‌ضرب چهارگان‌ها، برابر است با حاصل‌ضرب هنج‌های عامل‌های ضرب. در واقع داریم:

$$\text{هنج }(\alpha\beta)=(\alpha\beta)(\alpha\beta)^*=\alpha\beta\beta^*\alpha^*=(\alpha\alpha^*)(\beta\beta^*)=(\text{هنج }\alpha).(\text{هنج }\beta)$$

ویژگی‌های هنج‌ها، امکان ساده‌ای برای حل مسألهٔ مربوط به تقسیم چهارگان‌ها پدید می‌آورد. فرض کنید $\alpha=a+bi+cj+dk$، چهارگانی مخالف صفر باشد. در این صورت

$$(a+bi+cj+dk)\frac{۱}{a^۲+b^۲+c^۲+d^۲}(a-bi-cj-dk)=$$

$$=\frac{۱}{a^۲+b^۲+c^۲+d^۲}(a^۲+b^۲+c^۲+d^۲)=۱$$

یعنی چهارگان

$$\frac{۱}{a^۲+b^۲+c^۲+d^۲}(a-bi-cj-dk)=\alpha^{-۱}$$

وارون چهارگانِ مفروض α است.

با امکان به‌دست آوردن وارون یک چهارگان، به‌سادگی می‌توان خارج قسمت دو چهارگان را به‌دست آورد. دو چهارگان α و β را در نظر می‌گیریم و فرض می‌کنیم α مخالف صفر باشد. در این صورت، خارج قسمت‌های حاصل از تقسیم β بر α، با حل این معادله‌ها

به‌دست می‌آیند:

$$\alpha x = \beta \ , \ y\alpha = \beta$$

اگر دو طرف معادلهٔ اول را از سمت چپ در α^{-1} ضرب کنیم، به‌دست می‌آید:

$$x = \alpha^{-1}\beta$$

و اگر دو طرف معادلهٔ دوم را از سمت راست در α^{-1} ضرب کنیم، آن‌وقت

$$y = \beta\alpha^{-1}$$

چون حاصل‌ضرب‌های $\alpha^{-1}\beta$ و $\beta\alpha^{-1}$، در حالت کلی، باهم فرق دارند، برای چهارگان‌ها باید دو تقسیم در نظر گرفت: راست و چپ. این تقسیم‌ها، هردو جواب دارند، البته به‌جز حالت تقسیم بر صفر.

جبر بُردارها. با این که عمل با چهارگان‌ها تا اندازهٔ زیادی به عمل با عددهای مختلط شباهت دارد، از آن‌جا که در ضرب از قانون جابه‌جایی پیروی نمی‌کند، ویژگی‌های چهارگان‌ها به‌کلی غیر از ویژگی‌های عدد از آب درمی‌آید. برای نمونه، وقتی با جبر عددهای مختلط سروکار داریم، هر معادلهٔ درجه دوم دارای دو ریشه است. ولی فرض کنید بخواهیم معادلهٔ درجه دوم

$$x^2 + 1 = 0$$

در حوزهٔ چهارگان‌ها حل کنیم، در این صورت ۶ ریشه برای آن به‌دست می‌آید: $\pm i$، $\pm j$، $\pm k$، و اگر تجزیه و تحلیل دقیق‌تری انجام دهیم، متوجه می‌شویم، مجموعهٔ جواب‌های این معادله، مجموعه‌ای بسیار بزرگ است. این وضع، استفادهٔ از چهارگان‌ها را بسیار دشوار می‌کند و با وجود تلاش‌های *هامیلتون* و دیگر ریاضی‌دانان در وارد کردن چهارگان‌ها در ریاضیات و فیزیک، نقش آن‌ها در ریاضیات بسیار کم است و به‌هیچ وجه قابل مقایسه با نقش عددهای مختلط نیست.

با وجود این، چهارگان‌ها انگیزه‌ای برای پیشرفت جبر برداری شد که ابزاری بسیار ضروری در صنعت و فیزیک امروز است. مطلب بر سر این است که، در مکانیک و فیزیک، نقش اصلی به‌عهدهٔ مفهوم‌های سرعت، شتاب، نیرو و غیره است که برای تعیین رفتار آن‌ها به سه عدد نیاز داریم. دیدیم هر چهارگان را می‌توان به‌عنوان مجموعه‌ای از عدد حقیقی a و

بخش برداری $bi + cj + dk$ درنظر گرفت. از آن‌جاکه بخش برداری چهارگان با سه عدد معین می‌شود، برای تعیین رفتار مهم‌ترین کمیت‌های فیزیکی، کافی است تنها بخش برداری چهارگان‌ها را در نظر بگیریم.

بخش برداری $bi + cj + dk$ از چهارگان $a + bi + cj + dk$، از نظر هندسی، معرف برداری است که در دستگاه قائم محورهای مختصات، از مبداء آغاز شده است و روی محورهای مختصات تصویرهایی، به‌ترتیب برابر عددهای b ، c و d دارد. بنابراین، هر چهارگان را از نظر هندسی می‌توان مجموعه‌ای شامل یک عدد حقیقی و یک بردار در فضای سه‌بُعدی دانست. ببینیم با این تفسیر، عمل‌های روی چهارگان‌ها به‌چه صورتی درمی‌آیند.

دو چهارگان برداری را، که بخش اسکالر آن‌ها برابر صفر است، در نظر می‌گیریم:

$$xi+yj+zk\ ,\ \ x_1 i+y_1 j+z_1 k$$

با دید هندسی، با دو بردار سروکار داریم که از مبداء مختصات آغاز شده‌اند. مجموع این چهارگان‌ها، دوباره یک چهارگان برداری است:

$$(x+x_1)i+(y+y_1)j+(z+z_1)k$$

به‌سادگی دیده می‌شود، نمایش این مجموع، قطر متوازی‌الاضلاعی است که روی دو بردار اول ساخته شده باشد. بنابراین، جمع چهارگان‌های برداری، به‌خوبی پاسخ‌گوی عمل مشهور و جمع بردارها طبق قاعدهٔ متوازی‌الاضلاع است. به‌همین ترتیب، اگر چهارگان برداری را در عددی حقیقی ضرب کنیم، آن‌وقت چهارگان برداری حاصل هم در این عدد ضرب می‌شود.

در حالت ضرب چهارگان‌ها با موقعیت دیگری روبه‌رو می‌شویم. در واقع

$$(xi+yj+zk)(x_1 i+y_1 j+z_1 k)=-xx_1-yy_1-zz_1+(yz_1-y_1 z)i+(zx_1-z_1 x)j+(xy_1-x_1 y)k$$

یعنی برای حاصل‌ضرب دو بردار چهارگانی، یک چهارگان کامل به‌دست می‌آید که دارای بخش اسکالر و بخش برداری است.

بخش اسکالر حاصل‌ضرب چهارگان‌های برداری، به‌شرطی که علامت آن را عوض کنیم، حاصل‌ضرب اسکالر بردارها و بخش برداری آن حاصل‌ضرب برداری بردارهای مفروض نامیده می‌شود. حاصل‌ضرب اسکالر بردارهای α و β به‌صورت $(\alpha\beta)$ یا به‌طور

ساده $\alpha\beta$ و حاصل‌ضرب برداری آن‌ها به‌صورت $[\alpha\beta]$ نشان داده می‌شود. $\boldsymbol{i}$، $\boldsymbol{j}$، $\boldsymbol{k}$ را بردارهای واحد متناظر با چهارگان‌های i، j، k می‌گیریم، یعنی بردارهای واحدی که به‌ترتیب روی محورهای مختصات انتخاب شده‌اند. بنابر تعریف، اگر داشته باشیم:

$$\alpha = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} \quad \text{و} \quad \beta = x_1\boldsymbol{i} + y_1\boldsymbol{j} + z_1\boldsymbol{k}$$

آن‌وقت

$$(\alpha\beta) = xx_1 + yy_1 + zz_1$$

$$[\alpha\beta] = (yz_1 - y_1z)\boldsymbol{i} + (zx_1 - z_1x)\boldsymbol{j} + (xy_1 - x_1y)\boldsymbol{k}$$

به‌یاری این دستورها، تعبیر هندسی حاصل‌ضرب‌های اسکالر و برداری داده می‌شود: حاصل‌ضرب اسکالر دو بردار برابر است با حاصل‌ضرب طول‌های آن‌ها در کسینوس زاویه‌ای که باهم می‌سازند؛ حاصل‌ضرب برداری دو بردار عبارت است از برداری که طول آن برابر با عدد مساحت متوازی‌الاضلاعی است که روی این دو بردار ساخته شود و در امتداد عمود بر صفحهٔ همین متوازی‌الاضلاع قرار گیرد. جهت این بردار به‌گونه‌ای است که از آن‌جا، دوران از بردار مفروض اول به بردار دوم، در همان جهتی دیده شود که با قرار گرفتن روی محور Oz، دوران از محور Ox به محور Oy در آن جهت انجام بگیرد.

در مکانیک و فیزیک از عمل روی چهارگان‌ها استفاده نمی‌شود و، به‌جای آن، تنها از عمل روی بردارها، آن‌هم، روش خالص هندسی و به‌همان صورتی که هم‌اکنون شرح دادیم، صحبت می‌کنند.

در پایان از یک مسألهٔ مربوط به‌مکانیک نام می‌بریم که به‌ویژه، راه حل آن به‌کمک چهارگان‌ها، بسیار زیبا است. به‌خصوص حل همین مسأله، انگیزه‌ای برای کشف چهارگان‌ها شد.

جسم سختی را در نظر بگیرید که در آغاز به اندازهٔ زاویه‌ای مثل φ، در جهتی مفروض دور محور معینی می‌چرخد. OA که از نقطهٔ مفروض O می‌گذارد، پس از آن، دور محور OB که از همان نقطهٔ O می‌گذرد، به‌اندازهٔ زاویه‌ای مثل φ_1 دوران می‌کند. پرسش این است: دور چه محوری و با چه زاویه‌ای باید جسم را پیچاند تا یکباره از موقعیت نخست به موقعیت سوم برسد؟ این، مسألهٔ شناخته شده‌ای از مکانیک است که به جمع دوران‌های محدود، مربوط می‌شود. البته، این مسأله را می‌توان با روش عادی و به یاری هندسهٔ تحلیلی حل کرد،

کاری که *اولر* در سدهٔ هجدهم انجام داد: ولی اگر از چهارگان‌ها استفاده کنیم، به راه حلی روشن‌تر و زیباتر می‌رسیم.

$\xi = xi + yj + zk$ و $\alpha = a + bi + cj + dk$ دو چهارگان می‌کنیم که از آن‌ها، اولی را متغیر و دومی را مفروض فرض می‌کنیم. به‌سادگی و با محاسبه روشن می‌شود که عبارت $\alpha^{-1}\xi\alpha$، یک چهارگان برداری است. اکنون اگر چهارگان‌های ξ، $\alpha^{-1}\xi\alpha$ و بخش برداری چهارگان α را با $\vec{\xi}$، $\vec{\xi}_1$ و $\vec{\alpha}$ نشان دهیم، آن‌وقت معلوم می‌شود، از نظر هندسی می‌توان بردار $\vec{\xi}$ را از بردار $\vec{\xi}$ با دوران دور محوری که از بردار $\vec{\alpha}$ می‌گذرد، به اندازهٔ زاویه φ که با دستور

$$\cos\frac{\varphi}{2} = \frac{a}{\sqrt{a^2+b^2+c^2+d^2}}$$

پیدا می‌شود، به‌دست آورد. بنابراین چهارگان $\alpha = a + bi + cj + dk$ را می‌توان معرف دوران فضا به‌اندازهٔ زاویهٔ φ دور محور $\vec{\alpha} = bi + cj + dk$ دانست.

برعکس، با معلوم بودن محور دوران و زاویهٔ φ، می‌توان چهارگانی را جست‌وجو کرد که معرف این دوران باشد. این‌گونه چهارگان‌ها، مجموعه‌ای بی‌پایان را تشکیل می‌دهد، ولی اختلاف همهٔ آن‌ها باهم، تنها در ضرب عددی است.

اکنون دوران دیگری را به اندازهٔ زاویهٔ φ_1 و دور محوری مثل

$$\vec{\beta} = b_1 i + c_1 j + d_1 k$$

درنظر می‌گیریم. فرض کنید چهارگان $\beta = a_1 + b_1 i + c_1 j + d_1 k$ معرف این دوران باشد. بردار دل‌خواه $\vec{\xi} = xi + yj + zk$ زیر تأثیر دوران اول به بردار $\alpha^{-1}\vec{\xi}\alpha$ و این بردار، زیر تأثیر دوران دوم، به‌بردار $\beta^{-1}(\alpha^{-1}\vec{\xi}\alpha)\beta$ می‌رسد. بنابر قانون شرکت‌پذیری، این نتیجه را می‌توان با دستور

$$\beta^{-1}(\alpha^{-1}\vec{\xi}\alpha)\beta = (\alpha\beta)^{-1}\,\xi\alpha\beta$$

نشان داد. از آن‌جا که ضرب بردارها، یعنی چهارگان برداری $\vec{\xi}$ در چهارگان $(\alpha\beta)^{-1}$ از چپ، و چهارگان $\alpha\beta$ از راست، هم‌ارز دوران این بردار به زاویهٔ متناظر و دور محور متناظر است، به‌این نتیجه می‌رسیم که، دو دوران پشت سرهم، که با چهارگان‌های α و β مشخص می‌شوند، عبارت است از دورانی که حاصل‌ضرب $\alpha\beta$ معرف آن است. به‌زبان دیگر، جمع

دوران‌ها پاسخ‌گوی ضرب چهارگان‌های متناظر آن‌هاست.

چهارگان‌ها، به‌جز کاربردهایی که در هندسه و فیزیک دارند، در نظریهٔ عددها هم کاربردهای جالبی پیدا کرده‌اند. از کسانی که به‌ویژه در این زمینه کار کرده‌اند، باید از **یو.و.لی‌نیک** نام برد.

۱۲. جبرهای شرکت‌پذیر

تعریف کلی جبر (دستگاه‌های فرامختلط). عددهای فرامختلط به‌عنوان کمیت‌هایی تعریف می‌شدند که برای معرفی هریک از آن‌ها، چند عدد حقیقی لازم است؛ درضمن برای مشخص بودن عددهای فرامختلط را، به‌طور ساده، همچون دستگاهی از عددهای حقیقی در نظر می‌گرفتند. ولی این دیدگاه بی‌اندازه تنگ و محدود بود و به‌تدریج احساس شد، برای بررسی‌های نظری، به‌تعریفی کلی‌تر نیاز است؛ چنین تعریفی پدید آمد.

دستگاه S از کمیت‌ها را یک جبر (یا دستگاه فرامختلط) روی هیأت P گویند، وقتی که

الف) برای‌هرعضو a از هیأت P و هر کمیت α از دستگاه S، عضوی از این دستگاه به‌نام حاصل‌ضرب a در α معین باشد که آن را به‌صورت $a\alpha$ نشان می‌دهیم؛

ب) برای هر دو عضو α و β از دستگاه، مقداری از همین دستگاه به‌طور یکتا (منحصر به‌فرد) تعریف شود که مجموع دو مقدار اول نامیده و به‌صورت $\alpha+\beta$ معرفی می‌شود؛

ج) برای هر دو عضو α و β از دستگاه، مقداری از همین دستگاه به‌نام حاصل‌ضرب دو مقدار اول و به‌طور یکتا تعریف شده باشد که با $\alpha\beta$ نشان داده می‌شود.

و در ضمن، وقتی که این سه عمل دارای ویژگی‌های زیر باشند[1]:

(۱′ $\alpha+\beta=\beta+\alpha$،

(۲′ $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$،

(۳′ در دستگاه S، مقدار صفر θ وجود دارد با ویژگی $\alpha+\theta=\alpha$،

(۴′ $a(\alpha+\beta)=a\alpha+a\beta$،

---

۱. حرف‌های یونانی معرف کمیت‌های دل‌خواهی از دستگاه S و حرف‌های الفبای لاتینی معرف عضوهایی از هیأت P هستند.

۵′) $(a+b)\alpha = a\alpha + b\alpha$ ،

۶′) $(ab)\alpha = a(b\alpha)$ ،

۷′) $\theta\alpha = \theta$ ، $1.\alpha = \alpha$ (۱ ، عضو یکه یا واحد هیأت P است)،

۸′) بین مقدارهای دستگاه S، چنان مقدارهایی مثل α_1 ، α_2 ، ... ، α_n ، وجود دارد که هر مقدار دستگاه S را بتوان، به‌طور یکتا، برحسب آن‌ها، به‌این صورت نشان داد:

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n،$$

۹′) $(a\alpha)\beta = \alpha(a\beta) = a(\alpha\beta)$ ،

۱۰′) $(\beta+\gamma)\alpha = \beta\alpha + \gamma\alpha$ ، $\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma$.

نقشی که تاکنون برعهدهٔ عددهای حقیقی بود، در این تعریف، برعهدهٔ عضوهای هیأت دل‌خواه P است. در شرط ′۸ دیده می‌شود، هر مقدار فرامختلط با دستگاهی شامل n عضو a_1 ، a_2 ، ... ، a_n از هیأت P تعریف می شود که، با توجه به انتخاب هیأت P ، می‌تواند به‌وسیلهٔ n عدد مختلط، n عدد گویا، n عدد حقیقی یا هرگونهٔ دیگری از عددها، معین شود.

هشت شرطِ نخستین به‌این معنی است که S ، فضایی خطی با بُعد است (بند ۲ از بخش شانزدهم را ببینید)؛ این دستگاه روی هیأت P که هیأت ضریب‌های جبر نامیده می‌شود، بنا شده است.

شرط‌های ′۹ و ′۱۰ را می‌توان با برابری‌های

$$(a\beta + b\gamma)\,\alpha = a(\beta\alpha) + b\,(\gamma\alpha)\,,$$
$$\alpha(a\beta + b\gamma) = a(\alpha\beta) + b(\alpha\gamma)\,,$$

به‌هم پیوست که از آن‌ها نتیجه می‌شود: عمل ضرب، عملی خطی نسبت به‌هریک از عامل‌های ضرب است.

در سال‌های اخیر، از دو اصطلاح «دستگاه فرامختلط» و «جبر»، اصطلاح دوم را ترجیح می‌دهند، زیرا «دستگاه‌های فرامختلط» گاه شامل چنان عضوهایی هستند که، از نظر ویژگی‌های خود، به‌کلی از عددهای عادی دورند و اصطلاح «فرامختلط» برای آن‌ها نامناسب است. اصطلاح‌های «دستگاه‌های فرامختلط» و «عددهای فرامختلط» را امروز، تنها دربارهٔ ساده‌ترین حالت جبر، دستگاه چهارگان‌های معمولی به‌کار می‌برند.

باتوجه به شرط‌های ′۱ تا ′۱۰ روشن می‌شود که در جبر، قانون‌های جابه‌جایی و

شرکت‌پذیری ضرب را در نظر نمی‌گیرند؛ همچنین از وجود عضو یکه و عمل «تقسیم» صحبتی نشده است.

در هر جبر S پایه‌ای وجود دارد، یعنی دستگاهی از عضوهای α_1، α_2، ...، α_n که همهٔ عضوهای جبر را می‌توان برحسب آن‌ها، به‌طور یکتا به‌صورت ترکیب خطیِ

$$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$$

با ضریب‌هایی از هیأت ضریب‌های P نشان داد. هر جبر ممکن است دارای بی‌نهایت پایه باشد، ولی تعداد عضوهای هر پایه برای همهٔ آن‌ها یکی است و رتبهٔ جبر نامیده می‌شود.

دستگاه عددهای مختلط را به‌عنوان جبری که روی هیأت عددهای حقیقی ساخته شده است در نظر می‌گیریم و پایهٔ آن عددهای ۱ و i است. ولی زوج عددهای ۲ و $3i$ یا ۱ و $a+bi$ (a و b عددهایی حقیقی‌اند و $b \neq 0$) هم می‌توانند به‌عنوان پایه درنظر گرفته شوند.

ε_1، ε_2، ...، ε_n را پایهٔ یک جبر روی هیأتی مثل P می‌گیریم. بنابر تعریف، هر عضو جبر به‌طور یکتا به‌صورت

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n$$

نوشته می‌شود. اگر $\beta = b_1\varepsilon_1 + \ldots + b_n\varepsilon_n$ عضو دیگری از آن باشد، آن‌وقت باتوجه به ویژگی‌های $1'$ تا $6'$ داریم:

$$\alpha + \beta = (a_1 + b_1)\varepsilon_1 + (a_2 + b_2)\varepsilon_2 + \cdots + (a_n + b_n)\varepsilon_n$$

به‌همین ترتیب برای هر a از هیأت P:

$$a\alpha = aa_1\varepsilon_1 + aa_2\varepsilon_2 + \cdots + aa_n\varepsilon_n$$

بنابراین، عمل جمع مقدارهای جبر و عمل ضرب این مقدارها در عضوی از هیأت P، عمل‌هایی یک‌ارزشی‌اند، برای ضرب مقدارهای جبر باید هر بار، بنابر همان حالت عمل کرد؛ درضمن لازم نیست بدانیم مقدارهای دل‌خواه و غیرمشخص جبر چگونه درهم ضرب می‌شوند، بلکه کافی است تنها از قانون ضرب مقدارهای پایه‌ای ε_i اطلاع داشته باشیم. در واقع، باتوجه به‌ویژگی‌های $9'$ و $10'$

$$(a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n)(b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n) = \sum a_i b_j \cdot \varepsilon_i \varepsilon_j$$

هر حاصل‌ضرب $\varepsilon_i \varepsilon_j$، مقداری از جبر است و بنابراین می‌توان آن را برحسب عضوهای پایه نشان داد:

$$\varepsilon_i \, \varepsilon_j = c_{ij۱} \, \varepsilon_۱ + c_{ij۲} \, \varepsilon_۲ + \cdots + c_{ijn} \, \varepsilon_n$$

در این‌جا c_{ijk} به‌معنای عضوی از هیأت ضریب‌های P است که جبر روی آن ساخته شده است. اندیس اول به‌معنای شمارهٔ عامل اول ضرب، اندیس دوم به‌معنای شمارهٔ عامل دوم ضرب و اندیس سوم معرف شمارهٔ عضوی است که ضریب آن c_{ijk} است. ضریب‌های c_{ijk} را ثابت‌های ساختاری جبر گویند، زیرا با دانستن آن‌ها، همهٔ عمل‌های روی مقدارهای جبر، به‌طور کامل معین می‌شود.

تعداد ثابت‌های ساختاری جبر با رتبهٔ n را به‌سادگی می‌توان محاسبه کرد. هر ثابت، همراه با سه شمارهٔ i، j و k است. بنابراین، تعداد ثابت‌های ساختاری در جبر با رتبهٔ n، برابر است با تعداد سه‌تایی‌هایی که می‌توان با عددهای طبیعی ۱ ، ۲ ، ... ، n به‌دست آورد، یعنی برابر است با $n^۳$. برای نمونه، دستگاه عددهای مختلط روی میدان عددهای حقیقی، پایه‌ای دارد که شامل عددهای ۱ و i است. باتوجه به‌برابری‌های

$$۱.۱ = ۱.۱ + ۰.i \qquad , \qquad i.۱ = ۰.۱ + ۱.i \, ,$$

$$۱.i = ۰.۱ + ۱.i \qquad , \qquad i.i = -۱.۱ + ۰.i \, ,$$

ثابت‌های ساختاری متناظر، عبارت‌اند از:

$$c_{۱۱۱} = ۱ \, , \; c_{۱۱۲} = ۰ \, , \; c_{۲۱۱} = ۰ \, , \; c_{۲۱۲} = ۱ \, ,$$

$$c_{۱۲۱} = ۰ \, , \; c_{۱۲۲} = ۱ \, , \; c_{۲۲۱} = -۱ \, , \; c_{۲۲۲} = ۰$$

برعکس، فرض کنید $n^۳$ عضو از هیأتی مثل P داده شده باشد که به‌وسیلهٔ سه‌تایی‌هایی از عددهای طبیعی i، j و k شماره‌گذاری شده باشند (هر یک از عدهای i، j و k می‌تواند یکی از عددهای طبیعی ۱ تا n را اختیار کند). در این صورت می‌توان آن‌ها را به‌عنوان ثابت‌های ساختاری یک جبر روی هیأت P در نظر گرفت و برابری $\varepsilon_i \, \varepsilon_j = \sum_{k=۱}^{n} c_{ijk} \, \varepsilon_k$ را به‌عنوان تعریف قانون ضرب در این جبر پذیرفت.

دیدیم، هر جبری در حالت کلی، دارای مجموعهٔ نامتناهی از پایه‌های مختلف است.

ثابت‌های ساختاری بستگی به انتخاب نوع پایه دارد، بنابراین یک جبر مشخص، با دستگاه‌های مختلفی از ثابت‌های ساختاری معین می‌شود.

پس کدام جبرها باهم اختلاف دارند و کدام جبرها یکی هستند؟ در نظریهٔ جبر، دو جبر را روی یک هیأت P یکی می‌دانند اگر آن‌ها یکریخت (ایزومورف) باشند، یعنی به‌شرطی که کمیت‌های یک جبر را بتوان در تناظر یک‌به‌یک با کمیت‌های دیگری قرار داد و به‌شرطی که مجموع و حاصل‌ضرب دو کمیت دل‌خواه از جبر اول در تناظر یک‌به‌یک با مجموع و حاصل‌ضرب کمیت‌های متناظر جبر دوم باشد، همچنین حاصل‌ضرب عضوی از هیأت P در کمیتی از جبر اول متناظر با حاصل‌ضرب همان عضو هیأت P در عضو متناظر جبر دوم باشد.

این تعریف برای یکی دانستن جبرها نشان می‌دهد در نظریهٔ جبر، تنها به بررسی ویژگی‌هایی از کمیت‌ها و دستگاه کمیت‌های جبرها می‌پردازند که ضمن عبارت‌ها به‌صورت برخی ویژگی‌های سه عمل اصلی ظاهر می‌شوند. سخن کوتاه، نظریهٔ جبر ویژگی‌های عمل‌های مربوط به کمیت‌ها را بررسی می‌کند و هیچ توجهی به ماهیت و طبیعت کمیت‌های تشکیل دهندهٔ جبر ندارد.

به‌سادگی ثابت می‌شود، اگر دو جبر یکریخت (ایزومورف) باشند، آن‌وقت کمیت‌هایی که پایهٔ یکی از جبرها را تشکیل می‌دهند، متناظر با کمیت‌های تشکیل دهندهٔ پایهٔ جبر دوم‌اند؛ درضمن ثابت‌های ساختاری در دو پایهٔ متناظر، باهم برابرند. برعکس اگر دو جبر، به‌شرطی که روی یک هیأت تعریف شده باشند، و پایه‌های مناسبی متناظر با ثابت‌های ساختاری برابر داشته باشند، آن‌وقت این دو جبر یکریخت (ایزومورف) هستند.

بین جبرهای مختلف، جبرهای شرکت‌پذیر، یعنی جبرهایی که در آن‌ها عمل ضرب با قانون شرکت‌پذیری $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ سازگار است، نقشی مهم به‌عهده داشته‌اند و هم اکنون هم به‌عهده دارند. طرح ویژگی‌های این گونه جبرها به همین بند مربوط است. بین جبرهای شرکت‌ناپذیر، جبر *لی* از همه جالب‌تر است که برای آن ویژگی زیر برای ضرب وجود دارد:

$$\alpha\beta=-\beta\alpha \ , \ \ \alpha(\beta\gamma)+\beta(\gamma\alpha)+\gamma(\alpha\beta)=0$$

نوع‌های مختلف جبر *لی* از این جهت جالب‌اند که با گروه‌های *لی* (که در بند ۷ دربارهٔ آن‌ها صحبت کردیم)، بستگی نزدیکی دارند.

جبر ماتریس‌ها. پیش از این گفتیم در ادامهٔ نخستین دورهٔ پیشرفت نظریهٔ دستگاه‌های فرامختلط، توجه اصلی روی بررسی دستگاه‌های جداگانه‌ای بود که به دلیلی علاقهٔ پژوهشگران را به‌خود جلب می‌کرد. به بعضی از این دستگاه‌ها اشاره کرده‌ایم. به‌تقریب در میانهٔ سدهٔ نوزدهم، بررسی مربوط به جبر ماتریس‌ها که نقش اصلی را در نظریهٔ جبر به‌عهده دارند، آغاز شد. در این‌جا به‌کوتاهی از عمل‌های مربوط به ماتریس‌ها یاد می‌کنیم (بند ۱ بخش شانزدهم را ببینید).

ماتریس روی هیأت P، به مجموعه‌ای از عضوهای این هیأت گفته می‌شود که به‌صورت جدولی مستطیل شکل درآیند. دو ماتریس را برابر گویند، وقتی عضوهایی از آن‌ها، که در مکان‌های متناظر قرار دارند، باهم برابر باشند. در این‌جا، تنها به بررسی ماتریس‌های مربع می‌پردازیم، ماتریس‌هایی که در آن، تعداد سطرها با تعداد ستون‌ها برابر است. تعداد سطرها، یا تعداد ستون‌های ماتریس را، مرتبهٔ آن گویند.

برای این‌که دو ماتریس با مرتبه‌های برابر را باهم جمع کنیم، عضوهای متناظر آن‌ها را باهم جمع می‌کنیم. ضرب یک عدد در ماتریس، بنابه تعریف، به‌معنای ضرب همهٔ درایه‌های ماتریس (یعنی عضوهای ماتریس) در آن عدد است. ضرب یک ماتریس در ماتریس دیگر، تعریفی پیچیده‌تر دارد: حاصل دو ماتریس مرتبهٔ n، به ماتریسی با همان مرتبه گویند که در آن، درایهٔ سطر iام و ستون jام، برابر باشد با مجموع حاصل‌ضرب‌های درایه‌های سطر iام ماتریس اول در درایه‌های متناظر ستون jام ماتریس دوم. مثال

$$\begin{pmatrix} a & b \\ a_1 & b_1 \end{pmatrix} \begin{pmatrix} x & y \\ x_1 & y_1 \end{pmatrix} = \begin{pmatrix} ax+bx_1 & ay+by_1 \\ a_1x+b_1x_1 & a_1y+b_1y_1 \end{pmatrix}$$

این که چرا تعریف ضرب ماتریس‌ها را به‌این صورت درنظر گرفته‌اند، در بخش شانزدهم شرح داده شده است.

با توجه به‌این تعریف‌ها دربارهٔ ماتریس مرتبهٔ n با درایه‌هایی از هیأتی مثل P، دستگاهی از کمیت‌ها به‌وجود می‌آید که می‌توان آن‌ها را باهم جمع کرد، در عضوی از هیأت P ضرب و یا خود این کمیت‌ها را درهم ضرب کرد. محاسبه‌هایی ساده نشان می‌دهد، ویژگی‌های ′۱ تا ′۱۰، که معرف جبر بودند، در این‌جا برقرارند. به‌جز این، به‌سادگی ثابت می‌شود، ضرب ماتریس‌ها از قانون شرکت‌پذیری پیروی می‌کند. بنابراین، دستگاه همهٔ

ماتریس‌های مرتبهٔ n با درایه‌هایی از هیأت مفروض P، تشکیل جبری شرکت‌پذیر روی این هیأت می‌دهند.

روشن است برابری

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

ثابت می‌کند که چهار ماتریس سمت راست برابری، پایهٔ جبر ماتریس‌های مرتبهٔ دوم را تشکیل می‌دهند. در حالت کلی، اگر ماتریسی را که در سطر i و ستون j آن عدد ۱ و در بقیهٔ مکان‌ها عدد صفر قرار گرفته باشد، با ε_{ij} نشان دهیم، باتوجه به‌برابری

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum_{i,j} a_{ij}\ \varepsilon_{ij}$$

نشان می‌دهد، ماتریس ε_{ij}، پایهٔ جبر ماتریس‌های مرتبهٔ n را تشکیل می‌دهد. چون تعداد ماتریس‌های ε_{ij} برابر n^2 است، بنابراین رتبهٔ جبر ماتریس‌ها هم برابر n^2 است. جدول ضرب ماتریس‌های پایهٔ ε_{ij}، به‌این صورت است:

$$\varepsilon_{ij} \cdot \varepsilon_{jl} = \varepsilon_{il}\ ,\ \varepsilon_{ij} \cdot \varepsilon_{kl} = 0\ ,\ j \neq k\ ,\ i\ ,\ j\ ,\ k\ ,\ l = 1\ ,\ 2\ ,\ \dots\ ,\ n$$

جبر ماتریس‌ها شامل واحد است؛ نقش واحد را، ماتریس همانی به‌عهده دارد.

نمایش‌های جبرهای شرکت‌پذیر. فرض کنید هر کمیتی از یک جبر مثل A روی هیأت P، متناظر با کمیتی از یک جبر دیگر B روی همین هیأت P باشد. اگر در ضمن، مجموع و حاصل‌ضرب هر دو عضو جبر A، متناظر مجموع و حاصل‌ضرب عضوهای متناظر جبر B و حاصل‌ضرب هر عضو هیأت P در عضو دل‌خواهی از جبر A متناظر حاصل‌ضرب همان عضو هیأت P در عضو متناظر جبر B باشد، گویند جبر A به‌طور همریخت (*homomorphic*) در جبر B نگاشته شده است. نگاشت همریخت جبر شرکت‌پذیر در جبر ماتریس‌های مرتبهٔ n، نمایش مرتبهٔ n جبر A نامیده می‌شود. اگر عضوهای مختلف جبر A متناظر با ماتریس‌های مختلف باشند، آن‌وقت نمایش را صادق یا یکریخت (ایزومورف) گویند. وقتی نمایش جبر

A به‌وسیلهٔ ماتریس‌ها، یکریختی باشد، می‌توان عمل‌های روی کمیت‌های جبر را منجر به عمل‌ها روی ماتریس‌های متناظر کرد. به‌همین مناسبت، مسألهٔ پیدا کردن نمایش‌های یک جبر اهمیت زیادی دارد. در این‌جا از ساده‌ترین روش‌ها برای پیدا کردن نمایش جبر نام می‌بریم که البته نقش مهمی در نظریهٔ کلی به‌عهده دارند.

در جبر شرکت‌پذیر A، پایه‌ای مثل ε_1، ε_2، ... ، ε_n در نظر می‌گیریم و α را مقدار دل‌خواهی از A فرض می‌کنیم. $\varepsilon_1\alpha$، $\varepsilon_2\alpha$، ... ، $\varepsilon_n\alpha$ هم مقدارهایی از A هستند و، بنابراین، به‌صورت خطی برحسب ε_1، ε_2، ... ، ε_n قابل بیان باشند. فرض کنید

$$\varepsilon_1\alpha = a_{11}\,\varepsilon_1 + a_{12}\,\varepsilon_2 + \cdots + a_{1n}\varepsilon_n\,,$$
$$\varepsilon_2\alpha = a_{21}\,\varepsilon_1 + a_{22}\,\varepsilon_2 + \cdots + a_{2n}\varepsilon_n\,,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varepsilon_n\alpha = a_{n1}\,\varepsilon_1 + a_{n2}\,\varepsilon_2 + \cdots + a_{nn}\varepsilon_n$$

می‌بینیم، با ثابت نگه داشتن پایه، می‌توان به‌هر عضو α ماتریس معین $\|a_{ij}\|$ را متناظر کرد. محاسبهٔ ساده‌ای نشان می‌دهد، این تناظر، نمایشی از جبر A است. این نمایش را اغلب، نمایش عادی یا منظم جبر A گویند و روشن است، درجهٔ آن، برابر مرتبهٔ جبر است.

عددهای مختلط را می‌توان همچون جبری با مرتبهٔ ۲ روی هیأت عددهای حقیقی و با پایهٔ ۱ و i درنظر گرفت. برابری‌های

$$1.(a+bi)=a.1+b.i$$
$$i.(a+bi) = -b.1 + a.i$$

نشان می‌دهد، درنمایش عادی عددهای مختلط، عدد $a+bi$ متناظر با ماتریس $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ است. به‌همین ترتیب، نمایش چهارگان‌ها به‌صورت

$$a+bi+cj+dk \to \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$$

است. این نمایش‌ها برای عددهای مختلط و چهارگان‌ها، نمایش‌های صادق‌اند (یعنی با

خود جبر ایزومورف یا یکریخت‌اند). ولی می‌توان مثال‌هایی آورد که نشان دهند، نمایش عادی همیشه صادق نیست. ولی اگر جبر شامل یکه باشد، آن‌وقت نمایش عادی آن، بی‌تردید صادق است.

می‌توان ثابت کرد، هر جبر شرکت‌پذیر، داخل جبری همراه با یکه است. نمایش عادی جبر گسترده‌تر صادق است و، بنابراین، نمایش برای جبر مفروض هم صادق است. به‌این ترتیب، هر جبر شرکت‌پذیر دارای نمایش صادقی از ماتریس‌ها است.

این روش پیدا کردن نمایش، برای ساختن همهٔ نمایش‌های جبرها کافی نیست. روش ظریف‌تر مربوط به مفهوم ایده‌آل در جبر می‌شود که در ریاضیات امروزی نقشی جدی دارد.

دستگاه I عضوهای یک جبر را ایده‌آلی (یا آرمانیِ) راست گویند، وقتی که زیر فضای خطی جبر باشد و، درضمن، حاصل‌ضرب هر عضو از I در هر عضو جبر، دوباره عضوی از I باشد. به‌همین ترتیب (با جابه‌جا کردن ردیف عامل‌ها)، ایده‌آل چپ تعریف می‌شود. ایده‌آلی که درعین حال هم چپ و هم راست باشد، دوسویه (یا دوطرفه) نامیده می‌شود. به‌خودی خود روشن است، عضو صفر جبر، ایده‌آلی دوسویه است. همچنین، هر جبر را می‌توان به‌عنوان ایده‌آل دوسویهٔ خاص خودش درنظر گرفت. البته، به‌جز این دو ایده‌آل بدیهی، جبر می‌تواند شامل ایده‌آل‌های دیگری باشد که وجود آن‌ها، اغلب مربوط به ویژگی‌های جالب جبر است.

فرض کنید جبر شرکت‌پذیر A، دارای ایده‌آل راست I باشد. در این ایده‌آل، پایهٔ ε_1، ε_2، ... ، ε_m را انتخاب می‌کنیم. از آن‌جا که I، در حالت کلی، تنها بخشی از A است، بنابراین پایهٔ I، اغلب عضوهای کمتری نسبت به پایهٔ A دارد. α را عضو دل‌خواهی از A می‌گیریم. چون I ایده‌آل راست و ε_1، ε_2، ... ، ε_m جزئی از I است، بنابراین I شامل حاصل‌ضرب‌های $\varepsilon_1\alpha$، ... ، $\varepsilon_m\alpha$ هم خواهد بود، یعنی برحسب پایهٔ ε_1، ... ، ε_m به‌صورت خطی بیان می‌شوند:

$$\varepsilon_1 \alpha = a_{11} \varepsilon_1 + \cdots + a_{1m}\varepsilon_m ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\varepsilon_m \alpha = a_{m1} \varepsilon_1 + \cdots + a_{1mm}\varepsilon_m ,$$

اگر ماتریس $\|a_{ij}\|$ را متناظر با عضو α قرار دهیم، مثل قبل، نمایشی از جبر A به‌دست

می‌آید. درجهٔ این نمایش برابر تعداد عضوهای پایهٔ ایده‌آل I و بنابراین، در حالت کلی، کمتر از درجهٔ نمایش عادی است. روشن است درجهٔ نمایشی که به‌وسیلهٔ ایده‌آل به‌دست می‌آید، وقتی کمترین است که خود ایده‌آل کمین (مینیمال) باشد. از این‌جا، اهمیت و نقش خاص ایده‌آل‌های کمین در نظریهٔ جبر، معلوم می‌شود.

ساختار جبر. باتوجه به آن‌چه گفتیم، هر جبر شرکت‌پذیر A می‌تواند با ماتریس‌هایی از یک مرتبه، یکریخت (ایزومورف) باشد. در این نمایش مجموعهٔ ماتریس‌هایی که نگاره (یا تصویر) مقدارهای جبر A هستند، خود یک جبر است، ولی تنها بخشی از جبر همهٔ ماتریس‌های این مرتبه. اگر بخشی از مقدارهای جبر، خود یک جبر را تشکیل دهند، آن‌وقت آن را «زیرجبر» جبر مفروض گویند. بنابراین، می‌توان گفت، هر جبر شرکت‌پذیر، با یک زیرجبر ماتریس‌ها یکریخت (یا ایزومورف) است.

با این که این نتیجه‌گیری اهمیت جدی دارد، زیرا مسألهٔ مربوط به پیدا کردن همهٔ جبرها را، منجر به پیدا کردن همهٔ زیرجبرهای ممکن جبرهای ماتریسی می‌کند، به‌طور مستقیم به مسألهٔ مربوط به ساختار جبر پاسخ نمی‌دهد. پاسخ کلی به این مسأله را، برای نخستین بار، *ف.اِ.مُولین* (۱۸۶۱-۱۹۴۱) در سال‌های پایانی سدهٔ نوزدهم داد.

جبر را **ساده** گویند وقتی که شامل هیچ ایده‌آل دوسویه‌ای به‌جز صفر و خودش نباشد. *مُولین* ثابت کرد، همهٔ جبرهای شرکت‌پذیر سادهٔ با رتبهٔ ۲ یا بیشتر، روی هیأت عددهای مختلط، با جبر همهٔ ماتریس‌های با مرتبهٔ مناسب و روی همین هیأت، یکریخت (ایزومورف)اند.

به‌دنبال این بررسی‌ها *وِدِربرن* در آغاز سدهٔ بیستم، به نتیجه‌هایی رسید که به‌طور کامل، ساختار جبر را روی هر هیأت دل‌خواهی مشخص می‌کرد.

دستگاهی شامل عضوهای جبر A (به‌ویژه خود جبر A، بعضی از ایده‌آل‌ها یا زیرجبرهای آن)، پوچ‌توان (*nilpotent*) نامیده می‌شود، به‌شرطی که عدد طبیعی s وجود داشته باشد، به‌نحوی که حاصل‌ضرب هر s عضو از دستگاه برابر صفر شود. هر جبر شرکت‌پذیر یک ایده‌آل یکتای پوچ‌توانِ دوسویهٔ بیشین (ماکزیمال) دارد که رادیکال یا ریشگی جبر نامیده می‌شود. جبری که ریشگی آن برابر صفر باشد، **نیم‌ساده** نامیده می‌شود. می‌توان ثابت کرد، هر جبر نیم‌ساده به‌صورت خاصی به مجموع جبرهای ساده تبدیل می‌شود و در نتیجه، بررسی جبرهای نیم‌ساده، منجر به بررسی جبرهای ساده خواهد شد.

سرانجام، جبر A را **جبر تقسیمی** گویند، وقتی که در A، هر معادلهٔ به صورت $ax=b$ $(a\neq 0)$ دارای جواب باشد.

ساختار جبرهای ساده روی هیأت عددهای مختلط، به طور کامل به وسیلهٔ قضیهٔ ***مؤلین*** شرح داده می‌شود. اگر هم P هیأتی دل‌خواه باشد، آن وقت می‌توان از قضیهٔ کلی‌تر ***وِدِربرن*** استفاده کرد. هر جبر سادهٔ با رتبهٔ ۲ یا بیشتر روی هیأت P، با جبر همهٔ ماتریس‌ها با مرتبهٔ مناسب و عضوهایی از یک جبر تقسیمی روی همان هیأت P، یکریخت (ایزومورف) است. به این ترتیب، قضیهٔ ***ودربرن*** مسألهٔ مربوط به پیدا کردن جبرهای ساده روی میدان مفروض P را، منجر به پیدا کردن جبر تقسیمی روی هیأت P می‌کند. روی هیأت عددهای مختلط، تنها یک جبر تقسیمی وجود دارد: خود هیأت عددهای مختلط. از این جا، از قضیهٔ ***وِدِربرن*** نتیجه می‌شود. همهٔ جبرهای ساده روی هیأت عددهای مختلط، با جبر ماتریس‌ها روی همین هیأت یکریخت است، یعنی قضیهٔ ***مؤلین*** به دست می‌آید.

روی هیأت عددهای حقیقی، تنها سه جبر شرکت‌پذیر تقسیمی وجود دارد: خود هیأت عددهای حقیقی، هیأت عددهای مختلط و جبر چهارگان‌ها. اثبات این گزاره چندان ساده نیست و در این‌جا به آن نمی‌پردازیم. از این جا و با توجه به قضیهٔ ***وِدِربرن*** نتیجه می‌شود که هر جبر ساده روی هیأت عددهای حقیقی با جبر ماتریس‌های با مرتبهٔ مناسب یا روی هیأت عددهای حقیقی، یا روی هیأت عددهای مختلط و یا روی جبر چهارگان‌ها، یکریخت است.

از این مثال‌ها دیده می‌شود، چگونه ساختمان جبرهای نیم‌ساده، با قضیه‌های ***مؤلین*** و ***وِدِربرن*** گشوده می‌شود. برای آن چه به جبرهای باریشگی مربوط می‌شود، قضیهٔ اصلی ***وِدِربرن*** اهمیت جدی دارد. طبق این قضیه، با برخی محدودیت‌ها دربارهٔ هیأت ضریب‌ها، در هر جبر A باریشگی (رادیکال) R، زیرجبر نیم‌سادهٔ L وجود دارد که هر عضو جبر مفروض را می‌تواند به صورت یکتا به وسیلهٔ مجموع $\lambda+\rho$ $(\rho\in R\ ,\ \lambda\in L)$ نشان داده شود؛ در ضمن زیرجبر L به مفهومی، به صورت یکتا در درون جبر A معین می‌شود.

این قضیه‌های اصلی، تصور کاملی دربارهٔ امکان شناخت گونه‌های جبرهای شرکت‌پذیر به ما می‌دهد؛ در ضمن، مسألهٔ مربوط به ساختمان آن‌ها را، در اساس، منجر به مسألهٔ مشابه آن یعنی ساختمان جبرهای پوچ‌توان می‌کند. این نظریه‌ها، هنوز در مرحلهٔ بررسی بیشترند.

۱۳. جبرهای لی

در بند ۱۲ گفتیم، به‌جز نظریهٔ جبرهای شرکت‌پذیر، نظریهٔ مربوط به جبرهای *لی* هم پیشرفت زیادی کرده است، که در آن‌ها، ضرب‌ها با شرط‌های زیر سازگارند:

$$\alpha\beta = -\beta\alpha \ , \ \alpha(\beta\gamma) + \beta(\gamma\alpha) + \gamma(\alpha\beta) = 0$$

اهمیت این جبرها در آن است که بستگی نزدیکی با گروه‌های *لی* (بند ۷ را ببینید)، یعنی با مهم‌ترین خانوادهٔ گروه‌های پیوسته، دارند. در ضمن دیدیم، گروه‌های *لی* برای هندسهٔ امروزی نقشی اساسی به عهده دارند. جبر *لی* روی هیأت‌های همهٔ عددهای حقیقی و همهٔ عددهای مختلط، بیش از همه جالب است.

یکی از ساده‌ترین نمونه‌های جبر *لی* را در این جا می‌آوریم. مجموعهٔ همهٔ ماتریس‌های مربع با مرتبهٔ مفروض n را در نظر می‌گیریم. برای ماتریس A و B از این مجموعه، مفهوم جابه‌جاگر (یا تعویض‌گر) $AB - BA$ را تشکیل و به $[A \ , \ B]$ نشان می‌دهیم.

به سادگی می‌توان تحقیق کرد که

$$[A \ , B] = -[B \ , A] \ ,$$

$$[A \ , [B \ , C]] + [B \ , [C \ , A]] + [C \ , [A \ , B]] = 0$$

بنابراین مجموعهٔ همهٔ ماتریس‌های مربع با مرتبهٔ مفروض، نسبت به عمل جابه جاگری تشکیل یک جبر لی را می‌دهند. روشن است هر زیرجبری *لی*، یعنی هر مجموعه‌ای از ماتریس‌ها که نسبت به عمل‌های جمع، ضرب در عددی از هیأت پایه و عمل جابه‌جاگری (به مفهومی که آوردیم) بسته باشد، به نوبهٔ خود یک جبر *لی* را تشکیل می‌دهد.

این پرسش که، آیا برای همهٔ جبرهای مفروض انتزاعی *لی*، جبر ماتریسی یکریخت (ایزومورف) با آن وجود دارد، مدت‌ها بی‌پاسخ مانده بود تا سرانجام در سال ۱۹۳۵ به وسیلهٔ *ای.د.آدو*، شاگرد متخصص مشهور جبر *ن.گ.چه بُوتارِف* حل شد.

بدون این‌که موضوع را به تفصیل بررسی کنیم و بدون این که به تنظیم دقیق بستگی بین گروه‌های *لی* و جبر *لی* بپردازیم، در خط کلی خود، وقتی بتوان گروه *لی* و جبر *لی* را به صورت ماتریس‌ها نشان داد، اشاره می‌کنیم.

L را یک جبر ماتریسی *لی* می‌گیریم. هر ماتریس A را که به L تعلق دارد، با ماتریس

$$U=e^A=E+\frac{A}{1!}+\frac{A^2}{2!}+\cdots$$

متناظر می‌کنیم. در این صورت، مجموعهٔ همهٔ ماتریس‌هایی که به این ترتیب به دست می‌آید، نسبت به ضرب معمولی ماتریس‌ها، گروه *لی* را تشکیل می‌دهد. برعکس، برای هر گروه *لی* با تقریب یکریختی، یک جبر یکتای *لی* پیدا می‌شود، به نحوی که گروه متناظر با آن، با گروه مفروض یکریخت باشد.

برای سادگی کار، به جای تنظیم دقیق قضیهٔ مربوط به تناظر بین گروه‌ها و جبرهای *لی* ساده شدهٔ آن را می‌آوریم. در واقع، تناظر $U=e^A$ و A تنها وقتی وجود دارد که U به اندازهٔ کافی نزدیک به ماتریس همانی، و A به اندازهٔ کافی نزدیک به ماتریس صفر باشد. برای تنظیم دقیق‌تر قضیه، باید با مفهوم‌های بغرنج گروه موضعی و یکریختی (ایزومورفیسم) موضعی آشنا باشیم.

به این ترتیب، عبور از جبر *لی* به گروه متناظر آن، به یاری عملی شبیه پیدا کردن «آنتی لگاریتم» انجام می‌شود. برعکس، عبور از گروه به جبر، با عملی شبیه «لگاریتم گرفتن» به انجام می‌رسد.

اگر L بر جبر همهٔ ماتریس‌های مرتبهٔ n منطبق باشد، آن وقت گروه متناظر *لی*، گروه همهٔ ماتریس‌های ناتکین (وارون‌پذیر) است. زیرا هر ماتریس U که به واحد نزدیک باشد، می‌تواند به صورت $U=e^A$ نشان داده شود.

ماتریس $\|a_{ij}\|$ را «متقارن چاوله» (*skew – symmetric*) گویند وقتی عضوهای آن در رابطهٔ $a_{ji}=-a_{ij}$ صدق کنند. ماتریس‌های متقارن چاوله، تشکیل جبر *لی* می‌دهند، زیرا اگر A و B متقارن چاوله باشند، آن وقت ماتریس‌های

$$AB-BA=[A\,,B]\,,\alpha A+\beta B$$

هم متقارن چاوله از آب در می‌آیند.

به سادگی می‌توان تحقیق کرد، برای هر ماتریس متقارن چاوله A، عبارت e^A یک ماتریس متعامد (اورتوگونال) است؛ در ضمن هر ماتریس متعامد نزدیک به واحد می‌تواند به صورت نمایی مذکور نشان داده شود. بنابراین، جبر *لی* از گروه ماتریس‌های متعامد، عبارت است از جبر ماتریس‌های تقارن چاوله.

از هندسهٔ تحلیلی می‌دانیم، هر دوران فضا دور مبداء مختصات با ماتریس متعامد داده می‌شود؛ در ضمن حاصل‌ضرب دوران‌ها، متناظر با حاصل‌ضرب ماتریس‌های متناظر است. به زبان دیگر، گروه دوران‌های فضا دور یک نقطهٔ ثابت با گروه ماتریس‌های متعامد مرتبهٔ ۳ یکریخت است. از این جا نتیجه می‌گیریم، جبر *لی* برای گروه دوران‌های فضا، عبارت است از جبر همهٔ ماتریس‌های متقارن چاوله مرتبهٔ ۳، یعنی جبر *لی* از ماتریس‌های به صورت

$$A=\begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix}$$

از آن جا که هر یک از ماتریس‌ها، به طور کامل با سه عدد a ، b و c مشخص می‌شود، می‌توان آن را با برداری مثل $\boldsymbol{a}$ بیان کرد که روی محورهای مختصات، تصویرهایی برابر a ، b و c داشته باشد. در ضمن ترکیب خطی $\alpha A_1 + \beta A_2$ از این گونه ماتریس‌های A_1 و A_2، به روشنی متناظر با ترکیب خطی بردارهای متناظر $\alpha \boldsymbol{a}_1 + \beta \boldsymbol{a}_2$ است، و ماتریس جابه‌جاگر

$$[A_1 , A_2]=A_1A_2-A_2A_1=$$

$$=\begin{pmatrix} 0 & -a_1 & -b_1 \\ a_1 & 0 & -c_1 \\ b_1 & c_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -a_2 & -b_2 \\ a_2 & 0 & -c_2 \\ b_2 & c_2 & 0 \end{pmatrix}-\begin{pmatrix} 0 & -a_2 & -b_2 \\ a_2 & 0 & -c_2 \\ b_2 & c_2 & 0 \end{pmatrix}\begin{pmatrix} 0 & -a_1 & -b_1 \\ a_1 & 0 & -c_1 \\ b_1 & c_1 & 0 \end{pmatrix}$$

$$=\begin{pmatrix} 0 & b_2c_1-b_1c_2 & a_1c_2-a_2c_1 \\ b_1c_2-b_2c_1 & 0 & a_2b_1-a_1b_2 \\ a_2c_1-a_1c_2 & a_1b_2-a_2b_1 & 0 \end{pmatrix}$$

متناظر با برداری با تصویرهای $b_1c_2-b_2c_1$ ، $a_2c_1-a_1c_2$ و $a_1b_2-a_2b_1$ یعنی حاصل ضرب برداری بردارهای $\boldsymbol{a}_1$ و $\boldsymbol{a}_2$ است. به این نتیجهٔ جالب می‌رسیم که، مجموعهٔ بردارهای معمولی نسبت به جمع، ضرب در عدد و ضرب برداری یک جبر *لی* را می‌سازند که متناظر با گروه دوران‌های فضا دور نقطه‌ای ثابت است. این نتیجه‌گیری، یکبار دیگر ثابت می‌کند، مفهوم‌های هندسی تا چه اندازه با گروه دوران‌های فضا، به زبان دیگر با قانون‌های حرکت

جسم‌های صلب بستگی نزدیک دارد.

در پایان سدهٔ نوزدهم و آغاز سدهٔ بیستم، نتیجه‌هایی برای جبر *لی* به دست آمد که با نتیجه‌های اساسی جبرهای شرکت‌پذیر شباهت داشت، گرچه اثبات و تنظیم این نتیجه‌گیری‌ها، به دلیل پیچیده بودن آن‌ها، در این جا میسر نیست. از جمله، با تلاش‌های *لی*، *کی‌لینگ* و *کارتان* در آغاز سدهٔ بیستم ثابت شد، برای جبر *لی* می‌توان مفهوم‌های ریشگی و نیم‌ساده را به‌کار برد و همهٔ جبرهای سادهٔ *لی* را روی هیأت عددهای حقیقی و هیأت عددهای مختلط به دست آورد. در آغاز سال‌های ۳۰ سدهٔ بیستم، *کارتان* و *ویل* پایهٔ نظریهٔ نمایش جبر *لی* را به یاری ماتریس‌ها ریختند که وسیلهٔ نیرومندی برای حل بسیاری از مسأله‌ها بود. در سال‌های بعد، ریاضی‌دانان شوروی توانستند نتیجه‌های جالبی در زمینهٔ جبر *لی* پیدا کنند، از جمله نمایش جبرهای *لی* و حل مسألهٔ مربوط به زیرجبرهای نیم‌سادهٔ جبر *لی* و ساختن جبر با مفروض بودن ریشگی و جز آن را به پایان رساندند.

۱۴. حلقه

در بند ۱۱، تعریف کلی هیأت را، به عنوان مجموعهٔ دل‌خواهی از عضوها که دربارهٔ آن‌ها بتوان عمل‌های جمع و ضرب را با توجه به شرط‌های ۱ تا ۱۰ انجام داد، آوردیم. اگر از این تعریف، شرط ۱۰ را دربارهٔ وجود خارج قسمت و شرط‌های ۷ و ۸ را دربارهٔ جابه‌جایی‌پذیر و شرکت‌پذیر بودن ضرب کنار بگذاریم، به تعریف مفهوم حلقه می‌رسیم که یکی از مهم‌ترین مفهوم‌ها در جبر امروز است.

هر هیأت و همچنین هر جبر، که تنها نسبت به عمل‌های جمع و ضرب در نظر گرفته شود، یک حلقه است. ساده‌ترین مثال برای حلقه، مجموعهٔ عددهای درست گویا با عمل‌های معمولی جمع و ضرب است. نسبت به این عمل‌ها، مجموعهٔ عددهای به صورت

$$a+bi\,,\ a+b\sqrt{۲}\,,\ a+b\sqrt[۳]{۲}+c\sqrt[۳]{۴}\,,\ \ldots$$

هم یک حلقه را تشکیل می‌دهند، که در آن‌ها a ، b و c عددهای درست گویا هستند. عضوهای این حلقه‌ها عددند و بنابراین، خود حلقه هم، حلقه‌های عددی نامیده می‌شوند. برخی از ویژگی‌های حلقه‌ها را در بخش چهارم (جلد اول) و بخش دهم (جلد دوم) دیده‌ایم.

ولی خانواده‌های مهمی از حلقه‌های غیرعددی هم وجود دارد. از جمله می‌توان از مجموعهٔ چندجمله‌ای‌ها با متغیرهای مفروض x_1 ، x_2 ، ... ، x_n و ضریب‌هایی از یک حلقه یا هیأت، مجموعهٔ همهٔ تابع‌های پیوسته که در حوزهٔ (یا دامنه) معینی تعریف شده باشند، مجموعهٔ تبدیل‌های خطی فضای خطی یا فضای هیلبرتی به عنوان حلقه نسبت به عمل‌های عادی جمع و ضرب نام برد.

ویژگی‌های حسابی حلقه‌های عددی موضوع بررسی نظریهٔ عمیق عددهای جبری است که در مرز بین جبر خاص و نظریهٔ خاص عددها قرار دارد. ویژگی‌های حلقه‌های چندجمله‌ای‌ها، به نظریهٔ ایده‌آل‌های چندجمله‌ای‌ها مربوط می‌شود که بستگی نزدیکی با شاخه‌های عالی هندسهٔ تحلیلی دارد. سرانجام، حلقه‌های مربوط به تابع‌ها و تبدیل‌ها، نقشی اساسی در آنالیز تابعی (بخش نوزدهم را ببینید) دارد.

بر پایهٔ این نظریه و بسیاری از نظریه‌های دیگر، در جریان سده‌ها، نظریهٔ عمومی حلقه‌ها و نظریهٔ حلقه‌های توپولوژیک، به سرعت گسترش یافت.

با توجه به کمبود جا، در این‌جا تنها به نتیجه‌گیری‌های جداگانه‌ای از نظریهٔ حلقه‌ها اشاره می‌کنیم.

ایده‌آل‌ها. زیرمجموعهٔ I از عضوهای حلقهٔ K (که شرکت‌پذیر بودن آن‌ها اجباری نیست) ***ایده‌آل*** آن نامیده می‌شود، به شرطی که «تفاضل» هر دو عضو از I، باز هم عضوی از I باشد و به شرطی که حاصل‌ضرب‌های ax ، xa – عضو دل‌خواه a از I در عضو دل‌خواه x از K – در I قرار گیرد.

هر ایده‌آل چنان بخشی از حلقه است که خودش هم، نسبت به عمل‌های جمع و ضرب، یک حلقه است. چنین بخش‌هایی را «زیرحلقهٔ» حلقهٔ مفروض می‌نامند، یعنی هر ایده‌آل، در ضمن، یک زیرحلقه است. ولی عکس این حکم، همیشه درست نیست.

اشتراک هر دستگاه از ایده‌آل‌های حلقه، باز هم ایده‌آلی از این حلقه است، به‌ویژه اشتراک همهٔ ایده‌آل‌ها، که شامل عضو ثابتی از حلقه مثل a باشند، یک ایده‌آل حلقه است. این ایده‌آل را ایده‌آل اصلی تولید شده به وسیلهٔ عضو a گویند و با نماد (a) نشان می‌دهند.

به همین ترتیب، مفهوم ایده‌آلی که با دو یا چند عضو تولید می‌شوند، تعریف می‌شود. به سادگی ثابت می‌شود، اگر حلقهٔ شرکت‌پذیر و جابه‌جایی‌پذیر، دارای عضو یکه باشد، آن وقت ایده‌آلی که از عضوهای a_1 ، ... ، a_n پدید می‌آید، به طور ساده، مجموعهٔ همهٔ

عضوهای حلقه است که می‌توان به صورت مجموع $x_1a_1 + \cdots + x_na_n$ نوشت که در آن x_1، ...، x_n عضوهای دل‌خواهی از حلقه‌اند. به‌ویژه، وقتی ایده‌آل اصلی (a) مربوط به حلقهٔ شرکت‌پذیر و جابه‌جایی‌پذیر و همراه با یکه باشد، به طور ساده عبارت است از مجموعهٔ همهٔ عضوهای مضرب a یعنی xa.

در حلقهٔ همهٔ عددهای درست گویا، هر ایده‌آل در ضمن ایده‌آلی اصلی است. حلقه چندجمله‌ای‌ها با یک متغیر و ضریب‌هایی از یک هیأت، حلقهٔ عددهای مختلط به صورت $a+bi$ که در آن a و b عددهایی درست و گویا هستند و یک رشته حلقه‌های دیگر هم، دارای همین ویژگی هستند. ولی مجموعهٔ همهٔ چندجمله‌ای‌های با دو متغیر x و y و بدون مقدار ثابت، ایده‌آل اصلی در حلقهٔ همهٔ چندجمله‌ای‌های با متغیرهای x و y و ضریب‌های گویا نیست.

شبیه آن چه دربارهٔ زیرگروه‌های نرمال در نظریهٔ گروه‌ها داشتیم، برای هر ایده‌آل I از حلقهٔ K هم می‌توان حلقه خارج قسمت یا حلقهٔ رده‌های مانده‌ای K/I را ساخت که به این ترتیب انجام می‌گیرد. عضوهای a و b از حلقهٔ K را هم‌نهشت نسبت به ایده‌آل I می‌نامند و با نشانهٔ نمادین $a \equiv b(I)$ نشان می‌دهند، وقتی که تفاضل $a-b$ در I باشد. به سادگی روشن می‌شود که رابطهٔ «هم‌نهشتی» متقارن، بازتابی و ترایا است (بخش پانزدهم را ببینید)، و بنابراین، همهٔ عضوهای K به رده‌هایی تقسیم می‌شوند که به وسیلهٔ ایده‌آل I، نسبت به هم «هم‌نهشت»اند. اکنون این رده‌ها را به عنوان عضوهای مجموعهٔ تازه‌ای در نظر می‌گیریم، و برای آن‌ها مفهوم‌های جمع و ضرب را وارد کنیم ـ «مجموع» دو رده به رده‌ای گفته می‌شود که شامل مجموع یک عضو ردهٔ اول با یک عضو ردهٔ دوم باشد و به همین ترتیب حاصل‌ضرب دورده به رده‌ای گفته می‌شود که شامل حاصل‌ضرب یک عضو ردهٔ اول با یک عضو ردهٔ دوم باشد. از تعریف ایده‌آل‌ها نتیجه می‌شود، چنین تعریف مجموع و حاصل‌ضربی، در واقع به انتخاب عضوهای رده‌ها مربوط نیست و بنابراین، مجموعهٔ رده‌ها، یک حلقه را تشکیل می‌دهند.

نقش حلقهٔ خارج قسمت در نظریه‌ها حلقه‌ها، شبیه نقش گروه بهری در نظریهٔ گروه‌ها است. به‌ویژه ساختن حلقهٔ خارج‌قسمت از یک حلقهٔ مفروض روش ساده‌ای برای تشکیل حلقه‌های با ویژگی‌های مختلف به دست می‌دهد. به‌جز این می‌توان ثابت کرد هر حلقهٔ جابه‌جایی‌پذیر دل‌خواه K با حلقهٔ خارج‌قسمتی از حلقهٔ چندجمله‌ای‌های با ضریب‌های درست گویا با تعداد کافی متغیرها یکریخت (ایزومورف) است.

ویژگی‌های حسابی حلقه‌ها. در حلقه‌های عددی و در هیأت‌ها، حاصل‌ضرب چند عضو تنها وقتی می‌تواند برابر صفر شود که دست کم یکی از عامل‌های ضرب برابر صفر باشد. در حلقه‌های دل‌خواه، ممکن است این گزاره درست نباشد؛ از جمله حاصل ضرب دو ماتریس مخالف صفر، می‌تواند برابر صفر باشد. اگر در حلقه‌ای با فرض $a \neq 0$ و $b \neq 0$ داشته باشیم $ab = 0$، آن وقت a و b را بخشیاب‌های صفر گویند، ولی اگر چنین عضوهایی در حلقه وجود نداشته باشد، آن حلقه را ***حلقهٔ بدون بخشیاب‌های صفر*** می‌نامند.

ضمن بررسی قانون‌های بخش‌پذیری، در حلقه‌ها، به طور معمول فرض می‌کنند، حلقه جابه‌جایی‌پذیر باشد و بخشیاب‌های صفر نداشته باشد. چنین حلقه‌هایی را، حوزه‌های صحیح می‌نامند. حلقه‌های عددی و حلقه‌های چندجمله‌ای از جمله حوزه‌های صحیح‌اند.

K را حوزه‌ای صحیح فرض می‌کنیم. گویند عضو a در حوزهٔ K را بر عضو b بخش‌پذیر است، وقتی داشته باشیم:

$$a = bq\ ,\ \ q \in K$$

از این جا بلافاصله نتیجه می‌شود، مجموع عضوهای بخش‌پذیر بر b، بر b بخش‌پذیر است؛ همچنین حاصل‌ضرب چند عضو K، وقتی بر b بخش‌پذیر است که دست کم یکی از عامل‌های ضرب بر b بخش‌پذیر باشد. اگر مفهوم «***عضو اول***» را شبیه مفهوم عدد اول وارد کنیم، در نظریهٔ حلقه‌ها هم به‌دشواری‌هایی برمی‌خوریم که دشواری‌های بخش دهم (جلد دوم) را به‌یاد می‌آورد. به‌ویژه در این جا باید مفهوم عضوهای وابسته حلقه را وارد کنیم: عضوهای a و b وابسته‌اند وقتی که a بر b و b بر a بخش‌پذیر باشد. فرض می‌کنیم $a = bq_1$ و $b = aq_2$، در این صورت به‌دست می‌آید:

$$ab = ab.q_1q_2$$

یعنی $q_1q_2 = e$ که در آن، e عضو یکه حوزهٔ K است. به‌همین مناسبت، خارج قسمت‌های عضوهای وابسته را، ***بخشیاب‌های واحد*** یا عضوهای یکا گویند. هر عضوی از حوزه، بر هر بخشیابی از واحد بخش‌پذیر است. در حلقهٔ عددهای گویای درست، بخشیاب‌های واحد ± 1 و در حلقهٔ عددهای به‌صورت $a+bi$ (a و b عددهای درست‌اند)، بخشیاب‌های واحد ± 1 و $\pm i$ هستند.

هر عضو حوزهٔ صحیح K، تجزیه‌هایی به‌صورت $a = a\varepsilon . \varepsilon^{-1}$ دارد که در آن، ε بخشیابی

از واحد است. این گونه تجزیه‌ها را، بدیهی می‌دانند. اگر تجزیهٔ دیگری، به‌جز تجزیه‌های بدیهی، برای a وجود نداشته باشد، آن وقت a را عضو اول یا عضو تجزیه‌ناپذیر K گویند. در رابطه با یگانه بودن تجزیهٔ عددهای به‌صورت ضرب عددهای اول، جالب است چنان خانواده‌هایی از حلقه‌ها، و از آن جمله حلقه‌های جابه‌جایی‌ناپذیری را پیدا کنیم که این قضیهٔ مشهور مربوط به عددهای درست، دربارهٔ آن‌ها صادق باشد. برای نمونه، این قضیه دربارهٔ حلقه‌های ایده‌آل‌های اصلی، یعنی حوزه‌های صحیح که در آن‌ها همهٔ ایده‌آل‌ها اصلی‌اند، برقرار است.

در رابطه با مسألهٔ یکتا بودن تجزیه به عامل‌های اول بود که خود مفهوم مربوط به ایده‌آل‌ها پدید آمد. به‌تقریب در میانه‌های سدهٔ نوزدهم، **کومر** ریاضی‌دان آلمانی که تلاش می‌کرد ثابت کند، معادلهٔ $x^n+y^n=z^n$، جواب درست و مخالف صفر برای $n \geq ۳$ ندارد، به این اندیشه رسید که عددهای به صورت

$$a_۰+a_۱\xi+\cdots+a_n\xi^{n-۱}$$

را بررسی کند که در آن، $\xi=\cos\frac{۲\pi}{n}+i\sin\frac{۲\pi}{n}$، جواب معادلهٔ $x^n=۱$ و $a_۰$، ...، a_n عددهای درست معمولی‌اند. عددهایی که به‌این صورت باشند، یک حوزهٔ صحیح را تشکیل می‌دهند و **کومر** در آغاز، به‌عنوان یک گزارهٔ روشن، پذیرفت که در این حوزه، قضیهٔ مربوط به یگانه بودن تجزیه به عامل‌های اول، درست است. **کومر** براین اساس قضیهٔ ***فرما*** را ثابت کرد. ولی ضمن پژوهش‌های بعدی روشن شد که فرض مربوط به یگانه بودن تجزیه درست نیست. **کومر** که می‌خواست یگانه بودن تجزیه به عامل‌های اول را حفظ کند، نیاز به بررسی مسألهٔ تجزیهٔ عددهای حوزه به‌عامل‌هایی پیدا کرد که در خود حوزه نباشند و او این عددها را ایده‌آل‌ها نامید. بعدها برای ساختن نظریهٔ کلی، به‌جای عددهای ایده‌آلی، مجموعهٔ عضوهایی از حوزه در نظر گرفته شد که براین یا آن عدد ایده‌آلی بخش‌پذیر باشد؛ و همین‌ها بودند که نام ایده‌آل‌ها را به‌خود گرفتند.

کشف یگانه نبودن تجزیه به عامل‌های اول به‌ویژه برای حلقه‌ها، یکی از جنبه‌های جالبی است که در سدهٔ نوزدهم، منجر به ساختن و گسترده کردن نظریهٔ عددهای جبری شد.

یکی ازکاربردهای جالب این نظریه، به مسألهٔ مربوط به تبدیل عددهای درست معمولی به مجموع توان‌های دوم بود که در بخش دهم (جلد ۲) دربارهٔ آن صحبت شده است. نقش اصلی را در تکامل حلقه‌های عددی، کارهای ریاضی‌دانان شوروی، ***ا.ای.زُولُتو تارِف***،

گ.ف.وُرُونُوف، ای.م.وینوگرادوف و ن.گ.چه‌بوتارِف، به‌عهده داشتند.

چندگوناهای جبری. سرچشمهٔ دیگر نظریهٔ ایده‌آل‌ها، به هندسهٔ جبری مربوط می‌شود. در همان آغاز آشنایی با نظریهٔ منحنی‌های درجه دوم، این شگفتی پدید آمد که، به مجموعه‌ای از دو منحنی جدا از هم ـ شاخه‌های هذلولی ـ که ربطی به هم ندارند، نام یک منحنی ـ منحنی هذلولی ـ را گذاشته‌اند و، در ضمن دو خط راست را هم شاخه‌های جداگانه‌ای از منحنی درجه دوم به‌حساب می‌آورند. توضیح این مطلب در جبر نهفته است: اگر معادلهٔ منحنی‌ها را به صورت $f(x, y) = 0$ در نظر بگیریم و $f(x, y)$ را یک چندجمله‌ای نسبت به x و y فرض کنیم، دو حالت پیش می‌آید. در حالت اول، عبارت سمت چپ برابری، چندجمله‌ای درجه دومی است که قابل تجزیه نیست؛ و در حالت دوم، عبارت سمت چپ برابری را می‌توان به‌ضرب دو عبارت درجهٔ اول تبدیل کرد. در حالتی که منحنی را می‌توان به‌یاری چندجمله‌ای ساده نشدنی $f(x, y)$ نشان داده آن را تحویل‌ناپذیر و در حالت دوم، منحنی را تحویل‌پذیر نامیدند.

وقتی وارد در فضای سه‌بعدی بشویم، کار دشوارتر می‌شود. منحنی فضایی با دستگاهی شامل دو معادلهٔ $f(x, y, z) = 0$ و $g(x, y, z) = 0$ بیان می‌شود؛ در ضمن چندجمله‌ای‌های f و g که به وسیلهٔ یک منحنی معین می‌شوند، یگانه نیستند. در این‌جا، کدام منحنی‌ها را تحویل‌ناپذیر بنامیم؟

طبیعی‌ترین پاسخ را نظریهٔ ایده‌آل‌ها می‌دهد. f_1 ، f_2 و ... را مجموعه‌ای می‌گیریم که شامل چندجمله‌ای‌هایی از متغیرهای x ، y ، z و، در حالت کلی، با ضریب‌های مختلط باشد. مجموعه‌ای از نقطه‌های فضا (فضای مختلط)، که مختصات آن‌ها، همهٔ چندجمله‌ای‌های مفروض را برابر صفر کنند، چند گونای جبری نامیده می‌شود که با چندجمله‌ای‌های مفروض تعریف شده است. این چندگونا را M می‌نامیم و همهٔ چندجمله‌ای‌های شامل متغیرهای x ، y و z را در نظر می‌گیریم که در هر نقطهٔ M برابر صفر شوند. به‌سادگی دیده می‌شود، مجموعهٔ I شامل همهٔ این گونه چندجمله‌ای‌ها، در حلقهٔ چندجمله‌ای‌های شامل x ، y و z ، ایده‌آل است. این ایده‌آل، به‌جز این دارای این ویژگی است که اگر توان یک چندجمله‌ای در I باشد، آنگاه خود آن چندجمله‌ای هم در I است. به دست می‌آید که: در عین‌حال که مجموعه‌های مختلف چندجمله‌ای‌ها می‌توانند یک چندگونای جبری را معرفی کنند، تناظر بین چندگوناها و ایده‌آل‌های با ویژگی اضافی یاد

شده، تناظر یک‌به‌یک است.

به‌این ترتیب، ضمن بررسی ویژگی‌های چندگوناها، به‌طور طبیعی نه بی‌کم و بیش «معادله‌های» تصادفی آن‌ها و بلکه ایده‌آل متناظر آن‌ها مطالعه می‌شود. اگر ایده‌آل I بتواند به‌صورت اشتراک دو ایده‌آل I_1 و I_2 معرفی شود، آن‌وقت چندگونای M، اجتماع چندگوناهای M_1 و M_2 است که متناظر ایده‌آل‌های I_1 و I_2 هستند. بنابراین روشن می‌شود، چندگونای M را وقتی می‌توان تحویل‌ناپذیر دانست که، ایده‌آل I متناظر با آن را نتوان به‌صورت اشتراک دو ایده‌آل بزرگتر دیگر نشان داد. تبدیل منحنی به منحنی‌های با درجهٔ پایین‌تر، تجزیهٔ چند گونا به نمونه‌های تحویل‌ناپذیر، متناظر نمایش ایده‌آل آن به‌صورت اشتراک ایده‌آل‌های تجزیه‌نشدنی است. مسألهٔ مربوط به یگانه‌بودن و امکان چنین تجزیه‌ای، یکی از نخستین مسأله‌ها در چندگوناهای جبری و نظریهٔ کلی ایده‌آل‌ها است.

ساختار حلقه‌های جابه‌جایی‌ناپذیر. هر جبر، در عین حال یک حلقه است نسبت به تعریفی که از عمل‌های جمع و ضرب داده شده است. بنابراین، تعداد زیادی از ویژگی‌های مربوط به‌مفهوم‌ها و نتیجه‌گیری‌های نظریهٔ جبر، برای حلقه‌ها هم صادق است. با وجود این، استفاده از نتیجه‌گیری‌های ظریف‌تر جبر و از جمله قضیه *مولین ـ ودربرن* همراه با دشواری‌هایی بود و تنها در سال‌های نزدیک به نیمهٔ سدهٔ بیستم انجام گرفت. مطلب بیش از همه برسر این بود که بتوان چنان تعریفی برای ریشگی (رادیکال) حلقه پیدا کرد که حلقهٔ با ریشگی صفر نوعی شباهت با جبر نیم‌ساده داشته باشد، و در هر حالتی، از قضیه‌های نظریهٔ حلقه‌ها بتوان نظریهٔ ساختاری متناظر آن را در جبر، به‌عنوان حالت‌هایی خاص به‌دست آورد. در نظریهٔ حلقه‌ها، رادیکال‌های مشخصی وجود دارد که امکان می‌دهند، با برخی محدودیت‌ها، نظریهٔ درونی ساختمان حلقه‌های نیم‌ساده را بنا کرد. همان‌طور که پیش از این هم گفته‌ایم، نظریهٔ حلقه‌های جابه‌جایی‌ناپذیر تا اندازهٔ زیادی انگیزه‌ای برای درک اهمیت نظریهٔ حلقه‌های عمل‌گر در آنالیز تابعی بود.

۱۵. مشبکه‌ها

می‌دانیم، مجموعه‌ای را مرتب جزئی گویند که برای برخی از زوج عضوهای آن مشخص

باشد، کدام عضو قبل از دیگری یا به‌دنبال دیگری می‌آید. در ضمن فرض می‌شود که: ۱) هر عضو از خودش پیروی می‌کند یعنی به‌دنبال خودش می‌آید؛ ۲) اگر a از b، و b از a پیروی کند، به‌معنای آن است که a با b یکی است؛ ۳) وقتی a از b و b از c پیروی کند یعنی اگر a به‌دنبال b و b به‌دنبال c باشد)، آن وقت a از c پیروی می‌کند. رابطهٔ پیروی عضوها از یکدیگر را با نماد $\leq$ نشان می‌دهند.

نمونهٔ مهمی از مجموعهٔ مرتب جزئی، عبارت است از دستگاه همهٔ زیرمجموعه‌های یک مجموعه که در آن، رابطهٔ پیروی را به این معنا می‌گیریم که یک زیرمجموعه، جزئی (یا زیرمجموعه‌ای) از زیرمجموعهٔ دیگری است.

اگر رابطهٔ پیروی برای هر دو عضو مجموعهٔ مرتب جزئی معلوم باشد، آن وقت مجموعه را مرتب (خطی) گویند. نمونهٔ مجموعهٔ مرتب، مجموعهٔ عددهای حقیقی است که در آن، رابطهٔ $a\leq b$ به‌معنای آن است که عدد a از b بزرگتر نیست. ولی مجموعهٔ مرتب جزئی، که شامل همهٔ زیرمجموعه‌های یک مجموعه است، به‌شرطی که مجموعهٔ اصلی بیش از یک عضو داشته باشد، مجموعه‌ای مرتب نیست، زیرا زیرمجموعه‌های بدون عضو مشترک را نمی‌توان باهم مقایسه کرد.

فرض کنید عضوهای یک مجموعهٔ مرتب جزئی M دارای این ویژگی باشند که برای هر دو عضو a و b، عضو یگانه c وجود داشته باشد که کوچکترین عضوی باشد که از هردوی a و b بزرگتر است، یعنی به‌نحوی باشند که برای آن‌ها $a\leq c$ ، $b\leq c$ و به ازای هر d از M، اگر شرط $a \leq d$ و $b \leq d$ برقرار باشد، آنگاه داشته باشیم $c \leq d$. در این صورت M را نیم‌مشبکهٔ بالا و عضو c را «مجموع» a و b گویند. به‌سادگی می‌توان قانع شد، این عمل «جمع»، این ویژگی‌ها را دارد:

$$a+b=b+a\ ,\ \ (a+b)+c=a+(b+c)\ ,\ \ a+a=a \qquad (۱۳)$$

جالب این‌جاست که گزارهٔ عکس هم درست است. اگر در مجموعه‌ای، عمل جمع دارای ویژگی‌های (۱۳) باشد، آن وقت اگر بگوییم عضو a از عضو b پیروی می‌کند، به‌شرطی که $a+b=b$، آنگاه یک مجموعهٔ مرتب جزئی به‌دست می‌آید که در آن $a+b$ یگانه عضوی است که کوچکترین عضوی است که از هر دوی a و b بزرگتر است.

به‌همین‌ترتیب می‌توان، با در نظر گرفتن بزرگترین کوچکترها به جای کوچکترین بزرگترها، که در این‌جا «حاصل‌ضرب» عضوهای مفروض نامیده می‌شود، **نیم‌مشبکهٔ پایین** را

تعریف کرد. این عمل هم ویژگی‌هایی شبیه ویژگی «جمع» دارد:

$$ab=ba \quad \text{و} \quad (ab)c=a(bc) \quad \text{و} \quad aa=a \qquad (۱۴)$$

مجموعهٔ مرتب جزئی، وقتی هم نیم‌مشبکهٔ بالا و پایین باشد، آن را **مشبکه** گویند. در هر مشبکه، می‌توان دو عمل تعریف کرد که با شرط‌های (۱۳) و (۱۴) سازگارند. ولی این دو عمل به‌هم بستگی دارند، زیرا رابطهٔ $a \leq b$ را می‌توان در مشبکه، به‌هریک از صورت‌های $a+b=b$ و $ab=a$ نوشت. به‌زبان دیگر، در مشبکه‌ها، برابری‌های $a+b=b$ و $ab=a$ باید هم‌ارز باشند. معلوم می‌شود، شرط اخیر را می‌توان به‌زبان جبری و به صورت

$$a+ab=a \quad \text{و} \quad a(a+b)=a \qquad (۱۵)$$

نوشت و بررسی مشبکه‌ها را منجر به مسألهٔ خالص جبری دربارهٔ دستگاهی شامل دو عمل کرد، که شرط‌های (۱۳) و (۱۴) و (۱۵) پیروی می‌کنند. ارزش جبر برای بررسی مشبکه‌ها در این است که هر مشبکه مشخص را در حالت‌های مختلف به‌گونه‌ای به‌صورت رابطه‌های جبری بین عضوها تبدیل می‌کند و امکان استفاده از نظریهٔ گروه‌ها و نظریهٔ حلقه‌ها را فراهم می‌آورد.

همان‌طور که پیش از این گفتیم، مجموعهٔ همهٔ زیرمجموعه‌های یک مجموعه، یک مجموعهٔ مرتب جزئی است. به‌سادگی دیده می‌شود که، این مجموعه، یک مشبکه است؛ در ضمن در این جا، مجموع مشبکه‌ای عبارت از اجتماع و حاصل‌ضرب مشبکه‌ای، اشتراک زیر مجموعه‌های متناظر است. اگر به‌جای همهٔ زیرمجموعه‌ها، تنها برخی از زیرمجموعه‌ها را در نظر بگیریم، آن وقت می‌توان مشبکه‌های بکلی متفاوتی به‌دست آورد؛ از جمله، مجموعهٔ همهٔ زیرگروه‌ها، همچنین همهٔ زیرگروه‌های نرمال یک گروه دل‌خواه، مجموعهٔ همهٔ زیرحلقه‌ها و مجموعهٔ همهٔ ایده‌آل‌های یک حلقهٔ دل‌خواه و ... مشبکه‌اند. در حالت خاص، در مشبکه‌های همهٔ زیرگروه‌های نرمال یک گروه و همهٔ ایده‌آل‌های یک حلقه، به‌جز اتحادهای اصلی (۱۳)، (۱۴) و (۱۵)، به‌اصطلاح «قانون مدولی» هم برقرار است:

$$a(ab+c)=ab+ac$$

نظریهٔ مشبکه‌ها با قانون مدولی (مشبکه دِدِکنید)، فصل مهمی از نظریهٔ کلی مشبکه‌ها را تشکیل می‌دهد.

تعداد زیادی از قضیه‌های مربوط به نظریهٔ گروه‌ها و نظریه حلقه‌ها دربارهٔ زیرگروه‌ها، زیرگروه‌های نرمال و ایده‌آل را می‌توان طوری بازسازی کرد که به‌عنوان قضیه‌هایی دربارهٔ مشبکه زیرگروه‌ها یا ایده‌آل‌ها به‌کار روند. با برخی محدودیت‌ها، همین قضیه‌ها را برای مشبکه‌ای کلی هم می‌توان به‌کار برد. از این گونه قضیه‌ها را می‌توان در برخی قضیه‌های مهم نظریهٔ گروه‌ها، نظریهٔ حلقه‌ها و دیگر شاخه‌ها پیدا کرد. از طرف دیگر، استفاده از نظریهٔ مشبکه‌ها، برای پیدا کردن ویژگی‌های مشبکه‌های مشخصی، از جمله در نظریهٔ گروه‌ها و نظریهٔ حلقه‌ها، بسیار سودمند است.

نظریهٔ مشبکه‌ها چندان قدیمی نیست و در سال‌های ۲۰ و ۳۰ سدهٔ بیستم، هنوز نتوانسته بود آن اهمیتی را که از جمله نظریهٔ گروه‌ها داشت، کسب کند. ولی به‌تدریج و در سال‌های بعد، اعتبار خود را به‌دست آورد و به‌صورت شاخه‌ای از ریاضیات با مضمونی نیرومند، جای خود را باز کرد.

۱۶. دستگاه‌های جبری کلی

با مفهوم‌هایی که در بند قبل آوردیم، روشن می‌شود، چگونه استفاده از روش‌های جبری در مسأله‌های مختلف، ما را با گسترش دستگاه موضوع‌هایی روبرو می‌کند که در جبر بررسی می‌شود و چگونه مفهوم‌های مربوط به خود روش‌ها و عمل‌های جبری را تعمیم می‌دهد. در این زمینه، پیشرفت روش‌های اصل موضوعی که در کارهای ***لُباچوسکی*** برای پایه‌گذاری هندسه دیده می‌شود، همچنین تکامل نظریهٔ کلی مجموعه‌ها، نقشی اساسی داشته‌اند.

یکی از نتیجه‌های جدی این وضع، شکل‌گیری تدریجی مفهوم‌های کلی عمل‌های جبری، دستگاه‌های جبری و روی هم جمع شدن حقیقت‌های مهمی دربارهٔ دستگاه جبری معین است. به‌جای عمل‌های جبری مشخصی که از دبیرستان با آن‌ها آشنا هستیم و بخش عمدهٔ آن مربوط به عددهاست، در جبر امروز، برای عمل، مفهومی کلی در نظر می‌گیرند. دستگاهی از عضوهای S را مفروض بگیرید و فرض کنید بنابر قانونی، بتوان هر دستگاه a_1 ، a_2 ، ... ، a_m از m عضو S که با ردیف معینی انتخاب شده‌اند، با عنصر مشخص a از همین دستگاه در برابر هم قرار داد. در این صورت می‌گویند، در دستگاه S عملی mتایی داده شده و عنصر a نتیجه‌ای است از این عمل که روی عنصرهای a_1 ، a_2 ، ... ، a_m انجام

شده است. مجموعهٔ عنصرهایی را که یک یا چند عمل روی آن‌ها تعریف شده باشد، دستگاه جبری گویند. یکی از مسأله‌های اساسی جبر، بررسی و تنظیم دستگاه‌های جبری است. البته، مسأله‌ای که به‌این صورت در برابر ما قرار می‌گیرد، ماهیتی بی‌اندازه کلی دارد. تاکنون توانسته‌اند این مسأله را دربارهٔ مهم‌ترین و پرمضمون‌ترین نظریه‌ها، به‌وسیلهٔ دستگاه‌های اختصاصی جبر حل کنند. از جمله، از دستگاه‌های همراه با یک عمل، تنها نظریهٔ گروه‌ها پدید آمده و رشد کرده است (بندهای ۱ تا ۱۰ همین بخش را ببینید)؛ از دستگاه‌های همراه با دو عمل یا تعداد بیشتری عمل، نظریه‌های هیأت، جبر، حلقه و مشبکه اهمیت زیادی دارند. با وجود این، پژوهش و تلاش در زمینهٔ دستگاه‌های جبری مختلف همچنان ادامه دارد. پژوهش در بسیاری از زمینه‌های جبر کلاسیک هم، مثل بررسی همریخت‌ها، دستگاه‌های آزاد، دربارهٔ ریشگی و سایر جنبه‌هایی که به‌نحوی به نظریهٔ کلی دستگاه‌های جبری مربوط می‌شوند، کارهای جدی ادامه دارد؛ به‌نحوی که می‌توان گفت بحث و بررسی دستگاه‌های جبری، به‌عنوان نظریه‌ای تازه، شاخهٔ جدیدی را در جبر پدید آورده است.

وقتی دربارهٔ درون‌مایهٔ دانش‌جبر،به مفهوم کلی آن، صحبت می‌کنیم، اغلب از آن به‌عنوان دانشی که دربارهٔ عدم وجود یا فقدان پیوستگی بحث می‌کند، نام می‌برند و جبر را دانشی که به‌طور عمده دربارهٔ گسستگی‌ها بررسی می‌کند، می‌دانند. بی‌شک، این دیدگاه، یکی از مهم‌ترین ویژگی‌های عینی جبر را منعکس می‌کند. در دنیای واقع، پیوستگی و گسستگی بستگی دیالکتیکی جدی باهم دارند؛ ولی برای شناخت واقعیت‌ها، گاهی لازم است، آن را به بخش‌های جداگانه‌تقسیم و هر بخش‌را جدا از دیگری‌بررسی‌کنیم.بنابراین توجه یک‌سویهٔ جبر را به بررسی بستگی‌های ناپیوسته، نمی‌توان به‌عنوان نارسایی آن به‌حساب آورد.

در نمونهٔ نظریهٔ گروه‌ها دیدیم، نظام‌های جداگانهٔ جبری، نه‌تنها ابزاری برای محاسبه‌های فنی است، بلکه در ضمن زبانی برای بیان قانون‌های ژرف طبیعت است. با این همه، در کنار ارزش‌های عملی مستقیمی که جبر برای فیزیک، شیمی، بلورشناسی و دیگر دانش‌ها دارد، در خود ریاضیات هم، جایگاه شایسته‌ای به‌دست آورده است. اگر از زبان *ن.گ.چه‌بوتارِف*، جبردان مشهور شوروی سخن بگوییم، جبر را باید گهواره‌ای دانست که بسیاری از اندیشه‌ها و مفهوم‌های تازه را در ریاضیات پرورانده و تا اندازهٔ زیادی موجب باروری و تکامل چنان شاخه‌هایی از ریاضیات شده است که مبنایی برای پیشرفت دانش‌های مربوط به فیزیک و صنعت‌اند.

نمایه نام‌های خاص

فهرست عنوانهای جلد اوّل «جوهر، روش و کارایی ریاضیات»

فهرست عنوانهای جلد دوّم «جوهر، روش و کارایی ریاضیات»

عنوان	صفحه

۷۰

شرکت انتشارات فنی ایران

شابک : ۹۶۴-۶۲۳۲-۶۱-۲

ISBN 964-6232-61-2

9789646232617

۲۲۰۰ تومان