

Московский государственный технический университет им. Н. Э. Баумана

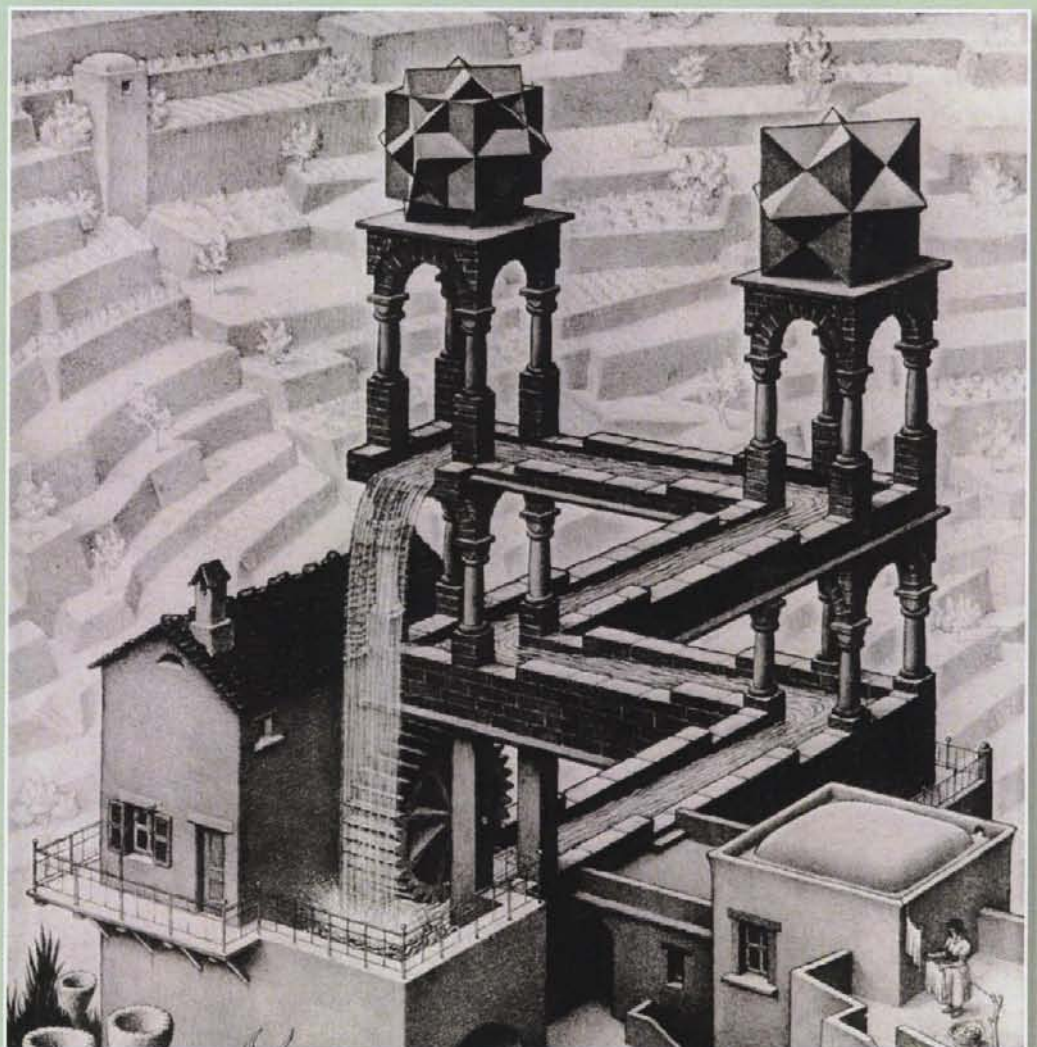
Научно-исследовательский институт
гиперкомплексных систем в геометрии и физике

Международный фонд развития исследований
в области финслеровой геометрии

Материалы Международной школы-семинара

Основы финслеровой геометрии и ее приложения в физике

Москва, 12 июля – 14 августа 2010 года



МГТУ им. Н. Э. Баумана – 2010

УДК 53:514:1

ББК 22.3

О 753

А в т о р ы :

Г.И. Гарасько, С.С. Кокарев, В.Н. Тришин, В. Балан, Н. Бринзей, С.В. Сипаров, В.М. Чернов, В.А. Панчелюга

Рецензент: профессор, д.ф.-м.н., *О.В. Пугачев*

Основы финслеровой геометрии и ее приложения в физике : Материалы
О 753 Международной школы-семинара для старшекурсников, аспирантов физико-математических факультетов и молодых ученых. Москва, 12 июля – 14 августа, 2010 г. Г.И.Гарасько, С.С.Кокарев, В.Н.Тришин, В.Балан, Н.Бринзей, С.В.Сипаров, В.М.Чернов, В.А.Панчелюга / Под общ. ред. Г.Ю.Богословского, В.О.Гладышева, Д.Г.Павлова. – М.: МГТУ им.Н.Э.Баумана, 2010. – 412 с.

ISBN 978-5-7038-3435-0

Данный сборник является подробным конспектом основных курсов лекций, читавшихся слушателям международных школ по основам финслеровой геометрии в 2008-2009 гг., организованных Международным фондом развития исследований в области финслеровой геометрии и НИИ гиперкомплексных систем в геометрии и физике при поддержке научно-учебного комплекса «Фундаментальные науки» МГТУ им. Н.Э. Баумана. Публикуемые материалы также составляют основу для работы IV-й Международной школы "Основы финслеровой геометрии" в 2010 году.

В сборник вошли лекции по основам финслеровой геометрии, элементам геометрии гладких многообразий, основам дифференциальной геометрии и группам Ли, псевдо-финслеровой геометрии, элементам комплексного анализа, современным проблемам ОТО, обобщенным n -арным законам композиции в алгебре и их связь с ассоциированными метрическими формами, а также по экспериментальным исследованиям неоднородности и анизотропности пространства-времени.

Для студентов, аспирантов и молодых специалистов, имеющих физико-математическое образование.

УДК 53:514:1

ББК 22.3

ISBN 978-5-7038-3435-0

© Составление. НИИ ГСГФ, 2010

© Оформление. МГТУ им. Н.Э.Баумана, 2010

Оглавление

Основы финслеровой геометрии (Г.И. Гарасько)	10
1. Координатное пространство	10
2. Элемент длины	11
3. Классическая механика	20
4. Канонические уравнения	24
5. Мировая функция	29
6. Теория поля	33
7. Гиперкомплексные числа	44
8. Поличисла H_3	47
9. Полилинейные формы	52
10. Антиопределители	57
11. Линейные пространства со скалярным полипроизведением	60
12. Норма и группа симметрии поличисел	64
13. Длина отрезка кривой в пространстве P_{k+2-m}	68
14. Экспоненциальное представление невырожденных поличисел	69
15. Функции поличисловой переменной	72
16. Пространство гиперкомплексных чисел H_4	76
17. Конформный потенциал	84
18. Пространство двойных чисел	87
19. Поличисловое пространство H_4	90
Элементы геометрии гладких многообразий (I): производные Ли и их приложения (С.С. Кокарев)	95
1. Введение и обозначения	95
2. Гладкие многообразия	97
3. Скалярные функции и векторные поля	101
4. Касательное и кокасательное пространства. 1-формы	104
5. Тензоры	107
6. Отображения многообразий и геометрических объектов	112
7. Интегральные кривые векторных полей и потоки на многообразиях	115
8. Производная Ли и ее свойства	118
9. Координатные формулы для производной Ли	121
10. Применения производной Ли (1): интегрирование дифференциальных уравнений	124

11. Применения производной Ли (2): тензор деформаций и изометрии многообразия.....	127
12. Применения производной Ли (3): группа движений пространства Минковского.....	130
13. Применения производной Ли (4): изометрии римановых метрик.....	131
14. Применения производной Ли (5): группа конформных симметрий евклидовой и псевдоевклидовой метрик ($n > 2$) ...	134
15. Применения производной Ли (6): группа конформных симметрий евклидовой и псевдоевклидовой плоскости.....	136
16. Применения производной Ли (7): группа изометрий однородных кубических метрик.....	137
16.1. Метрики с $\tau(G) = \tau_0(G) = 1$ (3 типа)	139
16.2. Метрики с $\tau(G) = 2$ (9 типов).....	140
16.3. Метрики с $\tau(G) = 3$ (13 типов)	142
16.4. Метрики с $\tau(G) = 4$ (10 типов)	144
16.5. Метрики с $\tau(G) = 5$ (5 типов).....	145
16.6. Метрики с $\tau(G) = 6$	146
16.7. Метрики с $\tau(G) = 6, 7, 8, 9, 10$	147
16.8. Инвариантная классификация метрик с нетривиальными изометриями	147
16.9. Связь с проективной классификацией кубичных форм.....	151
Основы дифференциальной геометрии и групп Ли (В.Н. Тришин)	155
1. Введение	155
2. Топологические пространства	155
2.1. Алгебраическая топология. Гомотопические группы.....	157
2.2. Гомологии	158
2.3. Гладкие многообразия.....	161
2.4. Определение многообразия. Гладкая структура	161
2.5. Векторные поля и дифференциальные формы. Когомологии де Рама.....	162
3. Группы Ли.....	167
4. Геометрия расслоений и калибровочные теории.....	171
4.1. Расслоения. Векторные и главные расслоения.....	171
4.2. Связности на расслоениях	180
4.3. Характеристические классы	186
5. Пространство струй и геометрия дифференциальных уравнений	189
6. Дополнительные структуры на многообразиях.....	192

6.1. Риманова метрика в теории гравитации	192
6.2. Финслеровы многообразия	194
6.3. Симплектическая и пуассонова структуры	195
6.4. Комплексные многообразия	197

Pseudo-Finsler Geometry.

Symmetries and transformation groups (<i>Vladimir Balan</i>)	200
1. Pseudo-Finsler Geometry	200
1.1. Finsler structures. Finsler metric tensor field.....	200
1.2. The non-linear connection of a Finsler structure	201
1.3. Linear connections and covariant derivation laws	201
1.4. Invariants of a Finsler structure	203
1.5. Examples of Finsler and pseudo-Finsler structures	203
1.6. Geodesics in (pseudo-)Finsler spaces.....	205
1.7. The flag curvature of a Finsler space	205
1.8. The Legendre transform of a Finsler structure	206
1.9. Exam quizzes - part I	207
2. Symmetries and transformation groups.....	207
2.1. Lie algebras and Lie groups	207
2.2. Homomorphisms of Lie groups.....	208
2.3. Lie algebras of Lie groups	209
2.4. Lie transformation groups	209
2.4.1. Transformation groups in the 2D plane.....	209
2.4.2. The isometries of R^n	210
2.4.3. Lorentz transformations	211
2.5. Complex transformation groups	211
2.6. Group actions, orbits and stabilizers.....	212
2.7. Symmetric spaces	213
2.8. Notable examples of actions	213
2.9. Infinitesimal isometries (Killing vector fields)	214
2.10. Groups of automorphisms of geometric structures	214
2.11. Invariant 4-tensors in R^4	214
2.12. G-structures associated to tensors	214
2.13. Riemannian isometries	215
2.14. Finslerian isometries	215
2.15. Seminar addenda I. The LU decomposition algorithm.....	215
2.16. Seminar addenda II. The QR decomposition algorithm	216

2.17. Exam quizzes - part II.....	217
Elements of Complex Analysis (<i>Nicoleta Brinzei</i>)	222
1. Complex Numbers	222
1.1. Definition and Operations	222
1.2. Polar Form.....	225
2. Complex Functions of a Real Variable.....	226
3. Functions of a Complex Variable	228
3.1. Limit. Continuity. Differentiability.....	228
3.2. Cauchy-Riemann Equations	232
3.3. Harmonic Functions.....	232
4. Elementary Functions.....	234
4.1. Polynomial and Rational Functions.....	234
4.2. Exponential Function.....	234
4.3. Trigonometric Functions.....	235
4.4. Complex Logarithm	236
4.5. Complex Powers.....	237
4.6. Branches of Multi-Valued Functions. Riemann Surfaces	238
5. Conformal Transformations	241
5.1. Definition and Properties.....	241
5.2. MHobius Transformations	243
6. Plane Vector Fields. Complex Potential	246
6.1. Plane Vector Fields.....	246
6.2. Complex Potential	248
6.3. Applications.....	250
6.3.1. In Electrostatics	250
6.3.2. In Fluid Mechanics	253
6.3.3. In Thermodynamics.....	256
6.4. The Use of Conformal Transformations	256
6.5. Other Examples	258
7. Integration.....	260
7.1. Definition and Basic Properties.....	260
7.2. Antiderivatives	262
7.3. Cauchy's Theorem	262
7.4. Cauchy's Integral Formula.....	263
7.5. Applications of Cauchy's Formula	266
8. Series. Power Series	269
8.1. Sequences	269

8.2. Series	271
8.3. Taylor Series	274
9. Laurent Series. Residues	275
9.1. Laurent Series	275
9.2. Residues	278
9.3. Classification of Isolated Singularities	280
Современные проблемы ОТО – физика и геометрия (С.В. Сипаров)	283
1. Введение: взаимоотношение математики и физики	283
1.1. Наука от Пифагора до Кеплера и Ньютона	283
1.2. Кривизна, силы, поля	285
1.3. Выводы	286
2. Некоторые геометрические аспекты общей теории относительности Эйнштейна ...	283
2.1. Тензоры	283
2.1.1. Фундаментальный тензор	283
2.1.2. Норма, метрика, геометрия	289
2.1.3. Построение тензоров	289
2.2. “Тождества Максвелла” и уравнения Максвелла	289
2.3. Вариационный принцип, простейший скаляр, геодезическая, принцип эквивалентности	291
2.4. Пример: построение основ. ОТО Эйнштейна (1916)	293
2.5. Закон гравитации Ньютона как приближенное решение уравнений ОТО ...	294
2.6. Классические тесты ОТО (красное смещение, отклонение луча, прецессия орбиты)	295
2.6.1. Геометрия	295
2.6.1. Течение времени	295
2.6.1. Отклонение луча света	295
2.6.1. Прецессия орбиты Меркурия	295
2.7. Точное решение Шварцшильда (1916)	297
2.8. ОТО и космология	298
2.9. Наблюдательные проблемы и “темные” понятия	300
2.10. Попытки модификации ОТО – классический (геометрический) подход	305
2.11. Анизотропная геометродинамика – наводящие соображения и первые шаги	307
2.11.1. Геодезическая	308
2.11.2. Снова “тождества Максвелла”	309
2.11.3. Гравитационная сила и ее компоненты	311
2.11.4. Кривые вращения и закон Талли-Фишера	312
2.12. Интерпретация известных наблюдений, космологические следствия и специфический эффект	314

Приложение 1	
Метрический тензор и метризация	316
Приложение 2	
Принцип наименьшего действия	319
Постулаты, определения, законы	319
Принцип наименьшего действия и канонические уравнения	323
Приложение 3	
Задача Эйнштейна	328
Приложение 4	329
Оптико-метрический параметрический резонанс как инструмент исследования геометрических свойств пространства-времени	329
Изотропное возмущение метрики Минковского	331
Анизотропное возмущение метрики Минковского	332
Обобщенные n-арные законы композиции в алгебре и их связь с ассоциированными метрическими формами (В.М. Чернов)	334
1. Введение	334
1.1. Классические композиционные алгебры и алгебры H_n	334
1.2. Основные идеи и определения	337
2. Обобщенные n-арные законы композиции в алгебре $R \oplus R \oplus R \oplus R$	341
2.1. Обобщенный кватернарный закон композиции в алгебре $R \oplus R \oplus R \oplus R$	341
2.2. Обобщенный тернарный закон композиции в алгебре $R \oplus R \oplus R \oplus R$	342
2.3. Обобщенный бинарный закон композиции в алгебре $R \oplus R \oplus R \oplus R$	343
3. Обобщения и открытые проблемы	344
Экспериментальные исследования и астрофизические наблюдения, свидетельствующие о неоднородности и анизотропии реального пространства-времени. (В.А. Панчелюга)	346
Введение	347
1. Об особенностях во временных рядах флуктуаций результатов измерений	347
1.1. Космофизическая обусловленность флуктуаций зрительного порога человека.	349
1.2. Н.А. Козырев и последователи.	354
1.3. В.А. Дубровский: спектры микросейсм и распределение звездных объектов.	355
1.4. Связь флуктуаций и выбросов в различных процессах с ходом λ_D – функции.	357

1.5. Работы Н.В. Клочека и сотрудников: Т-сигнал и его свойства.....	361
1.6. Исследования В.Н. Смирнова: детектор гравитационных воздействий.....	364
1.7. К.А. Труханов, Ю.А. Бауров: векторный космологический потенциал.	368
1.8. А.Г. Пархомов: исследование ритмов и флуктуаций в ходе различных процессов.....	371
1.8.1. Генераторы с кварцевыми резонаторами.	372
1.8.2. Инфранизкочастотные флуктуации в проводниках.....	373
1.8.3. Скорость счета α -частиц.	374
1.8.4. Скорость счета β -частиц.....	375
2. Феномен макроскопических флуктуаций и его основные свойства.	377
2.1. Феномен макроскопических флуктуаций. Краткая история.	378
2.2. Исходный экспериментальный материал для исследования феномена макроскопических флуктуаций.....	380
2.3. Методика обработки экспериментальных данных.	381
2.4. Построение распределения интервалов.	382
2.5. Основные свойства феномена макроскопических флуктуаций.	385
2.5.1. Эффект ближней зоны.	385
2.5.2. Универсальность феномена макроскопических флуктуаций.	385
2.5.3. Периодичность проявления феномена макроскопических флуктуаций.	386
2.5.4. Эффект местного времени.	387
2.5.4.1. Фрактальная структура пика местного времени.	388
2.5.4.2. Азимутальная выраженность эффекта местного времени [125-128.].	392
2.5.4.3. Эффект местного времени для движущихся источников флуктуаций [116, 130.].	397
2.5.5. Исчезновение суточных периодов при измерениях вблизи Северного полюса.	399
2.5.6. О зависимости результатов измерений от направления в пространстве	400
2.5.6.1. Неподвижный коллиматор направленный на Полярную звезду.	400
2.5.6.2. Опыты с неподвижными коллиматорами, направленными на восток и на запад.	400
2.5.6.3. Опыты с вращаемыми коллиматорами.	400
2.5.7. Выделенные формы гистограмм: новолуния и солнечные затмения.	402
2.5.8. Выделенные формы гистограмм: восходы и заходы Солнца и Луны.	402
2.5.9. Зеркальная симметрия гистограмм.	402

Основы финслеровой геометрии

Г.И. Гарасько

ГУП Всероссийский электротехнический институт, Москва, Россия.

E-mail: gri9z@mail.ru

1 Координатное пространство

Для применения математики в физических исследованиях и решении теоретических задач вначале необходимо перейти от физических понятий к математическим, например, ввести систему координат и описывать физическую систему геометрическими, алгебраическими или какими-либо иными математическими объектами в этом координатном пространстве. Будем считать, что самая трудная часть работы проделана, то есть для физической системы установлены система координат. Круг задач ограничен свойствами этого координатного пространства и свойствами тех объектов, которые это пространство содержит или которыми оно будет наполнено в результате построений. Важно подчеркнуть, что координаты могут иметь разные размерности.

Можно работать в некотором координатном пространстве, даже если не установлено взаимно однозначное соответствие между точками координатного пространства и событиями, но надо уметь ставить в соответствие теоретически получаемым результатам физически наблюдаемые и измеримые экспериментально величины, то есть действительные числа. Такую процедуру естественно назвать *интерпретацией*. Вполне бывает достаточно существование хотя бы частичной интерпретации. В таком подходе можно заниматься теоретическими и математическими исследованиями и лишь при получении некоторых результатов ставить вопрос о их интерпретации.

Все последующее будет развёртываться в непрерывном координатном пространстве n измерений, каждая точка $M \equiv M(x^1, x^2, \dots, x^n)$ которого это упорядоченный набор значений n независимых действительных переменных

$x^1, x^2, \dots, x^n$. Такое пространство будем называть *основным координатным пространством*, или просто *основным пространством*. При этом взамен переменных $x^1, x^2, \dots, x^n$ разрешается брать любые другие независимые переменные $x^{1'}, x^{2'}, \dots, x^{n'}$, при условии, чтобы между ними имелось взаимно однозначное соответствие и чтобы новые переменные были достаточно число раз дифференцируемыми функциями старых, ровно как и старые переменные - такого же рода функциями новых. Функциональную зависимость между координатами одной и другой системы

$$x^{i'} = f_i(x^1, x^2, \dots, x^n) \quad (1)$$

будем называть преобразованием координат. Скалярные функции $f_1, f_2, \dots, f_n$ дифференцируемы достаточно число раз и обратимы, поэтому якобиан

$$\frac{D(f_1, f_2, \dots, f_n)}{D(x^1, x^2, \dots, x^n)} \neq 0, \pm\infty. \quad (2)$$

Каждому преобразованию координат (1) соответствует преобразование самого пространства

$$x^i \rightarrow f_i(x^1, x^2, \dots, x^n), \quad (3)$$

то есть взаимно однозначное отображение координатного пространства самого на себя.

Важно отметить, что как при пассивном взгляде на преобразования (1), точки пространства остаются на месте, а меняются только численные значения их координат, так и в активном подходе (3), каждой точке пространства ставится в соответствие, вообще говоря, другая точка того же пространства в тех же координатах, в физических задачах условие (2) часто не выполняется во всём пространстве, а лишь в некоторой его области или областях, или же во всём пространстве, но за исключением особых точек. Для краткости изложения мы не всегда будем это оговаривать, но всегда будем иметь это в виду.

Рассмотрим какую-нибудь точку $M(x^1, x^2, \dots, x^n)$ нашего координатного пространства. Будем говорить, что в этой точке задан вектор ξ , если в каждой системе координат $x^1, x^2, \dots, x^n$ задана система действительных чисел $(\xi^1, \xi^2, \dots, \xi^n)$, преобразующихся при переходе от одной системы координат к другой $x^{1'}, x^{2'}, \dots, x^{n'}$ по закону

$$\xi^{i'} = \frac{\partial x^{i'}}{\partial x^i} \xi^i. \quad (4)$$

Совокупность всевозможных векторов ξ в данной точке M вместе с операциями покомпонентного сложения двух векторов и покомпонентного умножения вектора на действительное число образуют линейное пространство n измерений. Это пространство будем называть *касательным пространством* в точке M .

Пусть ε - бесконечно малая величина, не зависящая от выбора системы координат. Рассмотрим векторы ξ касательного пространства в точке M , умноженные на ε . Каждый такой вектор имеет бесконечно малые компоненты. Сопоставим вектору $\varepsilon\xi$ бесконечно малое смещение из точки $M(x^1, x^2, \dots, x^n)$ в бесконечно близкую точку $M'(x^1 + dx^1, x^2 + dx^2, \dots, x^n + dx^n)$ так, чтобы дифференциалы координат dx^i совпадали с координатами вектора $\varepsilon\xi$:

$$dx^i = \varepsilon \xi^i. \quad (5)$$

Смещение MM' , построенное по закону (5), не зависит от выбора системы координат, так как закон преобразования дифференциалов dx^i в точности таков, как и закон преобразования (4) компонент вектора $\varepsilon\xi$ при переходе от одной системы координат к другой.

Геометрический смысл касательного пространства: бесконечно малые векторы касательного пространства в точке M реализуются как бесконечно малые смещения MM' в основном пространстве.

Всегда будем подразумевать, что основное координатное пространство в каждой точке снабжено касательным пространством.

Среди всех таких основных пространств для физиков на первом месте, конечно, стоит четырёхмерное пространство событий, то есть четырёхмерное пространство-время x^0, x^1, x^2, x^3 .

2 Элемент длины

Пусть в каждой точке $M(x^1, x^2, \dots, x^n)$ основного координатного пространства и для всех или некоторых бесконечно малых смещений dx из точки M в бесконечно близкую точку $M'(x + dx)$ определено бесконечно малое действительное неотрицательное число - элемент длины, то есть расстояние $dl = L(dx; x)$ от точки M до всех или только некоторых соседних бесконечно близких точек M' . Тогда такое пространство будем называть *метрическим*

пространством, а функцию $L(dx; x)$ - метрической функцией. В общем случае метрическая функция содержит $2n$ действительных аргументов. Элемент длины может иметь свою размерность, не совпадающую с размерностью ни одной из координат, например, координата x^1 может иметь размерность времени [сек], координаты $x^2, \dots, x^n$ - размерность длина [см], а величина l ("длина") может иметь размерность действия [эрг · сек].

Если в метрическом пространстве найдётся такая система координат $x^1, x^2, \dots, x^n$, в которой элемент длины dl не зависит от точки пространства $dl = L(dx^i)$, то такое пространство называется *плоским метрическим пространством*.

Иногда удобно считать, что начала всех векторов касательного пространства, относящегося к точке M основного пространства, находятся в одной и той же точке, которую назовём центром, причём можно представлять, что центр совпадает с точкой M основного пространства. Такое касательное пространство $\mathcal{A}_n(M)$ называется *центроаффинным касательным пространством* в точке M , при этом векторы ξ в нём являются радиус-векторами, а их компоненты (координаты) $\xi^1, \xi^2, \dots, \xi^n$ совпадают с координатами точек этого центроаффинного касательного пространства.

Используя метрическую функцию, запишем уравнение специальной гиперповерхности в соответствующем центроаффинном касательном пространстве

$$L(\xi; x) = l_1, \quad (6)$$

где l_1 - единица длины. Такую гиперповерхность будем называть *индикатрисой*, соответствующее уравнение (6) - уравнением индикатрисы, а удовлетворяющие уравнению (6) векторы ξ_{ind} - единичными векторами.

Геометрический смысл: индикатриса суть геометрическое место концов единичных радиус-векторов соответствующего центроаффинного касательного пространства.

Евклидовым пространством размерности n называется плоское метрическое пространство, для которого найдётся такая система координат $x^1, x^2, \dots, x^n$, что элемент длины в ней вычисляется по формуле

$$dl = \sqrt{(dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2}. \quad (7)$$

Системы координат, в которых элемент длины в евклидовом пространстве имеет такой вид, называют декартовыми прямоугольными. Таким образом, евклидовое пространство - это такое метрическое пространство, для которого найдётся такая система координат, что все индикатрисы будут сферическими гиперповерхностями одного и того же радиуса.

В римановом пространстве $x^1, x^2, \dots, x^n$ элемент длины выражается формулой

$$dl = \sqrt{g_{ij}(x^1, x^2, \dots, x^n) dx^i dx^j}, \quad (8)$$

где $g_{ij}(x)$ - метрический тензор риманова пространства. Плоское риманово пространство является евклидовым. Все индикатрисы риманова пространства суть эллипсоиды, у которых параметры зависят от точки основного пространства. Фундаментальное свойство римановых и, в частности, евклидовых пространств состоит в том, что все индикатрисы в них это замкнутые выпуклые гладкие гиперповерхности, поэтому для бесконечно малого смещения по любому направлению определена длина.

В четырёхмерном пространстве-времени $x^0 \equiv ct, x^1, x^2, x^3$ классической СТО определена псевдоевклидова геометрия с элементом длины

$$dl = \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2}. \quad (9)$$

Индикатрисы для такой геометрии - трехмерные двуполостные гиперboloиды, каждая односвязная область которых есть вогнутая гиперповерхность. Такое метрическое пространство принято называть пространством Минковского. Если потребовать, чтобы приращение временной координаты было всегда положительной величиной $dx^0 > 0$, то получившееся метрическое пространство будет иметь в каждой точке основного пространства индикатрису, состоящую только из одной полости двуполостного гиперboloида, расположенной в конусе будущего.

В четырёхмерном пространстве событий ОТО, которое является псевдоримановым пространством с сигнатурой $(1, -1, -1, -1)$, определен элемент длины

$$dl = \sqrt{g_{ij}(x^0, x^1, x^2, x^3)dx^i dx^j}, \quad (10)$$

с метрическим тензором $g_{ij}(x^0, x^1, x^2, x^3)$, компоненты которого зависят от точки пространства, причем для каждой точки всегда найдётся такая система координат $x^{0'}, x^{1'}, x^{2'}, x^{3'}$, в которой в данной точке

$$(g_{i'j'}) = \text{diag}(1, -1, -1, -1) \equiv (\overset{\circ}{g}_{i'j'}). \quad (11)$$

Индикатрисы в таком пространстве являются вогнутыми гиперповерхностями, трёхмерными двуполостными гиперboloидами, параметры которых зависят от точки пространства. Плоское четырёхмерное псевдориманово пространство с сигнатурой $(1, -1, -1, -1)$ суть пространство Минковского.

Метрические функции, стоящие справа в формулах (7) - (10), выражаются через квадратичные формы дифференциалов координат. Это отличительное свойство именно евклидовых, псевдоевклидовых, римановых и псевдоримановых пространств, которые кратко можно называть *квадратичными геометриями*, или *квадратичными пространствами*. Отказавшись от этого свойства, получим геометрии (пространства), которые естественно назвать *неквадратичными*.

Обратим внимание на то, что не для всякого бесконечно малого смещения в пространстве Минковского в частности, а в общем случае для любого псевдориманово пространства, элемент длины определён, то есть соответствующий вектор бесконечно малого смещения измерим. Подставим в правую часть формулы (9) бесконечно малое смещение

$$(dx^0, dx^1, dx^2, dx^3) = (1, 1, 1, 1) \cdot \varepsilon, \quad (12)$$

где $\varepsilon > 0$ - бесконечно малая действительная положительная величина. Так как элемент длины должен быть действительным неотрицательным числом, то элемент длины для такого бесконечно малого смещения для данной метрической функции не определен. В неквадратичных пространствах аналогичная ситуация также возможна.

Пусть два бесконечно малых смещения сонаправлены, то есть

$$\delta x^i = a dx^i, \quad (13)$$

где $a > 0$ - действительное положительное число. Тогда подставив эти смещения в метрические функции, стоящие справа в формулах (7) - (10), получим

$$L(\delta x^1, \dots, \delta x^n; x^1, \dots, x^n) = aL(dx^1, \dots, dx^n; x^1, \dots, x^n), \quad (14)$$

то есть рассматриваемые метрические функции являются положительно однородными первой степени по первым n аргументам, и если смещение dx измеримо, то и смещение δx измеримо, и наоборот. Это свойство, которое мы сохраним при переходе к неквадратичным геометриям, можно сформулировать иначе: длина суммы двух сонаправленных бесконечно малых смещений есть сумма их длин. Таким образом, задание метрической функции в общем случае вырезает в пространстве первых n аргументов некий конус (или конусы) бесконечно малых смещений вместе со всеми им сонаправленными векторами, для которых элемент длины определен. В квадратичных пространствах без дополнительных ограничений направления измеримых векторов в касательных пространствах, как и во многих неквадратичных пространствах, граница таких конусов измеримых векторов состоит из векторов, имеющих нулевую длину и носящих название *изотропных векторов* или *изотропных направлений*, а сама граница в таких случаях называется *изотропным конусом*. В физике изотропные направления обычно связывают с направлениями распространения электромагнитных волн и направлениями мировых линий элементарных частиц, не имеющих массы покоя и движущихся со скоростью света.

Если вектор dx измерим, то $L(dx; x)$ - расстояние от точки $M(x^1, x^2, \dots, x^n)$ до точки $M'(x^1 + dx^1, x^2 + dx^2, \dots, x^n + dx^n)$, поэтому при обобщении вида метрических функций следует сохранить свойство:

$$L(0, 0, \dots, 0; x^1, x^2, \dots, x^n) = 0, \quad (15)$$

которое выполняется для приведённых выше метрических функций квадратичных пространств и которое означает, что индикатриса не может проходить через центр центроаффинного касательного пространства. Для неквадратичных пространств подстановка в метрическую функцию нулевого смещения может приводить к неопределённости, поэтому нулевое смещение следует понимать как смещение $(\delta x^1, \dots, \delta x^n)$ (13) при $a \rightarrow 0$ и условии, что для бесконечно малого смещения $(dx^1, \dots, dx^n)$ элемент длины определен.

Частные производные

$$p_i \equiv \frac{\partial L(\xi; x)}{\partial \xi^i} \quad (16)$$

метрической функции по первым n аргументам образуют ковариантный тензор, называемый обобщённым импульсом. Обязательно существует хотя бы одна функциональная зависимость

$$\Phi(p_1, p_2, \dots, p_n; x^1, x^2, \dots, x^n) = 0, \quad (17)$$

так как частные производные (16) являются однородными функциями ξ нулевой степени, то есть зависят не более, чем от $(n - 1)$ -ой независимой величины, например, от отношений компонент вектора ξ , а число самих компонент p_i равно n . Будем предполагать всегда, что такая функциональная зависимость только одна. Эту единственную функциональную зависимость принято называть *тангенциальным уравнением индикатрисы*, а функцию $\Phi(p; x)$ в таком уравнении - функцией Финслера.

Запишем тангенциальные уравнения индикатрисы для метрических пространств, рассмотренных выше.

Евклидово пространство. Метрическая функция в декартовых прямоугольных координатах n -мерного евклидова пространства записана в правой части формулы (7),

$$L(dx) = \sqrt{(dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2}. \quad (18)$$

Тогда компоненты обобщённого импульса определяются по формулам:

$$p_i = \frac{dx^i}{\sqrt{(dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2}}. \quad (19)$$

Очевидно, что они функционально зависимы,

$$p_1^2 + p_2^2 + \dots + p_n^2 - 1 = 0, \quad (20)$$

и эта функциональная зависимость единственная, хотя функция Финслера при этом может иметь разный вид, например:

$$\Phi'(p) = p_1^2 + \dots + p_n^2 - 1, \quad \Phi''(p) = \sqrt{p_1^2 + \dots + p_n^2} - 1. \quad (21)$$

Обе эти функции Финслера относятся к одной и той же евклидовой геометрии с метрической функцией (18).

Риманово пространство. Метрическая функция n -мерного риманова пространства в произвольных координатах записана в правой части формулы (8),

$$L(dx; x) = \sqrt{g_{ij}(x^1, x^2, \dots, x^n) dx^i dx^j}. \quad (22)$$

Компоненты обобщённого импульса определяются по формулам:

$$p_i = \frac{g_{ij}(x) dx^j}{\sqrt{g_{kl}(x) dx^k dx^l}}. \quad (23)$$

Если $\det(g_{ij}(x)) \neq 0$, то можно построить тензор с двумя верхними индексами $g^{ij}(x)$, при этом компоненты обобщённого импульса связаны единственной функциональной зависимостью

$$g^{ij}(x) p_i p_j - 1 = 0. \quad (24)$$

Функция Финслера при этом может иметь разный вид, например:

$$\Phi'(p) = g^{ij}(x) p_i p_j - 1, \quad \Phi''(p) = \sqrt{g^{ij}(x) p_i p_j} - 1. \quad (25)$$

Обе эти функции Финслера относятся к одной и той же римановой геометрии с метрической функцией (22).

Пространство Минковского. Метрическая функция в специальных координатах четырёхмерного пространства Минковского записана в правой части формулы (9),

$$L(dx) = \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2} \equiv \sqrt{{}^o g_{ij} dx^i dx^j}. \quad (26)$$

Тогда компоненты обобщённого импульса определяются по формулам:

$$p_i = \frac{{}^o g_{ij} dx^j}{\sqrt{{}^o g_{kl} dx^k dx^l}}. \quad (27)$$

Компоненты обобщённого импульса функционально зависимы,

$$p_0^2 - p_1^2 - p_2^2 - p_3^2 - 1 = 0, \quad (28)$$

и эта функциональная зависимость единственная, хотя функция Финслера при этом может иметь разный вид, например:

$$\Phi'(p) = p_0^2 - p_1^2 - p_2^2 - p_3^2 - 1, \quad \Phi''(p) = \sqrt{p_0^2 - p_1^2 - p_2^2 - p_3^2} - 1. \quad (29)$$

Обе эти функции Финслера относятся к одной и той же геометрии Минковского с метрической функцией (26).

Рассмотрим метрическую функцию, лишь числовым множителем mc , где m - масса покоя материальной частицы, а c - скорость света, отличающуюся от метрической функции (26) пространства Минковского,

$$L(dx) = mc \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2}. \quad (30)$$

Такая метрическая функция, взятая со знаком минус, совпадает с элементом действия для свободной релятивистской частицы массой m , то есть действием между двумя бесконечно близкими событиями [11]. Компоненты обобщённого импульса в этом случае определяются по формулам:

$$p_i = mc \frac{{}^o g_{ij} dx^j}{\sqrt{{}^o g_{kl} dx^k dx^l}}. \quad (31)$$

Они функционально зависимы,

$$p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2 c^2 = 0, \quad (32)$$

и эта функциональная зависимость единственная. В качестве функции Финслера можно взять, например, одну из функций:

$$\Phi'(p) = p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2 c^2, \quad \Phi''(p) = \sqrt{p_0^2 - p_1^2 - p_2^2 - p_3^2} - mc. \quad (33)$$

Обе эти функции Финслера относятся к одной и той же геометрии с метрической функцией (30).

Псевдориманово пространство. Для псевдориманово n -мерного пространства формально справедливы все формулы (22) - (25), если метрический тензор $g_{ij}(x)$ невырожден, то есть $\det(g_{ij}(x)) \neq 0$.

Метрическое пространство будем называть *финслеровым*, если метрическая функция $L(\xi^1, \dots, \xi^n; x^1, \dots, x^n) \geq 0$ является положительно однородной первой степени по первым n аргументам и если

$$\text{rang} \left(\frac{\partial^2 L(\xi; x)}{\partial \xi^i \partial \xi^j} \right) = n - 1. \quad (34)$$

Именно условие (34) гарантирует единственность функциональной зависимости между компонентами p_i обобщённого импульса. Приведём пример метрического пространства y^1, y^2, y^3, y^4 , которое не является финслеровым, так как между компонентами обобщённого импульса в этом пространстве имеются две функциональные зависимости, а не одна, то есть условие (34) не выполняется.

Пусть в пространстве y^1, y^2, y^3, y^4 определена метрическая функция $L(dy)$ следующего вида:

$$L(dy) = \sqrt{(dy^1)^2 - (dy^2)^2} + \sqrt{(dy^3)^2 - (dy^4)^2}.$$

Тогда компоненты обобщённого импульса определяются формулами:

$$p_1 = \frac{dy^1}{\sqrt{(dy^1)^2 - (dy^2)^2}}, \quad p_2 = -\frac{dy^2}{\sqrt{(dy^1)^2 - (dy^2)^2}},$$

$$p_3 = \frac{dy^3}{\sqrt{(dy^3)^2 - (dy^4)^2}}, \quad p_4 = -\frac{dy^4}{\sqrt{(dy^3)^2 - (dy^4)^2}}$$

- и между ними имеют место две функциональные зависимости:

$$p_1^2 - p_2^2 - 1 = 0, \quad p_3^2 - p_4^2 - 1 = 0.$$

Таким образом, метрическое пространство y^1, y^2, y^3, y^4 не является финслеровым, так как условие (34) не выполняется.

Важно подчеркнуть, что центроаффинное пространство $\mathcal{A}_n(M)$ в любой фиксированной точке M финслерова пространства является плоским метрическим пространством. Расстояние l_{21} между двумя точками

$R_{(1)}(\xi_{(1)}^1, \xi_{(1)}^2, \dots, \xi_{(1)}^n)$ и $R_{(2)}(\xi_{(2)}^1, \xi_{(2)}^2, \dots, \xi_{(2)}^n)$ в нём определяется метрической функцией основного метрического

пространства, последние n аргументов которой суть координаты фиксированной точки M :

$$l_{21} = L(\xi_{(2)}^1 - \xi_{(1)}^1, \dots, \xi_{(2)}^n - \xi_{(1)}^n; x^1, \dots, x^n). \quad (35)$$

Длина $|\xi|$ вектора или радиус-вектора ξ вычисляется соответственно по формуле

$$|\xi| = L(\xi^1, \dots, \xi^n; x^1, \dots, x^n). \quad (36)$$

Длину радиус-вектора, а значит и расстояние между двумя точками центроаффинного пространства можно вычислить, не зная метрической функции, если в центроаффинном пространстве построена (задана) индикатриса. Вектор ξ определяет луч, выходящий из центра,

если такой луч индикатрису не пересекает, то будем считать, что длина данного вектора и всех ему сонаправленных не определена, то есть они неизмеримы. Пусть вектор ξ измерим. Найдём точку пересечения этого луча с индикатрисой, тем самым построим радиус-вектор $\xi_{(ind)}$, который имеет смысл вектора единичной длины в данном направлении. Этот вектор сонаправлен вектору ξ , то есть существует такое действительное неотрицательное число l , что

$$\xi^i = l \cdot \xi_{(ind)}^i, \quad (37)$$

тем самым длина вектора ξ определяется геометрией основного пространства и выражается формулой

$$|\xi|_{fins} = l. \quad (38)$$

Если представить, что пространство $\mathcal{A}_n(M)$ является евклидовым, а координаты $\xi^1, \xi^2, \dots, \xi^n$ - декартовы прямоугольные, то из формулы (37) следует,

$$|\xi|_{fins} = \frac{|\xi|_{ev}}{|\xi_{ind}|_{ev}}, \quad (39)$$

где ξ_{ind} - единичный вектор, сонаправленный вектору ξ , а $|\xi|_{ev}$, $|\xi_{ind}|_{ev}$ - длины векторов ξ , ξ_{ind} в евклидовом смысле.

Итак, римановы и псевдоримановы пространства с метрическими тензорами, удовлетворяющими условию

$$\det(g_{ij}) \neq 0,$$

являются финслеровыми пространствами.

Приведем несколько примеров финслеровых пространств, которые не являются квадратичными.

В четырёхмерном пространстве Галилея (t, x, y, z) , где t - время, а x, y, z - пространственные декартовы прямоугольные координаты, с материальной частицей массы m в потенциальном поле $U(t, x, y, z)$ можно определить следующую метрическую функцию:

$$L(dt, dx, dy, dz; x, y, z) = mc^2 dt - m \cdot \frac{(dx)^2 + (dy)^2 + (dz)^2}{2dt} + U dt. \quad (40)$$

Она является положительно однородной первой степени по первым четырём аргументам и определяет в пространстве (t, x, y, z) неквадратичную метрическую геометрию. Так как потенциал $U(t, x, y, z)$ может в общем случае принимать отрицательные значения, конус бесконечно малых смещений, для которых элемент длины определен, лучше задать двумя неравенствами:

$$mc^2(dt)^2 - m \cdot \frac{(dx)^2 + (dy)^2 + (dz)^2}{2} > 0, \quad (41)$$

$$(mc^2 + U)(dt)^2 - m \cdot \frac{(dx)^2 + (dy)^2 + (dz)^2}{2dt} > 0, \quad (42)$$

что позволяет гарантированно "включать" и "выключать" потенциал $U(t, x, y, z)$. Такое финслерово пространство соответствует классической механике одной нерелятивистской частицы: скорость частицы всегда много меньше скорости света, а потенциальная энергия много меньше внутренней энергии mc^2 . Из этих условий классической нерелятивистской механики автоматически следует выполнение неравенств (41), (42).

Остановимся более подробно на связи финслеровой геометрии с метрической функцией (40), и лагранжевым формализмом для описания движения нерелятивистской частицы в потенциальном поле [10]. Будем считать, что $dt > 0$, вынесем dt в левой части формулы (40) из метрической функции, а в правой части за скобки и сократим левую и правую части на dt , в результате получим:

$$L(1, v_x, v_y, v_z; x, y, z) = mc^2 - m \cdot \frac{v_x^2 + v_y^2 + v_z^2}{2} + U, \quad (43)$$

то есть связь между функцией Лагранжа $L'(v_x, v_y, v_z; x, y, z)$ и метрической функцией L (40) определяется формулой

$$L'(v_x, v_y, v_z; x, y, z) = mc^2 - L(1, v_x, v_y, v_z; x, y, z). \quad (44)$$

Функцию Лагранжа можно умножить на любое действительное число, в данном случае это $-(-1)$, и прибавить полную производную по времени любой функции времени и координат (в данном случае это $mc^2 \cdot t$), при этом уравнения движения остаются теми же, поэтому в качестве функции Лагранжа при описании движения нерелятивистской частицы в потенциальном поле можно использовать метрическую функцию L (40) со специально выбранными аргументами, а именно, $L(1, v_x, v_y, v_z; x, y, z)$.

В четырёхмерном пространстве событий в x^0, x^1, x^2, x^3 с метрическим тензором $g_{ij}(x)$, который, как считается в ОТО, описывает гравитационное поле, и электромагнитным полем, описываемым 4-потенциалом $A_i(x)$, а также с частицей (материальной точкой) массой m и зарядом e , определена метрическая функция вида

$$L(dx; x) = cm\sqrt{g_{ij}(x)dx^i dx^j} + \frac{e}{c}A_i dx^i + Mc dx^0, \quad (45)$$

где $M > 0$ достаточно большая постоянная. Такая метрическая функция является положительно однородной функцией первой степени по первым четырём аргументам. Конус измеримых бесконечно малых смещений при достаточно большом значении постоянной $M > 0$ определяется неравенствами:

$$g_{ij}dx^i dx^j > 0, \quad dx^0 > 0. \quad (46)$$

Тем самым такое четырёхмерное пространство событий является неквадратичным финслеровым пространством.

Метрическая функция Бервальда - Моора

$$L(dx^1, dx^2, dx^3, dx^4) = \sqrt[4]{dx^1 dx^2 dx^3 dx^4} \quad (47)$$

определяет в четырёхмерном пространстве x^1, x^2, x^3, x^4 , которое будем обозначать H_4 , плоскую финслерову геометрию, не являющуюся квадратичной. Измеримые бесконечно малые смещения определяются неравенством

$$dx^1 dx^2 dx^3 dx^4 > 0 \quad (48)$$

и заполняют 8 несвязанных конусов. Индикатриса определяется уравнением

$$\xi^1 \xi^2 \xi^3 \xi^4 = 1 \quad (49)$$

и состоит из 8-ми несвязанных полостей гиперboloида 4-го порядка. Если конус измеримых направлений определить неравенствами

$$dx^1 > 0, \quad dx^2 > 0, \quad dx^3 > 0, \quad dx^4 > 0, \quad (50)$$

то мы получим уже другую плоскую финслерову геометрию с индикатрисой, состоящей из одной полости гиперboloида 4-го порядка, расположенной в конусе будущего.

Все финслеровы пространства делятся на два класса: 1) основное координатное пространство не содержит ни в какой системе координат ни одной координаты, которую можно было бы назвать физическим временем или аналогом физического времени, обычно такие пространства имеют невогнутую индикатрису; 2) найдется такая система координат и в ней выделенная координата, которая обладает качествами физического времени, обычно такие пространства имеют вогнутую индикатрису или, хотя бы часть индикатрисы является вогнутой гиперповерхностью.

Наглядным представителем первого класса является евклидова геометрия, а второго - псевдоевклидова геометрия.

3 Классическая механика

Каждому понятию классической механики соответствует некоторое понятие финслеровой геометрии. С одной стороны эти понятия схожи, а с другой между такими понятиями существуют более или менее серьёзные различия. Обычно таким парным понятиям в классической механике и финслеровой геометрии присваивают разные термины, например, *действие между двумя событиями*, принадлежащих одной и той же мировой линии, в классической механике и *длина отрезка кривой* в финслеровой геометрии. Можно говорить, что каждое финслерово пространство наделено классической механикой, так как в каждом финслеровом пространстве определены и работают все понятия и методы классической механики при соответствующей их трансформации к понятиям и методам финслеровой геометрии.

Пусть в основном координатном финслеровом пространстве задана кривая

$$x^i = x^i(\tau), \quad \tau \in [\tau_1, \tau_2], \quad (51)$$

где τ - монотонно возрастающий параметр вдоль кривой, то есть $d\tau > 0$, тогда длина отрезка кривой между точками $x^i(\tau_1)$ и $x^i(\tau_2)$ вычисляется следующим образом:

$$l_{21} = \int_{M_1}^{M_2} L(dx^1, \dots, dx^n; x^1, \dots, x^n) = \int_{\tau_1}^{\tau_2} L(\dot{x}^1, \dots, \dot{x}^n; x^1, \dots, x^n) d\tau, \quad (52)$$

если все смещения dx вдоль кривой измеримы, и где

$$\dot{x}^i = \frac{dx^i}{d\tau} \quad (53)$$

- компоненты вектора, касательного к кривой, или компоненты вектора скорости по параметру τ .

Если основное координатное пространство описывает физическую систему, то кривая (51) суть мировая линия, а величина l_{21} (длина отрезка мировой линии) может рассматриваться с точностью до постоянного множителя и аддитивной добавки под интегралом вида $df(x^1, x^2, \dots, x^n)$, где f - произвольная функция координат, как действие. Кривые, имеющие экстремальную длину при фиксированных начальной и конечной точках, удовлетворяют системе уравнений Лагранжа - Эйлера:

$$\delta l_{12} = 0 \quad \Rightarrow \quad \frac{d}{d\tau} \frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = 0 \quad (54)$$

- и называются *экстремальными*. Экстремали могут быть и мировыми линиями, и траекториями движения в n -мерном пространстве событий, и геодезическими. Геодезические, конечно, являются экстремальными, но обратное верно не всегда. В несколько расширенном смысле будем иногда называть экстремали "геодезическими". Параметр τ вдоль кривой естественно назвать параметром эволюции, если пространство содержит временные переменные.

Итак, мировая линия физической системы есть экстремаль основного финслерова пространства. Принципу наименьшего действия, или стационарного действия классической механики соответствует в финслеровой геометрии принцип экстремальности отрезка кривой.

Введем обозначения

$$p_i = \frac{\partial L}{\partial \dot{x}^i}, \quad F_i = \frac{\partial L}{\partial x^i}. \quad (55)$$

Заметим, что величины p_i мы определяли (16) как частные производные функции $L(dx; x)$ по первым n аргументам dx^i ,

$$p_i = \frac{\partial L(dx; x)}{\partial (dx^i)}, \quad (56)$$

а в (55) определяем как частные производные функции $L(\dot{x}; x)$ по первым n аргументам $\dot{x}^i$. Это не приводит ни к ошибкам, ни к путанице в силу того, что метрическая функция является положительно однородной первой степени по первым n аргументам и параметр эволюции, как мы условились выше, монотонно возрастает, то есть $d\tau > 0$. В новых обозначениях уравнения Лагранжа - Эйлера принимают ньютоновский вид:

$$\frac{d}{d\tau} p_i = F_i \quad (57)$$

- то есть величины p_i естественно называть компонентами обобщённого импульса, что мы и делали выше, а F_i - компонентами обобщённой силы. Из формул (54), (55) следует: если метрическая функция и параметр эволюции являются скалярами, то обобщённый импульс p и обобщённая сила F - ковариантные тензоры, а уравнения движения Лагранжа - Эйлера (54) и уравнения движения Ньютона (57) имеют явную ковариантную запись.

Так как метрическая функция является однородной первой степени по первым n аргументам, то, во-первых,

$$\frac{\partial L(dx; x)}{\partial (dx^i)} dx^i = L(dx; x) \quad \Rightarrow \quad L(dx; x) = p_i \cdot dx^i, \quad (58)$$

или

$$\frac{\partial L(\dot{x}; x)}{\partial \dot{x}^i} \dot{x}^i = L(\dot{x}; x) \quad \Rightarrow \quad L(\dot{x}^1, \dots, \dot{x}^n; x^1, \dots, x^n) = p_i \cdot \dot{x}^i, \quad (59)$$

а во-вторых, компоненты обобщённого импульса - однородные функции нулевой степени по первым n аргументам, то есть n функций зависит от $(n - 1)$ или меньше независимых комплексов, а значит, как уже отмечалось выше, между ними существует хотя бы одна функциональная зависимость. Из требования (34) следует, что такая зависимость единственная, её принято записывать следующим образом:

$$\Phi(p_1, p_2, \dots, p_n; x^1, x^2, \dots, x^n) = 0. \quad (60)$$

В финслеровой геометрии такая зависимость называется тангенциальным уравнением индикатрисы, а функцию $\Phi(p; x)$ принято называть функцией Финслера. Зависимость (60) имеет тот же смысл, что связь между обобщённой энергией (функцией Гамильтона) и компонентами обобщённого импульса в классической механике. Если известна функция $\Phi(p; x)$, но не известна соответствующая ей метрическая функция, то метрическую функцию $L(\xi^1, \dots, \xi^n; x^1, \dots, x^n)$ можно попытаться найти как решение дифференциального уравнения в частных производных

$$\Phi\left(\frac{\partial L}{\partial \xi^1}, \dots, \frac{\partial L}{\partial \xi^n}; x^1, x^2, \dots, x^n\right) = 0. \quad (61)$$

Оказывается [1], что метрическую функцию можно всегда однозначно восстановить по известной функции Финслера, разрешая систему неявных функциональных уравнений, не прибегая к дифференциальному уравнению в частных производных (61).

Рассмотрим длину l_{12} некоторой экстремали как функцию точки верхнего предела интегрирования при фиксированной нижней точке. Обозначим эту величину $S(x^1, \dots, x^n)$, в классической механике её принято называть действием как функцией координат. Обобщённые импульсы, отнесённые к текущей точке на экстремали, также будем считать функциями точки основного пространства $p_i(x^1, \dots, x^n)$. Из формул (52) и (58) следует, что

$$S(x^1, \dots, x^n) = \int_{M_1}^{M(x^1, \dots, x^n)} p_i dx^i. \quad (62)$$

Траектории движения материальных частиц и траектории эволюции физической системы с конечным числом степеней свободы в классической механике обладают замечательным свойством [10]:

$$p_i(x) = \frac{\partial S(x)}{\partial x^i}. \quad (63)$$

В финслеровой геометрии это же свойство формулируется как принадлежность экстремалей материальных частиц и физических систем с конечным числом степеней свободы некоторой *нормальной конгруэнции экстремалей* [1]. Подставим выражения (63) в тангенциальное уравнение индикатрисы (60), получим нелинейное дифференциальное уравнение в частных производных первого порядка относительно одной неизвестной функции $S(x)$,

$$\Phi\left(\frac{\partial S}{\partial x^1}, \frac{\partial S}{\partial x^2}, \dots, \frac{\partial S}{\partial x^n}; x^1, x^2, \dots, x^n\right) = 0, \quad (64)$$

решением которого является функция $S(x^1, \dots, x^n)$ (действие, как функция координат) и которое в классической механике называется уравнением Гамильтона - Якоби. В финслеровой геометрии всякая нормальная конгруэнция экстремалей определяется функцией $S(x)$,

которая должна удовлетворять нелинейному дифференциальному уравнению в частных производных (64) [1].

Рассмотрим четырёхмерное финслерово пространство, которое соответствует одной нерелятивистской частице, движущейся в потенциальном поле $U(x, y, z)$, независимом от времени, с метрической функцией (40). Для удобства будем использовать в данной случае двойные обозначения $t, x^1 \equiv x, x^2 \equiv y, x^3 \equiv z$, которые позволяют обозначать через p_t компоненту импульса, соответствующую координате t (времени), а через $p_1 \equiv p_x, p_2 \equiv p_y, p_3 \equiv p_z$, компоненты импульса, соответствующие координатам x^1, x^2, x^3 . Тогда, используя формулу (56), имеем:

$$p_t = mc^2 + m \cdot \frac{(dx)^2 + (dy)^2 + (dz)^2}{2(dt)^2} + U(x, y, z), \quad (65)$$

$$p_\alpha = -m \cdot \frac{dx^\alpha}{dt}. \quad (66)$$

Компонента импульса p_t - это полная энергия частицы в потенциальном поле вместе с её внутренней энергией mc^2 . Компоненты импульса p_α отличаются от принятых в классической механике лишь знаком.

Выберем в качестве параметра эволюции время. Тогда в качестве функции Лагранжа можно использовать метрическую функцию $L(dt, dx, dy, dz; x, y, z)$, но с другими первыми четырьмя аргументами

$$L(1, \dot{x}, \dot{y}, \dot{z}; x, y, z) = mc^2 - m \cdot \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{2} + U(x, y, z). \quad (67)$$

В этом случае уравнения движения Лагранжа классической механике запишутся следующим образом:

$$\frac{d}{dt} p_\alpha = \frac{\partial U}{\partial x^\alpha}, \quad (68)$$

или

$$m\ddot{x}^\alpha = -\frac{\partial U}{\partial x^\alpha}. \quad (69)$$

Это обычные уравнения движения Ньютона для частицы массой m в потенциальном поле $U(x, y, z)$.

Исключая из четырёх соотношений (65) и (66) компоненты скорости $\dot{x}^\alpha$, приходим к тангенциальному уравнению индикатрисы:

$$p_t - \left\{ mc^2 + \frac{p_1^2 + p_2^2 + p_3^2}{2m} + U \right\} = 0. \quad (70)$$

Функцию, стоящую в фигурных скобках, принято называть в классической механике функцией Гамильтона и обозначать

$$H(p_1, p_2, p_3; x^1, x^2, x^3) = mc^2 + \frac{p_1^2 + p_2^2 + p_3^2}{2m} + U. \quad (71)$$

Тогда тангенциальное уравнение индикатрисы запишется следующим образом:

$$p_t - H(p_1, p_2, p_3; x^1, x^2, x^3) = 0. \quad (72)$$

Для данной физической системы уравнение Гамильтона - Якоби имеет вид

$$\frac{\partial S}{\partial t} - \left\{ mc^2 + \frac{1}{2m} \left[\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 + \left(\frac{\partial S}{\partial z} \right)^2 \right] + U \right\} = 0. \quad (73)$$

Так как тангенциальное уравнение (70) не зависит от времени явно, то не зависит и гамильтониан (71). Система консервативна, полная энергия сохраняется во времени, то есть $p_t = E = \text{const}$; поэтому зависимость действия от времени сводится к слагаемому $(E \cdot t)$:

$$S(t, x, y, z) = S_0(x, y, z) + E \cdot t. \quad (74)$$

Подставив это выражение для действия в (73), получим для укороченного действия $S_0(x, y, z)$ уравнение Гамильтона - Якоби в виде

$$mc^2 + \frac{1}{2m} \left[\left(\frac{\partial S_0}{\partial x} \right)^2 + \left(\frac{\partial S_0}{\partial y} \right)^2 + \left(\frac{\partial S_0}{\partial z} \right)^2 \right] + U(x, y, z) = E. \quad (75)$$

Заметим, что формулы (73) и (74) отличаются от общепринятых [10] знаком между двумя аддитивными членами, это связано с тем, что в данном геометрическом подходе

$$\frac{\partial S}{\partial t} = E > 0, \quad (76)$$

где $E > 0$ - полная энергия системы.

4 Канонические уравнения

Для определения финслеровой геометрии в основном координатном пространстве $x^1, x^2, \dots, x^n$ надо задать либо метрическую функцию $L(dx; x)$, либо тангенциальное уравнение индикатрисы:

$$\Phi(p_1, p_2, \dots, p_n; x^1, x^2, \dots, x^n) = 0. \quad (77)$$

Кривая (зададим её параметрически

$$x^i = x^i(\tau), \quad \tau \in [\tau_1, \tau_2], \quad (78)$$

где τ - монотонно возрастающий параметр вдоль кривой) только тогда будет экстремалью, когда функциональная зависимость, зафиксированная в тангенциальном уравнении индикатрисы, не будет нарушаться, то есть

$$\frac{d}{d\tau} \Phi(p_1, p_2, \dots, p_n; x^1, x^2, \dots, x^n) = 0, \quad (79)$$

или

$$\frac{\partial \Phi}{\partial p_i} \dot{p}_i + \frac{\partial \Phi}{\partial x^i} \dot{x}^i = 0. \quad (80)$$

Если

$$\dot{x}^i = \frac{\partial \Phi}{\partial p_i} \cdot \lambda(p; x), \quad \dot{p}_i = -\frac{\partial \Phi}{\partial x^i} \cdot \lambda(p; x), \quad (81)$$

где $\lambda(p; x) \neq 0, \pm\infty$ - произвольная функция p и x , то вдоль таких кривых выполняется условие (79), и всегда выполняется соотношение (77), если этому соотношению удовлетворяют начальные значения $x^i(\tau_1), p_i(\tau_1)$.

Уравнения (81) можно получить математически строго, если задачу на экстремум функционала (52) сформулировать несколько иначе. Так как метрическая функция является однородной первой степени по первым n аргументам, то

$$L(\dot{x}; x) = p_i \dot{x}^i, \quad (82)$$

поэтому функционал (52) принимает вид

$$l_{12} = \int_{\tau_1}^{\tau_2} p_i \dot{x}^i d\tau \quad (83)$$

при условии выполнения (77). То есть имеет место задача на условный экстремум функционала, для решения которой следует составить новый функционал [5], [8]

$$\check{l}_{12} = \int_{\tau_1}^{\tau_2} [p_i \dot{x}^i - \check{\lambda}(p; x) \Phi(p; x)] d\tau \quad (84)$$

и именно для него решать задачу на экстремум для $2n$ неизвестных функций $x^i(\tau)$ и $p_i(\tau)$ при нулевых вариациях кривой (78) на концах отрезка, то есть

$$\delta x^i(\tau) \Big|_{\tau=\tau_1} = 0, \quad \delta x^i(\tau) \Big|_{\tau=\tau_2} = 0, \quad (85)$$

причём функцию $\check{\lambda}(p; x)$ следует подобрать так, чтобы выполнялась функциональная зависимость (77). Из требования экстремальности функционала (84)

$$\delta \check{l}_{12} = 0$$

с учётом того, что

$$\delta \check{l}_{12} = p_1 \delta x^1 \Big|_{\tau_1}^{\tau_2} + \int_{\tau_1}^{\tau_2} \left[-\dot{p}_i - \frac{\partial(\check{\lambda}\Phi)}{\partial x^i} \right] \delta x^i d\tau + \int_{\tau_1}^{\tau_2} \left[\dot{x}^i - \frac{\partial(\check{\lambda}\Phi)}{\partial p_i} \right] \delta p_i d\tau,$$

и учётом условий (85), получаем уравнения Эйлера - Лагранжа:

$$\dot{x}^i = \frac{\partial(\check{\lambda}\Phi)}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial(\check{\lambda}\Phi)}{\partial x^i}. \quad (86)$$

Выберем $\check{\lambda} \equiv \lambda$ и учтём условие (77), тогда полученные уравнения совпадают с уравнениями (81). Напомним: для выполнения функциональной зависимости (77) необходимо и достаточно, чтобы эта функциональная зависимость выполнялась только в начальной точке экстремали.

Итак, доказано, что уравнения (81) являются уравнениями для определения экстремалей в финслеровом пространстве.

Произвол в выборе функции $\lambda(p; x)$ связан лишь с возможностью различной параметризации экстремалей.

Уравнения (81) в классической механике называются *каноническими уравнениями* движения [10]. Можно утверждать, что функция Финслера $\Phi(p; x)$ играет ту же (аналогичную) роль в финслеровой геометрии, что и функция Гамильтона в классической механике. Если заменить $\Phi \rightarrow H$ и положить $\lambda(p; x) \equiv 1$, то уравнения (81) совпадают по виду с уравнениями Гамильтона классической механики, но в финслеровом подходе к классической механике время становится одной из координат, что приводит к возможности ковариантной записи уравнений движения.

В классической механике, если известна функция Лагранжа, функция Гамильтона получается однозначно. В финслеровой геометрии при заданной метрической функции тангенциальное уравнение индикатрисы можно записать во многих эквивалентных формах, что приводит к разным функциям Финслера. При этом экстремали останутся теми же, изменится лишь их параметризация, как если бы произошла замена параметра эволюции $\tau \rightarrow \tau'(\tau)$ своя для каждой мировой линии. Этот произвол аналогичен выбору функции $\lambda(p; x) \neq 0$ в канонических уравнениях (81), что также можно рассматривать как перепараметризацию экстремалей.

Если известно действие $S(x)$ как функция координат, то с учётом (63) нахождение экстремалей сводится к решению всего n дифференциальных уравнений первого порядка:

$$\dot{x}^i = \left. \frac{\partial \Phi}{\partial p_i} \right|_{p_j = \frac{\partial S}{\partial x^j}} \cdot \lambda(x), \quad (87)$$

где $\lambda(x) \neq 0, \pm\infty$ - некоторая функция. Если же исключить из этих уравнений параметр эволюции, то останется система из $(n-1)$ -го дифференциального уравнения первого порядка. Для этого следует разделить все уравнения (87) с индексами $i = 1, 2, \dots, n-1$ на первое с индексом $i = 0$, если x^0 - временная или аналогичная временной координата, и учесть, что

$$\frac{\dot{x}^\alpha}{\dot{x}^0} = \frac{dx^\alpha}{dx^0}, \quad \alpha = 1, 2, \dots, n-1. \quad (88)$$

Функция $\lambda(x)$ исчезает из системы уравнений так же, как и параметр эволюции:

$$\frac{dx^\alpha}{dx^0} = \left. \frac{\frac{\partial \Phi}{\partial p_\alpha}}{\frac{\partial \Phi}{\partial p_0}} \right|_{p_j = \frac{\partial S}{\partial x^j}}. \quad (89)$$

Пусть $f(p; x)$, $g(p; x)$, $h(p; x)$ - некоторые функции обобщённых импульсов p и обобщённых координат x . Найдём, как изменяются такие величины вдоль экстремалей, то есть найдем их производную по параметру эволюции:

$$\frac{df}{d\tau} = \frac{\partial f}{\partial p_i} \dot{p}_i + \frac{\partial f}{\partial x^i} \dot{x}^i. \quad (90)$$

Подставляя $\dot{p}_i$ и $\dot{x}^i$ из канонических уравнений (81), получим

$$\frac{df}{d\tau} = \left(-\frac{\partial f}{\partial p_i} \frac{\partial \Phi}{\partial x^i} + \frac{\partial f}{\partial x^i} \frac{\partial \Phi}{\partial p_i} \right) \lambda(p; x). \quad (91)$$

Скобками Пуассона $\{g, f\}$ произвольных функций $f(p; x)$, $g(p; x)$ называется выражение

$$\{g, f\} = \frac{\partial g}{\partial p_i} \frac{\partial f}{\partial x^i} - \frac{\partial g}{\partial x^i} \frac{\partial f}{\partial p_i}. \quad (92)$$

Тогда с помощью скобок Пуассона производную $\frac{df}{d\tau}$ величины $f(p; x)$ по параметру эволюции τ можно записать следующим образом:

$$\frac{df}{d\tau} = \{\Phi, f\} \cdot \lambda(p; x). \quad (93)$$

Функции координат и компонент обобщённого импульса, которые остаются постоянными при движении физической системы, называются *интегралами движения*.

Для того чтобы величина f была интегралом движения, необходимо и достаточно выполнение условия

$$\{\Phi, f\} = 0. \quad (94)$$

Это следует из формулы (93), так как $\lambda(p; x) \neq 0$.

Очевидно, что

$$\{f, f\} = 0,$$

в частности

$$\{\Phi, \Phi\} = 0,$$

то есть функция Финслера всегда является интегралом движения.

Перечислим ряд простейших свойств скобок Пуассона, которые следуют непосредственно из определения.

Если переставить функции в скобках Пуассона, то скобки меняют знак:

$$\{g, f\} = -\{f, g\}. \quad (95)$$

Если одна из величин в скобках Пуассона - число, например, a , то скобка равна нулю:

$$\{a, f\} = 0. \quad (96)$$

Скобки Пуассона обладают линейным свойством по каждому из аргументов; например, для первого аргумента имеем

$$\{ag + bh, f\} = a\{g, f\} + b\{h, f\}, \quad (97)$$

где a и b - числа.

Если один из аргументов скобок Пуассона является произведением двух функций, то имеет место следующая формула

$$\{hg, f\} = h\{g, f\} + \{h, f\}g. \quad (98)$$

Если одна из функций в скобках Пуассона совпадает с координатой или компонентой обобщённого импульса, то скобки сводятся к частной производной соответственно по импульсу или координате:

$$\{f, x^i\} = \frac{\partial f}{\partial p_i}, \quad \{p_i, f\} = \frac{\partial f}{\partial x^i}. \quad (99)$$

Важное значение имеют скобки Пуассона, когда оба аргумента являются координатой или компонентой обобщённого импульса:

$$\{x^i, x^j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{p_i, x^j\} = \delta_i^j. \quad (100)$$

Для скобок Пуассона справедливо тождество Якоби:

$$\{h, \{g, f\}\} + \{g, \{f, h\}\} + \{f, \{h, g\}\} = 0. \quad (101)$$

Пусть функции $f(p; x; u, \dots, v)$ и $g(p; x; u, \dots, v)$ зависят ещё и от некоторого числа независимых параметров $u, \dots, v$, ни один из которых, конечно, не является параметром эволюции, обозначим через z любой аргумент этих функций: p_i, x^j или $u, \dots, v$. Тогда справедлива следующая формула:

$$\frac{\partial}{\partial z} \{g, f\} = \left\{ \frac{\partial g}{\partial z}, f \right\} + \left\{ g, \frac{\partial f}{\partial z} \right\}. \quad (102)$$

Вычислим производную от $\{g, f\}$ по параметру эволюции. В этом случае формула (102) не применима, но согласно (93) имеем

$$\frac{d}{d\tau} \{g, f\} = \{\Phi, \{g, f\}\} \lambda. \quad (103)$$

Воспользуемся тождеством Якоби, тогда

$$\frac{d}{d\tau} \{g, f\} = \{\{\Phi, g\}, f\} \lambda + \{g, \{\Phi, f\}\} \lambda, \quad (104)$$

или

$$\frac{d}{d\tau} \{g, f\} = \left\{ \frac{1}{\lambda} \frac{dg}{d\tau}, f \right\} \lambda + \left\{ g, \frac{1}{\lambda} \frac{df}{d\tau} \right\} \lambda. \quad (105)$$

Из формулы (105) следует теорема Пуассона.

Теорема Пуассона: если $f(p; x)$ и $g(p; x)$ - два интеграла движения, то и их скобки Пуассона также являются интегралом движения, то есть вдоль мировой линии

$$\{g, f\} = \text{const}. \quad (106)$$

Скобки Пуассона от двух интегралов движения могут дать новый (функционально независимый от первых) интеграл движения, но далеко не всегда. Число функционально независимых интегралов движения не может быть больше $(2n - 1)$. Если бы число функционально независимых интегралов движения оказалось равным $2n$, то, приравняв их к постоянным и разрешив относительно $p_1, p_2, \dots, p_n, x^1, x^2, \dots, x^n$, мы получили бы тривиальный результат: $p_i = \text{const}, x^j = \text{const}$; то есть любая функция $f(p; x)$ являлась бы интегралом движения и время в том числе.

В основном координатном пространстве можно переходить от одних координат к другим (1), и при этом мы остаёмся в том же самом координатном пространстве с той же самой финслеровой геометрией. Физическая система, которая описывается этим координатным пространством с метрической функцией в нём, также не изменяется. При этом уравнениями экстремалей являются уравнения Эйлера - Лагранжа (54) или каноническими уравнениями (81). В последние формально координаты и импульсы входят равноправно, хотя и разделяются на два подмножества, поэтому их принято называть *канонически сопряжёнными величинами*. При решении системы канонических уравнений (81) никто не запрещает вместо неизвестных функций $x^i(\tau)$ и $p_j(\tau)$ ввести другие неизвестные функции

$$X^i = X^i(x^1, \dots, x^n; p_1, \dots, p_n), \quad P_i = P_i(x^1, \dots, x^n; p_1, \dots, p_n), \quad (107)$$

но при этом совсем не обязательно, что $X^i(\tau)$ и $P_j(\tau)$ будут обобщёнными координатами и обобщёнными импульсами хоть какого-то финслерова пространства, пусть даже другого. В то же время наша физическая система не изменилась. Если всё же $X^i(\tau)$ и $P_j(\tau)$ будут обобщёнными координатами и обобщёнными импульсами какого-то финслерова пространства, то экстремали будут решениями системы канонических уравнений с заменой

$$x^i, p_i, \Phi(p; x), \lambda(p; x) \quad \rightarrow \quad X^i, P_i, \bar{\Phi}(P; X), \bar{\lambda}(X; P). \quad (108)$$

Итак, если мы не хотим выходить за рамки множества финслеровых пространств необходимо (но не достаточно) ограничиться такими преобразованиями (107), которые переводят систему канонических уравнений в систему канонических уравнений. Естественно такие преобразования назвать каноническими.

Среди множества канонических преобразований существуют такие, которые переводят финслерово пространство с одной метрической функцией в другое финслерово пространство с другой метрической функцией, при этом физическая система остаётся той же. Таким образом, одной и той же физической системе соответствует, вообще говоря, много качественно различных финслеровых геометрий.

5 Мировая функция

Представим себе, что пространство событий x^0, x^1, x^2, x^3 (или его некоторая область) заполнено материальными точками, которые как-то взаимодействуют между собой, но не распадаются и не сливаются, не объединяются. Рассмотрим эволюцию такого пространства событий в бесконечно близкое состояние. В этом случае каждая точка (x^0, x^1, x^2, x^3) перейдёт в бесконечно близкую точку $(x^0 + dx^0, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)$, и всё пространство-время (или некоторая его область) будет зачерчено мировыми линиями, причём через каждую точку проходит одна и только одна мировая линия.

Если через каждую точку пространств (или некоторую его область) проходит одна и только одна мировая линия, говорят, что в пространстве (в этой области пространства) задана *конгруэнция мировых линий*.

Если известна конгруэнция мировых линий, то известно, по каким траекториям движутся все материальные объекты в пространстве-времени, а значит известно и поле скоростей в каждой точке. Следует сразу отметить два факта. Во-первых, конгруэнция мировых линий, соответствующая физическому Миру, не может быть произвольной, а должна удовлетворять

некоторым требованиям. Во-вторых, для описания физического Мира не достаточно знания траекторий движения всех материальных точек, необходимо ещё знание хотя энергетических характеристик материальных объектов, то есть, например, поля обобщённых импульсов.

Пара векторных полей {поле скоростей; поле импульсов} может быть заменена на пару скалярных функций {функция Финслера; действие как функция координат} $\equiv \{\Phi(p; x); S(x)\}$, функция $S(x)$ при этом удовлетворяет дифференциальному уравнению (1.3.14) и в финслеровой геометрии имеет смысл длины отрезка мировой линии от некоторой фиксированной точки этой мировой линии до точки $M(x)$, принадлежащей той же мировой линии. Напомним, как поля $\{\dot{x}(x); p(x)\}$ выражаются через поля $\{\Phi(p; x); S(x)\}$:

$$\dot{x}^i = \frac{\partial \Phi}{\partial p_i} \Big|_{p_j = \frac{\partial S}{\partial x^j}} \cdot \lambda(x), \quad p_j = \frac{\partial S}{\partial x^j}. \quad (109)$$

Конгруэнция мировых линий является, конечно, конгруэнцией экстремалей.

Конгруэнции экстремалей, которые могут быть получены согласно формулам (109), называются *нормальными* [1].

Всегда будем предполагать, что конгруэнция мировых линий любой физической системы является нормальной в рассматриваемой области изменения координат.

Итак, пара полей $\{\dot{x}(x); p(x)\}$ определяет нам физический Мир в некоторой координатной области, если в этой области поле скоростей задаёт *нормальную* конгруэнцию мировых линий.

Если физический Мир известен, то функция $S(x)$ определяется с точностью до постоянной. Существует ли какой-нибудь произвол при этом в выборе финслеровой геометрии, или она определяется физическим Миром однозначно?

Покажем, что одному и тому же физическому Миру могут соответствовать больше, чем одна финслерова геометрия.

Рассмотрим две геометрии, первая из которых четырёхмерная риманова (8) или псевдориманова геометрия (10) ОТО. В таких геометриях тангенциальное уравнение индикатрисы имеет вид

$$g^{ij}(x)p_i p_j - 1 = 0. \quad (110)$$

В качестве второй геометрии рассмотрим геометрию, конформно связанную с геометрией (47), запишем её метрическую функцию ковариантно в изотропном базисе:

$$L(dx; x) = \kappa(x) \sqrt[4]{g_{ijkl} dx^i dx^j dx^k dx^l}, \quad (111)$$

где $\kappa(x) > 0$ - некоторая скалярная функция, а числовой тензор

$$g_{ijkl} = \begin{cases} \frac{1}{24}, & \text{если все индексы разные,} \\ 0, & \text{во всех остальных случаях.} \end{cases} \quad (112)$$

Тогда тангенциальное уравнение индикатрисы можно записать следующим образом:

$$g^{ijkl} p_i p_j p_k p_l - \left(\frac{\kappa}{4}\right)^4 = 0, \quad (113)$$

где числовой тензор

$$g^{ijkl} = \begin{cases} \frac{1}{24}, & \text{если все индексы разные,} \\ 0, & \text{во всех остальных случаях,} \end{cases} \quad (114)$$

то есть в изотропном базисе

$$(g^{ijkl}) = (g_{ijkl}).$$

Пусть $S(x)$ - некоторая скалярная функция, удовлетворяющая условию

$$g^{ijkl} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} \frac{\partial S}{\partial x^k} \frac{\partial S}{\partial x^l} \equiv \frac{\partial S}{\partial x^1} \frac{\partial S}{\partial x^2} \frac{\partial S}{\partial x^3} \frac{\partial S}{\partial x^4} > 0. \quad (115)$$

Зададим коэффициент растяжения-сжатия формулой:

$$\kappa(x) \equiv 4 \sqrt[4]{g^{ijkl} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} \frac{\partial S}{\partial x^k} \frac{\partial S}{\partial x^l}}. \quad (116)$$

В этом случае функция $S(x)$ автоматически удовлетворяет дифференциальному уравнению (1.3.14), записанному для финслеровой геометрии (111), и определяет нормальную конгруэнцию мировых линий (109):

$$\dot{x}^i = g^{ijkl} \frac{\partial S}{\partial x^j} \frac{\partial S}{\partial x^k} \frac{\partial S}{\partial x^l} \lambda(x), \quad p_j = \frac{\partial S}{\partial x^j}, \quad (117)$$

где $\lambda(x) \neq 0$ - некоторая функция.

Определим дважды контравариантный тензор $g^{ij}(x)$ следующей формулой:

$$g^{ij}(x) = g^{ijkl} \frac{\partial S}{\partial x^k} \frac{\partial S}{\partial x^l}. \quad (118)$$

Так как его определитель

$$\det(g^{ij}(x)) = -\frac{3}{12^4} \left(\frac{\partial S}{\partial x^1} \frac{\partial S}{\partial x^2} \frac{\partial S}{\partial x^3} \frac{\partial S}{\partial x^4} \right)^2 \neq 0 \quad (119)$$

не равен нулю, в силу условия (115), то можно однозначно построить дважды ковариантный тензор $g_{ij}(x)$ и рассматривать его как метрический тензор риманова или псевдориманова пространства. Рассмотрим геометрию, конформно связанную с такой геометрией, то есть геометрию с элементом длины

$$ds = \left(\frac{\kappa(x)}{4} \right)^2 \sqrt{g_{ij}(x) dx^i dx^j}, \quad (120)$$

где $\kappa(x)$ определяется формулой (116). Это риманова или псевдориманова геометрия с метрическим тензором

$$G_{ij} = \left(\frac{\kappa(x)}{4} \right)^4 g_{ij}(x), \quad (121)$$

и элементом длины

$$ds' = \sqrt{G_{ij}(x)dx^i dx^j}.$$

В этой геометрии тангенциальное уравнение индикатрисы имеет вид

$$g^{ij}(x)p_i p_j - \left(\frac{\kappa(x)}{4}\right)^4 = 0. \quad (122)$$

Запишем в этой геометрии уравнение (64) для действия $\check{S}(x)$ как функции координат

$$g^{ij}(x)\frac{\partial \check{S}}{\partial x^i}\frac{\partial \check{S}}{\partial x^j} - \left(\frac{\kappa}{4}\right)^4 = 0, \quad (123)$$

или, учитывая (118), получим

$$g^{ijkl}(x)\frac{\partial \check{S}}{\partial x^i}\frac{\partial \check{S}}{\partial x^j}\frac{\partial S}{\partial x^k}\frac{\partial S}{\partial x^l} - \left(\frac{\kappa}{4}\right)^4 = 0. \quad (124)$$

Из (116) следует, что одним из решений этого уравнения является функция $\check{S}(x) \equiv S(x)$. Запишем уравнения для мировых линий в геометрии (120):

$$\dot{x}^i = g^{ij}(x)\frac{\partial S}{\partial x^j} \cdot \check{\lambda}(x). \quad (125)$$

Учитывая (118), получим (117), если заменить $\check{\lambda}$ на λ .

Таким образом, в одной и той же координатной системе имеют место две качественно разные геометрии, которые описывают один и тот же физический Мир $\{\dot{x}(x); p(x)\}$.

Выше упоминалось, что конкретной физической системе соответствуют бесконечное множество качественно различных финслеровых геометрий, которые связаны между собой каноническими преобразованиями, то есть преобразованиями специального вида в фазовом пространстве p, x .

Теперь мы доказали более сильное утверждение: в одном и том же координатном пространстве существует класс финслеровых геометрий, которые имеют общую конгруэнцию мировых линий вместе с общим полем обобщённых импульсов.

Итак, одному и тому же физическому Миру соответствует не одна финслерова геометрия, а некоторый класс финслеровых геометрий.

Очевидно, что функция $S(x)$ наряду с требованиями типа (115) должна удовлетворять некому фундаментальному уравнению поля одному и тому же для всего класса финслеровых геометрий, которые соответствуют данному физическому Миру, или данной замкнутой физической системе. Такую функцию $S(x)$ будем называть *Мировой функцией* и обозначать $S_W(x)$.

Отметим также, что из выше изложенного просматривается связь между Мировой функцией $S_W(x)$ и коэффициентом растяжения сжатия $\kappa(x)$, который возникает при переходе от некоего исходного финслерова пространства к конформно связанному с ним финслеру пространству.

6 Теория поля

Наиболее удобный способ получения уравнений поля как в классической теории [11], так и в теории квантованных полей [12] связан с понятиями лагранжиана, действия, а также с принципом стационарного (или наименьшего) действия, то есть с принципом экстремума действия. При этом однозначно устанавливается связь [7] между непрерывными преобразованиями, относительно которых инвариантно действие, и физическими законами сохранения, которые могут быть проверены экспериментально.

Если x^0, x^1, x^2, x^3 - координаты, $f(x) \equiv f(x^0, x^1, x^2, x^3)$ - скалярное поле в пространстве Минковского, а

$$\mathfrak{L} \equiv \mathfrak{L} \left(f(x); \frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \frac{\partial f}{\partial x^3} \right), \quad (126)$$

- лагранжиан, то интеграл от лагранжиана по некоторому четырехмерному объему V пространства-времени,

$$I[f] = \int_V \mathfrak{L} dx^0 dx^1 dx^2 dx^3 \quad (127)$$

принято называть действием. Считая, что вариации функции поля δf равны нулю на границе объема интегрирования, и требуя экстремальности действия, то есть

$$\delta I[f] = 0, \quad (128)$$

с помощью известной математической процедуры получим уравнение Лагранжа - Эйлера, которое и есть уравнение поля:

$$\frac{\partial}{\partial x^i} \frac{\partial \mathfrak{L}}{\partial \left(\frac{\partial f}{\partial x^i} \right)} - \frac{\partial \mathfrak{L}}{\partial f} = 0. \quad (129)$$

Обычно лагранжиан подбирают так, чтобы получить известные уравнения поля, или строят, обеспечивая заданную симметрию и выполнение некоторых дополнительных требований, например, чтобы получались линейные дифференциальные уравнения в частных производных второго порядка. Построение качественно новых лагранжианов, описывающих новые (особенно нелинейные) физические процессы является, в известном смысле, искусством.

Функционалу (127) можно придать чисто геометрический смысл и считать его не интегралом в пространстве Минковского от лагранжиана $\mathfrak{L}$, а объемом, если $\mathfrak{L} > 0$, в некотором более сложном пространстве, в котором элемент объема имеет вид:

$$dV = \mathfrak{L} dx^0 dx^1 dx^2 dx^3. \quad (130)$$

Тогда уравнения поля получаются из требования экстремальности любого объема в этом новом пространстве, которое будем считать финслеровым. Это утверждение относится и к полю Мировой функции $S_W(x)$.

Рассмотрим финслерово пространство $x^1, x^2, \dots, x^n$ с метрической функцией

$$L(dx; x) \equiv L(dx^1, dx^2, \dots, dx^n; x^1, x^2, \dots, x^n). \quad (131)$$

В таком пространстве элемент длины ds по определению есть

$$ds = L(dx^1, dx^2, \dots, dx^n; x^1, x^2, \dots, x^n). \quad (132)$$

Более наглядно метрические свойства финслерова пространства можно описать с помощью понятия индикатрисы, которая определяется в каждой точке $M(x^1, x^2, \dots, x^n)$ основного пространства в соответствующем касательном центроаффинном пространстве $\xi^1, \xi^2, \dots, \xi^n$ как геометрическое место концов единичных радиус-векторов $\xi_{(ind)}$. Такая гиперповерхность описывается уравнением

$$L(\xi^1, \xi^2, \dots, \xi^n; x^1, x^2, \dots, x^n) = 1. \quad (133)$$

Задание системы индикатрис в каждой точке основного пространства, или, что тоже самое, задание множеств единичных векторов, полностью определяет финслерову геометрию. Для того, чтобы вычислить длину вектора $(dx^1, dx^2, \dots, dx^n)$, надо найти сонаправленный вектору dx единичный вектор $\xi_{(ind)}$, тогда числовой коэффициент ds в формуле

$$dx^i = ds \cdot \xi_{(ind)}^i \quad (134)$$

и есть длина вектора dx . Из этой формулы следует, что элемент длины

$$ds = \frac{|dx|_{ev}}{|\xi_{(ind)}|_{ev}}, \quad (135)$$

где $|dx|_{ev}$, $|\xi|_{ev}$ - длины векторов $(dx^1, dx^2, \dots, dx^n)$ и $(\xi^1, \xi^2, \dots, \xi^n)$, вычисляемые так, как если бы касательное центроаффинное пространство было бы евклидовым, а система координат в нём - декартова прямоугольная.

Если при выполнении последних предположений можно вычислить *объем индикатрисы*, то есть n -мерный объем, который зачерчивает единичный вектор $\xi_{(ind)}$ в касательном пространстве $\xi^1, \xi^2, \dots, \xi^n$, пробегая всю индикатрису, то в финслеровом пространстве можно аналогично (135) определить элемент объема dV по формуле [21]

$$dV = const \cdot \frac{dx^1 dx^2 \dots dx^n}{(V_{ind})_{ev}}, \quad (136)$$

где $(V_{ind})_{ev}$ - объем индикатрисы, вычисленный в предположении, что касательное центроаффинное пространство евклидово, а координаты декартовы прямоугольные. Таким образом определенный элемент объема инвариантен относительно преобразования координат.

Докажем последнее утверждение. Для этого рассмотрим переход от произвольных криволинейных координат $x^{1'}, x^{2'}, \dots, x^{n'}$ к другим произвольным криволинейным координатам $x^1, x^2, \dots, x^n$. Тогда согласно известным формулам [9] преобразования элемента объема имеем

$$dx^{1'} dx^{2'} \dots dx^{n'} = \left| \frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)} \right| dx^1 dx^2 \dots dx^n,$$

где

$$\frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)}$$

- якобиан. С другой стороны,

$$(V'_{ind})_{ev} = \int_{ind}^n d\xi^{1'} d\xi^{2'} \dots d\xi^{n'} = \int_{ind}^n \left| \frac{D(\xi^{1'}, \xi^{2'}, \dots, \xi^{n'})}{D(\xi^1, \xi^2, \dots, \xi^n)} \right| d\xi^1 d\xi^2 \dots d\xi^n,$$

но так как производные

$$\frac{\partial \xi^{i'}}{\partial \xi^j} = \frac{\partial x^{i'}}{\partial x^j},$$

не зависят от координат касательного пространства, что следует из правила их преобразования

$$\xi^{i'} = \frac{\partial x^{i'}}{\partial x^j} \xi^j,$$

то и якобиан

$$\frac{D(\xi^{1'}, \xi^{2'}, \dots, \xi^{n'})}{D(\xi^1, \xi^2, \dots, \xi^n)} = \frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)}$$

также не зависит от координат касательного пространства, и поэтому

$$(V'_{ind})_{ev} = \left| \frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)} \right| \int_{ind}^n d\xi^1 d\xi^2 \dots d\xi^n,$$

или

$$(V'_{ind})_{ev} = \left| \frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)} \right| (V_{ind})_{ev}. \quad (*)$$

Таким образом, при переходе к новой системе координат в правой части формулы (136) в числителе и знаменателе появляется один и тот же коэффициент - модуль якобиана перехода, сокращая в числителе и знаменателе этот коэффициент, получим

$$\frac{dx^{1'} dx^{2'} \dots dx^{n'}}{(V'_{ind})_{ev}} = \frac{dx^1 dx^2 \dots dx^n}{(V_{ind})_{ev}}.$$

Тем самым доказано, что при переходе от одних координат к другим формула (136) остаётся неизменной.

Если в формуле (136) положить $const$ равной единице, то при вычислении объёма его численное значение равно числу объёмов индикатрис, "уместившихся" в нём.

Рассмотрим n -мерное риманово пространство, метрическая функция в этом случае имеет вид

$$L(dx; x) = \sqrt{g_{ij}(x) dx^i dx^j}, \quad (137)$$

а уравнение индикатрисы соответственно -

$$g_{ij}(x) \xi^i \xi^j = 1. \quad (138)$$

Это уравнение определяет гиперповерхность второго порядка, а именно, эллипсоид. Если считать пространство $\xi^1, \xi^2, \dots, \xi^n$ евклидовым, то, как известно, объём такого эллипсоида есть

$$(V_{ind})_{ev} = \frac{const'}{\sqrt{\det(g_{ij})}}. \quad (139)$$

Докажем это. С помощью некоего линейного преобразования уравнение эллипсоида (138) в евклидовом пространстве всегда можно привести к виду

$$\frac{(\xi^{1'})^2}{R^2} + \frac{(\xi^{2'})^2}{R^2} + \dots + \frac{(\xi^{n'})^2}{R^2} = 1,$$

то есть в этой специальной системе координат

$$\det(g_{i'j'}(x')) = \frac{1}{R^{2n}},$$

и формула (139) выполняется. Учитывая, что

$$g_{i'j'}(x') = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} g_{ij}(x)$$

и

$$\det(g_{i'j'}(x')) = \left[\frac{D(x^{1'}, x^{2'}, \dots, x^{n'})}{D(x^1, x^2, \dots, x^n)} \right]^{-2} \det(g_{ij}(x)), \quad (**)$$

подставим (*) и (**) в выполняющуюся, как мы выше выяснили, для эллипсоида формулу в евклидовом пространстве в специальной системе координат

$$(V'_{ind})_{ev} = \frac{const'}{\sqrt{\det(g_{i'j'})}}.$$

В результате получим формулу (139) в произвольной системе координат. Тем самым доказано выполнение формулы (139) в произвольной системе координат.

Подставляя выражение (139) в формулу (136), получим элемент объема в произвольном римановом пространстве

$$dV = const \cdot \sqrt{\det(g_{ij})} \, dx^1 dx^2 \dots dx^n, \quad (140)$$

что совпадает с обычным определением инвариантного элемента объема в римановом пространстве.

Для псевдоримановых пространств без специальных дополнительных условий на индикатрису

$$(V_{ind})_{ev} = \infty \quad \Rightarrow \quad dV = 0 \cdot dx^1 dx^2 \dots dx^n, \quad (141)$$

но можно провести цепочку рассуждений, которая позволяет и для псевдоримановых пространств предложить инвариантный элемент объема в виде формулы, аналогичной формуле (140). Точно такие же рассуждения приходится приводить и для получения инвариантного элемента объема в неквадратичных финслеровых пространствах, в которых имеет место проблема (141). При этом удобно исходить из некоторого плоского пространства, близкого к пространству, элемент объема которого мы хотим определить.

Проведем эти рассуждения для конкретного примера: псевдориманово пространства с сигнатурой $(+, -, -, -)$, то есть пространства Минковского x^0, x^1, x^2, x^3 , в котором метрическая функция имеет вид

$$L(dx) = \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2} \equiv \sqrt{g_{ij}^o dx^i dx^j}, \quad (142)$$

а тангенциальное уравнение индикатрисы

$$(\xi^0)^2 - (\xi^1)^2 - (\xi^2)^2 - (\xi^3)^2 = 1 \quad (143)$$

является уравнением второго порядка и определяет гиперповерхность, которая является двуполостным гиперболоидом, поэтому проблема определения объема индикатрисы имеет место. Так как и метрическая функция, и уравнение индикатрисы не зависят от точки пространства, то каким бы образом мы не регуляризовали соответствующий интеграл, мы получим действительное число, одно и то же во всех точках пространства. Обозначим это число $(V_{ind})_{ev}$. Для того, чтобы в пространстве Минковского получить инвариантный элемент объема по формуле (136) величину $(V_{ind})_{ev}$ следует записать в виде

$$(V_{ind})_{ev} = \frac{const'}{\sqrt{-det(\overset{o}{g}_{ij})}}. \quad (144)$$

Перейдем от координат x^0, x^1, x^2, x^3 к некоторым криволинейным координатам $x^{0'}, x^{1'}, x^{2'}, x^{3'}$, тогда $\overset{o}{g}_{ij}$ перейдет в $g(x')_{i'j'}$, а значит элемент объема в пространстве Минковского в криволинейных координатах $x^{0'}, x^{1'}, x^{2'}, x^{3'}$ будет определяться формулой

$$dV = const \cdot \sqrt{-det(g(x')_{i'j'})} \, dx^{0'} dx^{1'} dx^{2'} dx^{3'}, \quad (145)$$

но все же пока мы остались в том же самом пространстве Минковского.

Рассмотрим псевдориманово пространство конформно связанное с пространством Минковского

$$ds = \kappa(x) \cdot \sqrt{\overset{o}{g}_{ij} dx^i dx^j}, \quad (146)$$

где $\kappa(x) > 0$, которое (пространство) никаким преобразованием координат не может быть переведено в пространство Минковского. Уравнение индикатрисы для такого псевдориманово пространства можно записать следующим образом:

$$(\xi^0)^2 - (\xi^1)^2 - (\xi^2)^2 - (\xi^3)^2 = \frac{1}{\kappa^2(x)}. \quad (147)$$

Сравнивая это уравнение с уравнением (143), видим, что гиперповерхность, определяемая уравнением (147), получается из гиперповерхности, определяемой (143), общим масштабным преобразованием с коэффициентом $\kappa^{-1}(x)$, поэтому если индикатрисе (143) мы приписывали объем (144), то индикатрисе (147) следует приписать объем

$$(V_{ind})_{ev} = \frac{const'}{\kappa^4(x) \sqrt{-det(\overset{o}{g}_{ij})}} = \frac{const'}{\sqrt{-det(g(x)_{ij})}}, \quad (148)$$

где

$$g(x)_{ij} \equiv \kappa^2(x) \overset{o}{g}_{ij}. \quad (149)$$

Из выше проведенных рассуждений следует, что псевдориманову пространству с метрическим тензором $g(x)_{ij}$ и сигнатурой $(+, -, -, -)$ можно приписать инвариантный элемент объема вида

$$dV = \text{const} \cdot \sqrt{-\det(g(x)_{ij})} \quad dx^0 dx^1 dx^2 dx^3, \quad (150)$$

что и принято в ОТО.

К проблеме (141) в псевдоримановых пространствах можно подойти более строго, но при этом приходится рассматривать пространства более общие, чем псевдоримановы. Поясним это на примере пространства Минковского. Если вместо пространства Минковского с метрической функцией (142) взять финслерово пространство с метрической функцией

$$L(dx) = \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2} + q_0 dx^0 \quad (151)$$

и условием $dx^0 \geq 0$, где $q_0 > 0$, то у такого пространства объем индикатрисы $(V_{ind})_{ev}$ конечное действительное число, зависящее от параметра q_0 и стремящееся к ∞ , когда параметр q_0 стремится к 0. Выбирая параметр q_0 достаточно малым, но неравным нулю, получим $(V_{ind}(q_0))_{ev} < \infty$ - конечное действительное число, зависящее от параметра q_0 , которое затем можно использовать для получения значений физических величин при $q_0 \rightarrow 0$.

Будем считать, что в любом финслеровом пространстве, в котором имеет место проблема (141), она может быть разрешена тем или иным способом, то есть может быть получен объем индикатрисы в каждой точке основного пространства, если считать касательное пространство евклидовым, а координаты декартовыми прямоугольными. Тогда можно утверждать, что из самой геометрии финслерова пространства, если метрическая функция содержит некоторые поля, автоматически получается лагранжиан [21]

$$\mathcal{L} = \frac{\text{const}}{(V_{ind})_{ev}}, \quad (152)$$

а значит и уравнения поля, и законы сохранения [7]. Несомненно, это справедливо и для Мировой функции, то есть Мировая функция также должна являться решением некоторого уравнения поля.

Замечание. Постоянные, которые фигурируют в формулах (136), (139), ..., (152), можно опускать или выбирать из соображений размерности. Все эти постоянные не входят в полевые уравнения, но наиболее правильно выбирать их так, чтобы при получении, например, закона сохранения энергии, получалось правильное значение энергии с учётом размерности.

Сформулируем кратко основные результаты.

Принцип самодостаточности финслеровой геометрии:

если метрическая функция финслерова пространства содержит неопределённые поля, то эти поля должны удовлетворять уравнениям поля, следующих из экстремальности любого объема финслерова пространства при отсутствии вариации полей на границе рассматриваемого объема.

Основные формулы.

Связь между элементом объема dV финслерова пространства и полевым лагранжианом $\mathcal{L}$ имеет вид

$$dV = \text{const} \quad \mathcal{L} \quad dx^1 dx^2 \dots dx^n,$$

а полевой лагранжиан выражается через объём индикатрисы по формуле

$$\mathfrak{L} = \frac{const}{(V_{ind})_{ev}}.$$

Рассмотрим несколько простейших примеров получения уравнений поля, решениями которых будет являться Мировая функция $S_W(x)$. Здесь же можно проследить связь между Мировой функцией и коэффициентом растяжения-сжатия $\kappa(x)$, с помощью которого метрическая функция одного финслерова пространства превращается в метрическую функцию другого финслерова пространства.

Пространство, конформно связанное с n -мерным евклидовым пространством, обладает элементом длины

$$ds = \kappa(x) \cdot \sqrt{(dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2}, \quad (153)$$

где $\kappa(x) > 0$. Так как в этом случае

$$\sqrt{\det(g_{ij})} = \kappa^n(x), \quad (154)$$

лагранжиан имеет вид

$$\mathfrak{L} = \kappa^n(x). \quad (155)$$

Для того, чтобы получить из такого лагранжиана уравнение поля, надо выразить скалярное поле $\kappa(x)$ через другое поле так, что лагранжиан уже будет содержать производные нового поля.

Обобщенные импульсы в пространстве (153) определяются формулой

$$p_i = \kappa(x) \frac{dx^i}{\sqrt{(dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2}}, \quad (156)$$

а тангенциальное уравнение индикатрисы можно записать следующим образом:

$$p_1^2 + p_2^2 + \dots + p_n^2 - \kappa^2(x) = 0. \quad (157)$$

Скалярная функция $S(x)$, которая в пространстве $x^1, x^2, \dots, x^n$ определяет нормальную конгруэнцию геодезических и которую в классической механике принято называть действием как функцией координат, должна удовлетворять уравнению Гамильтона - Якоби

$$\left(\frac{\partial S}{\partial x^1}\right)^2 + \left(\frac{\partial S}{\partial x^2}\right)^2 + \dots + \left(\frac{\partial S}{\partial x^n}\right)^2 = \kappa^2(x), \quad (158)$$

но мы это уравнение будем рассматривать как соотношение для выражения поля $\kappa(x)$ через поле $S(x)$. Таким образом,

$$\mathfrak{L} = \left[\left(\frac{\partial S}{\partial x^1}\right)^2 + \left(\frac{\partial S}{\partial x^2}\right)^2 + \dots + \left(\frac{\partial S}{\partial x^n}\right)^2 \right]^{\frac{n}{2}}, \quad (159)$$

а уравнение поля (129) принимает вид:

$$\frac{\partial}{\partial x^i} \left\{ \frac{\partial S}{\partial x^i} \left[\left(\frac{\partial S}{\partial x^1}\right)^2 + \left(\frac{\partial S}{\partial x^2}\right)^2 + \dots + \left(\frac{\partial S}{\partial x^n}\right)^2 \right]^{\frac{n}{2}-1} \right\} = 0. \quad (160)$$

Отметим, что для $n > 2$ это уравнение является нелинейным дифференциальным уравнением в частных производных второго порядка.

Запишем для пространства, конформно связанного с двумерной евклидовой плоскостью (x, y) , уравнение (160),

$$\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} = 0, \quad (161)$$

то есть функция $S(x, y)$ должна удовлетворять уравнению Лапласа, а значит она является компонентой аналитической функции комплексной переменной. Тогда

$$\kappa(x, y) = \sqrt{\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2} \quad (162)$$

- это коэффициент конформного преобразования элемента длины евклидова пространства, так как

$$ds' = \sqrt{(x')^2 + (y')^2} = \kappa(x, y) \sqrt{x^2 + y^2} \quad (163)$$

при конформном преобразовании

$$x' = u(x, y), \quad y' = \pm v(x, y), \quad (164)$$

когда функция S является одной из компонент аналитической функции $u + iv$ комплексной переменной $x + iy$.

Найдем решение уравнения (160) в предположении, что функция S зависит только от радиуса

$$r = \sqrt{(x^1)^2 + (x^2)^2 + \dots + (x^n)^2}. \quad (165)$$

Это удобнее сделать, записав элемент объема в сферических координатах и проинтегрировав по всем углам,

$$dV_r = r^{n-1} \left| \frac{dS}{dr} \right|^n dr \quad \Rightarrow \quad \mathfrak{L}_r = r^{n-1} \left| \frac{dS}{dr} \right|^n. \quad (166)$$

Тогда уравнение поля запишется следующим образом:

$$\frac{d}{dr} \left[r^{n-1} \left| \frac{dS}{dr} \right|^{n-1} \right] = 0. \quad (167)$$

Интегрируя это уравнение, получим

$$\frac{dS}{dr} = \frac{C}{r}, \quad S = C \ln \frac{r}{r_0} + C_0, \quad (168)$$

$C \neq 0$, C_0 , $r_0 > 0$ - действительные числа, тогда

$$\kappa(x) = \left| \frac{dS}{dr} \right| = \frac{|C|}{r}. \quad (169)$$

Геометрические из нормальной конгруэнции геодезических, описываемой функцией $S(r)$ (168), в таком пространстве определяются уравнениями

$$\dot{x}^i = \frac{\partial S}{\partial x^i} \cdot \lambda(x), \quad (170)$$

где $\lambda(x) \neq 0$ - некоторая функция, $\dot{x}^i$ - производная x^i по параметру τ вдоль геодезической. Выберем $\lambda(x) = \frac{r^2}{C}$, тогда из (170) следует, что

$$\dot{x}^i = x^i. \quad (171)$$

Пусть $j > 1$, тогда

$$\frac{dx^j}{dx^1} = \frac{x^j}{x^1} \Rightarrow x^j = C^j x^1, \quad (172)$$

то есть геодезические из нормальной конгруэнции геодезических, которая (конгруэнция) определяется Мировой функцией (168), в таком пространстве - это прямые линии, проходящие через начало координат с направляющим вектором $(1, C^2, C^3, \dots, C^n)$. Так как в качестве параметра эволюции можно выбрать любую координату x^i , то направляющим вектором может быть любой вектор $(k^1, k^2, \dots, k^n)$ нашего евклидова пространства.

Пространство, конформно связанное с псевдоевклидовым пространством с сигнатурой $(+, -, -, \dots, -)$, обладает элементом длины

$$ds = \kappa(x) \cdot \sqrt{(dx^0)^2 - (dx^1)^2 - \dots - (dx^{n-1})^2}, \quad (173)$$

где $\kappa(x) > 0$. Так как в этом случае

$$\sqrt{(-1)^{n-1} \det(g_{ij})} = \kappa^n(x), \quad (174)$$

лагранжиан имеет вид

$$\mathcal{L} = \kappa^n(x). \quad (175)$$

Для того, чтобы получить из такого лагранжиана уравнение поля, надо выразить скалярное поле $\kappa(x)$ через другое поле так, чтобы лагранжиан содержал производные нового поля.

Обобщенные импульсы в пространстве (173) определяются формулами

$$\left. \begin{aligned} p_0 &= \frac{\kappa(x) dx^0}{\sqrt{(dx^0)^2 - (dx^1)^2 - \dots - (dx^{n-1})^2}}, \\ p_\mu &= \frac{\kappa(x) dx^\mu}{\sqrt{(dx^0)^2 - (dx^1)^2 - \dots - (dx^{n-1})^2}}, \end{aligned} \right\} \quad (176)$$

$\mu = 1, 2, \dots, (n-1)$, а тангенциальное уравнение индикатрисы можно записать следующим образом:

$$p_0^2 - p_1^2 - \dots - p_{n-1}^2 - \kappa^2(x) = 0. \quad (177)$$

Скалярная функция $S(x)$, которая в пространстве $x^0, x^1, \dots, x^{n-1}$ определяет нормальную конгруэнцию геодезических, должна удовлетворять уравнению Гамильтона - Якоби

$$\left(\frac{\partial S}{\partial x^0} \right)^2 - \left(\frac{\partial S}{\partial x^1} \right)^2 - \dots - \left(\frac{\partial S}{\partial x^{n-1}} \right)^2 = \kappa^2(x), \quad (178)$$

но это уравнение можно рассматривать как связь между скалярными полями $\kappa(x)$ и $S(x)$. Таким образом,

$$\mathfrak{L} = \left[\left(\frac{\partial S}{\partial x^0} \right)^2 - \left(\frac{\partial S}{\partial x^1} \right)^2 - \dots - \left(\frac{\partial S}{\partial x^{n-1}} \right)^2 \right]^{\frac{n}{2}}, \quad (179)$$

а уравнение поля (129) принимает вид:

$$\begin{aligned} & \frac{\partial}{\partial x^0} \left\{ \frac{\partial S}{\partial x^0} \left[\left(\frac{\partial S}{\partial x^0} \right)^2 - \left(\frac{\partial S}{\partial x^1} \right)^2 - \dots - \left(\frac{\partial S}{\partial x^{n-1}} \right)^2 \right]^{\frac{n}{2}-1} \right\} - \\ & - \frac{\partial}{\partial x^\mu} \left\{ \frac{\partial S}{\partial x^\mu} \left[\left(\frac{\partial S}{\partial x^0} \right)^2 - \left(\frac{\partial S}{\partial x^1} \right)^2 - \dots - \left(\frac{\partial S}{\partial x^{n-1}} \right)^2 \right]^{\frac{n}{2}-1} \right\} = 0. \end{aligned} \quad (180)$$

Отметим, что для $n > 2$ это уравнение является нелинейным дифференциальным уравнением в частных производных второго порядка.

Для пространства, конформно связанного с двумерной псевдоевклидовой плоскостью (x, y) , уравнение (180) даёт

$$\frac{\partial^2 S}{\partial x^2} - \frac{\partial^2 S}{\partial y^2} = 0, \quad (181)$$

то есть в двумерном случае уравнение поля (180) - это волновое уравнение, общее решение которого имеет вид:

$$S = f_1(x + y) + f_2(x - y),$$

где f_1, f_2 произвольные дважды дифференцируемые функции одной действительной переменной.

Для пространства, конформно связанного с пространством Минковского $n = 4$, используя метрический тензор $\overset{o}{g}_{ij}$, получим: связь между функцией $S(x)$ и коэффициентом растяжения-сжатия $\kappa(x)$ -

$$\overset{o}{g}^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = \kappa^2(x), \quad (182)$$

лагранжиан -

$$\mathfrak{L} = \left(\overset{o}{g}^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} \right)^2, \quad (183)$$

уравнение поля -

$$\overset{o}{g}^{kl} \frac{\partial}{\partial x^k} \left[\frac{\partial S}{\partial x^l} \left(\overset{o}{g}^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} \right) \right] = 0. \quad (184)$$

Запишем уравнение поля для Мировой функции в финслеровом пространстве, которое является конформно связанным с пространством H_4 , или с четырёхмерным пространством Бервальда - Моора. Метрическая функция такого пространства определяется формулой (111), тангенциальное уравнение можно записать в виде (113), а коэффициент растяжения-сжатия

$\kappa(x)$ выражается через Мировую функцию $S_W(x)$ по формуле (116). Таким образом, лагранжиан для получения уравнения поля для Мировой функции в специальной изотропной системе координат запишется следующим образом:

$$\mathfrak{L} = \text{const} \frac{\partial S_W}{\partial x^1} \frac{\partial S_W}{\partial x^2} \frac{\partial S_W}{\partial x^3} \frac{\partial S_W}{\partial x^4}. \quad (185)$$

Уравнение Лагранжа - Эйлера (129), то есть уравнение поля принимает вид:

$$\begin{aligned} & \frac{\partial}{\partial x^1} \left(\frac{\partial S_W}{\partial x^2} \frac{\partial S_W}{\partial x^3} \frac{\partial S_W}{\partial x^4} \right) + \frac{\partial}{\partial x^2} \left(\frac{\partial S_W}{\partial x^1} \frac{\partial S_W}{\partial x^3} \frac{\partial S_W}{\partial x^4} \right) + \\ & + \frac{\partial}{\partial x^3} \left(\frac{\partial S_W}{\partial x^1} \frac{\partial S_W}{\partial x^2} \frac{\partial S_W}{\partial x^4} \right) + \frac{\partial}{\partial x^4} \left(\frac{\partial S_W}{\partial x^1} \frac{\partial S_W}{\partial x^2} \frac{\partial S_W}{\partial x^3} \right) = 0. \end{aligned} \quad (186)$$

Это нелинейное дифференциальное уравнение в частных производных второго порядка с дополнительным условием

$$\kappa(x) > 0. \quad (187)$$

В силу этого условия, мы исключаем, например, такие функции в качестве возможных решений, как:

$$f_1(x^1), \quad \dot{f}_{34}(x^1, x^2), \quad \bar{f}_1(x^2, x^3, x^4), \quad (188)$$

где f_1 , $\dot{f}_{34}$, $\bar{f}_1$ - произвольные функции одной, двух и соответственно трёх действительных переменных, так как, хотя они удовлетворяют уравнению (186), для них не выполняется условие (187).

Непосредственной подстановкой в уравнение (186) проверяется, что этому уравнению, например, удовлетворяет функция

$$S_W = \frac{1}{4} [f_1(x^1) + f_2(x^2) + f_3(x^3) + f_4(x^4)], \quad (189)$$

где функции f_i должны удовлетворять условию

$$\frac{\partial f_1}{\partial x^1} \frac{\partial f_2}{\partial x^2} \frac{\partial f_3}{\partial x^3} \frac{\partial f_4}{\partial x^4} > 0, \quad (190)$$

а в остальном произвольны.

Ещё один пример решения задачи (186), (187). Будем искать решение в виде:

$$S_W(x) = u(x^1, x^2) + v(x^3, x^4). \quad (191)$$

Тогда, подставляя (191) в (186), получим

$$\frac{\partial v}{\partial x^3} \frac{\partial v}{\partial x^4} \frac{\partial^2 u}{\partial x^1 \partial x^2} + \frac{\partial u}{\partial x^1} \frac{\partial u}{\partial x^2} \frac{\partial^2 v}{\partial x^3 \partial x^4} = 0. \quad (192)$$

Это уравнение с частично разделяющимися переменными. Так как из (187) следует условие

$$\frac{\partial u}{\partial x^1} \frac{\partial u}{\partial x^2} \frac{\partial v}{\partial x^3} \frac{\partial v}{\partial x^4} > 0, \quad (193)$$

разделим левую и правую части уравнения (192) на

$$\frac{\partial u}{\partial x^1} \frac{\partial u}{\partial x^2} \frac{\partial v}{\partial x^3} \frac{\partial v}{\partial x^4}.$$

Получим в правой части уравнения два аддитивных члена, первый из которых зависит только от x^1, x^2 , а второй - только от x^3, x^4 , поэтому каждый из них есть постоянная, причём, если первый равен $(-\lambda)$, то второй должен быть равен λ :

$$\frac{\frac{\partial^2 u}{\partial x^1 \partial x^2}}{\frac{\partial u}{\partial x^1} \frac{\partial u}{\partial x^2}} = -\lambda, \quad \frac{\frac{\partial^2 v}{\partial x^3 \partial x^4}}{\frac{\partial v}{\partial x^3} \frac{\partial v}{\partial x^4}} = \lambda. \quad (194)$$

Интегрируя эту систему уравнений, имеем

$$S_W = \frac{1}{\lambda} \ln \left| \frac{f_1(x^1) + f^2(x^2)}{f_3(x^3) + f_4(x^4)} \right|, \quad (195)$$

где функции f_i должны удовлетворять условию (187), то есть

$$\kappa^4(x) \equiv \frac{4^4}{\lambda^4} \frac{\dot{f}_1 \dot{f}_2 \dot{f}_3 \dot{f}_4}{(f_1 + f_2)^2 (f_3 + f_4)^2} > 0, \quad (196)$$

или

$$\lambda^4 \neq 0, \pm\infty, \quad f_1 + f_2, f_3 + f_4 \neq 0, \pm\infty, \quad \dot{f}_1 \dot{f}_2 \dot{f}_3 \dot{f}_4 > 0, \quad (197)$$

а в остальном произвольны.

Важно отметить, что решение S_W (195) не может быть представлено в виде (189).

7 Гиперкомплексные числа

Гиперкомплексная система чисел - это линейная алгебра с единицей. Более развёрнутое определение звучит следующим образом.

Линейное пространство над полем действительных чисел с бинарной операцией, линейной по каждому сомножителю, и единицей называется *гиперкомплексной алгеброй*, или *системой гиперкомплексных чисел* над полем действительных чисел.

Гиперкомплексные числа, кроме базисных элементов, обычно будем обозначать большими латинскими буквами $A, B, C, \dots, X, Y, Z$; действительные числа - в основном малыми латинскими и греческими буквами. Базисные (символьные) элементы системы гиперкомплексных чисел в основном будем обозначать $e_1, e_2, \dots, e_n$, где n - размерность гиперкомплексного пространства.

Если базис $e_1, e_2, \dots, e_n$ фиксирован, то любое гиперкомплексное число однозначно представимо в виде линейной комбинации базисных элементов

$$X = x^1 e_1 + x^2 e_2 + \dots + x^n e_n. \quad (198)$$

Умножение гиперкомплексного числа на действительное число и сложение гиперкомплексных чисел происходит покомпонентно по тем же законам, что и элементов линейного пространства [6], так как система гиперкомплексных чисел и есть линейное пространство. Бинарная операция полностью определяется законом умножения базисных элементов:

$$e_i \cdot e_j = p_{ij}^l \cdot e_l, \quad (199)$$

то есть числовым тензором p_{ij}^l , или *структурным тензором*. При этом бинарная операция *линейна* по каждому сомножителю:

$$(aA + bB)(cC + dD) = acAC + adAD + bcBC + bdBD. \quad (200)$$

Если представить, что начала всех векторов $X \in P_n$ находятся в одной фиксированной точке $\mathcal{O}$, то компоненты таких векторов (радиус-векторов) определяют нам n -мерное координатное пространство:

$$X = x^1 e_1 + x^2 e_2 + \dots + x^n e_n \quad \leftrightarrow \quad (x^1, x^2, \dots, x^n) \quad (201)$$

с координатами $x^1, x^2, \dots, x^n$ и началом координат в точке $\mathcal{O}(0, 0, \dots, 0)$. В этом координатном пространстве, как и в координатном пространстве свободных векторов (198), определена бинарная операция:

$$X \cdot Y = Z, \quad x^i x^j p_{ij}^l = z^l. \quad (202)$$

Система гиперкомплексных чисел называется коммутативной, если для любых двух чисел X и Y

$$XY = YX \quad \Leftrightarrow \quad p_{ij}^l = p_{ji}^l, \quad (203)$$

и ассоциативной, если для любых трёх чисел X, Y, Z

$$(XY)Z = X(YZ) \quad \Leftrightarrow \quad p_{ij}^l p_{lr}^s = p_{jr}^l p_{il}^s. \quad (204)$$

Если система гиперкомплексных чисел и коммутативна, и ассоциативна, то имеют место и другие соотношения для тензора p_{ij}^l , например,

$$p_{ij}^l p_{lr}^s = p_{rj}^l p_{li}^s. \quad (205)$$

Пусть ϵ^i - коэффициенты разложения единицы в базисе $e_1, e_2, \dots, e_n$, тогда справедливы следующие формулы:

$$\epsilon^i p_{ij}^l = \delta_j^l, \quad \epsilon^j p_{ij}^l = \delta_i^l. \quad (206)$$

Если гиперкомплексная алгебра не содержит нильпотентных элементов, то есть таких чисел $N \neq 0$, что $N^m = 0$, где $m > 1$ - некоторое натуральное число, то будем называть такую алгебру гиперкомплексных чисел *невыврожденной*.

Если для двух гиперкомплексных чисел $A \neq 0$ и $B \neq 0$ их произведение равно нулю $AB = 0$, то оба этих числа принято называть *делителями нуля*. В этом случае

$$a^i b^j p_{ij}^l = 0 \quad \Rightarrow \quad \det(b^j p_{ij}^l) = 0, \quad \det(a^i p_{ij}^l) = 0.$$

Достижения, которые были получены в теории комплексных чисел и особенно теории функций комплексной переменной, заставляют нас стремиться к тому, чтобы и в других метрических пространствах, не только на евклидовой плоскости, сопоставить взаимно однозначно

каждой точке пространства число, принадлежащее некоторой гиперкомплексной системе. Можно работать с некоммутативными и даже неассоциативными гиперкомплексными алгебрами, но при этом возникает ряд трудностей особенно в построении аналитических функций. Существует другой путь [18] - [20], которого будем придерживаться и мы: выделять самые удобные, самые "простые" числа, а именно невырожденные *ассоциативно-коммутативные гиперкомплексные числа*, или кратко - невырожденные *поличисла* P_n , и концентрировать усилия на их изучении и применении в физике.

Теорема Фробениуса [3]: Любая ассоциативная алгебра с делением изоморфна одной из трех: алгебре действительных чисел, алгебре комплексных чисел или алгебре кватернионов.

Так как алгебра кватернионов не коммутативна, то из теоремы Фробениуса следует, что невырожденные поличисла P_n , размерности $n > 2$ всегда содержат делители нуля, то есть найдутся такие числа $X, Y \neq 0$, что

$$XY = YX = 0, \quad (207)$$

поэтому на множестве поличисел размерности больше двух нельзя определить классическое понятие нормы числа, но можно определить близкое понятие, которое следовало бы назвать *квазинормой* или каким-то другим термином, мы же будем использовать старый термин.

Если на некотором подмножестве D_n поличисел P_n для любого числа X из этого подмножества задано действительное число $|X|$, причем для любых $X, Y \in D_n$ выполняются следующие три требования:

1. $|X| \geq 0$;
2. если $a > 0$ - действительное число, то $|aX| = a|X|$;
3. если $XY \in D_n$, то $|XY| = |X||Y|$

- будем говорить, что на множестве поличисел P_n задана норма, а D_n - область ее определения.

Понятие функции F на гиперкомплексной алгебре $HC \ni X, dX$ общего вида с значениями в той же самой алгебре $F(X) \in HC$ определяется обычным образом.

Если e - базис, а x - координаты гиперкомплексного числа X (198) в этом базисе, то

$$F(X) = f^i(x^1, x^2, \dots, x^n) e_i, \quad (208)$$

где f^i - n произвольных функций от n действительных переменных. Область определения функции $F(X)$ суть пересечение областей определения всех функций $f^1, f^2, \dots, f^n$.

Пусть dX - произвольное бесконечно малое гиперкомплексное число, то есть

$$dX = dx^i e_i, \quad dx^i \rightarrow 0. \quad (209)$$

Если

$$F(X + dX) - F(X) = F'(X) dX + o(dx^i), \quad (210)$$

то $F(X)$ - называют *аналитической функцией гиперкомплексной переменной*, а $F'(X)$ - её *производной*.

Из определения следует, что, во-первых, для того, чтобы $F(X)$ была аналитической функцией необходимо, чтобы функции f^i были хотя бы гладкими функциями всех n действительных переменных. Во-вторых, если гиперкомплексная система некоммутативна, то возникаю

серьёзные трудности с введением понятия аналитичности, ещё больше трудностей возникает при отсутствии ассоциативности. Это одна из причин, по которым в дальнейшем будут рассматриваться только ассоциативно-коммутативные гиперкомплексные числа, или кратко *поличисла*.

Для поличисел определение (210) является естественным и порождает естественный алгоритм построения, если не всех, то большого класса аналитических функций поличисловой переменной. Любая функция $F(X)$ поличисловой переменной, построенная как полином или сходящийся ряд, является аналитической функцией той же поличисловой переменной. Приведём два примера, аналитическими функциями поличисловой переменной являются:

$$F(X) = C X^2 + B X + C \quad (211)$$

и

$$e^{AX} = 1 + AX + \frac{1}{2!}(AX)^2 + \frac{1}{3!}(AX)^3 + \dots + \frac{1}{m!}(AX)^m + \dots, \quad (212)$$

где A, B, C - поличисловые постоянные. Все пространства невырожденных поличисел являются метрическими пространствами, а именно, финслеровыми пространствами, причём, если размерность пространства больше двух, то это метрические неквадратичные пространства.

Конечно, самым интересным вопросом для физиков является вопрос: какую поличисловую систему сопоставить пространству событий, то есть четырёхмерному пространству-времени? Одна из возможностей - это поличисла H_4 [18] - [20], которые изоморфны алгебре действительных квадратных диагональных матриц 4×4 и которые являются финслеровым пространством с метрической функцией Бервальда - Моора (47).

8 Поличисла H_3

В качестве примера невырожденных поличисел рассмотрим алгебру действительных квадратных диагональных матриц 3×3 . Эта алгебра изоморфна невырожденным поличислам H_3 .

Любой элемент $X \in H_3$ можно рассматривать как матрицу вида

$$X = \begin{pmatrix} a^1 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^3 \end{pmatrix} \equiv a^1 e_1 + a^2 e_2 + a^3 e_3, \quad (213)$$

где a^1, a^2, a^3 - действительные числа, а $e_1, e_2, e_3 \in H_3$ - так называемый, изотропный базис:

$$e_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (214)$$

Роль единичного элемента выполняет единичная матрица, то есть сумма базисных элементов e_1, e_2 и e_3 :

$$e_1 + e_2 + e_3 \equiv \hat{1}.$$

Так как для данных поличисел, если m - натуральное число, из

$$X^m = 0 \quad \Rightarrow \quad X = 0,$$

то поличисла H_3 являются невырожденными.

Для произвольной квадратной матрица A , как известно, определена функция

$$\exp(A) \equiv \hat{1} + A + \frac{1}{2!}A^2 + \dots + \frac{1}{m!}A^m + \dots \quad (215)$$

Здесь $\hat{1}$ - единичная матрица. Тогда для X (213) имеем

$$\exp(X) = \begin{pmatrix} e^{a^1} & 0 & 0 \\ 0 & e^{a^2} & 0 \\ 0 & 0 & e^{a^3} \end{pmatrix}. \quad (216)$$

Рассмотрим ещё один базис $1, j, k$, который связан с изотропным базисом преобразованием

$$\begin{pmatrix} 1 \\ j \\ k \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \quad (217)$$

и попытаемся представить поличисло X в экспоненциальной форме, то есть следующим образом:

$$X = x^0 1 + x^1 j + x^2 k = \varrho \cdot \exp(\alpha j + \beta k), \quad (218)$$

где α, β - любые действительные числа, а действительное число $\varrho > 0$. Воспользовавшись формулой (216), получим

$$a^1 e_1 + a^2 e_2 + a^3 e_3 = \varrho e^{\alpha+\beta} e_1 + \varrho e^{-\alpha} e_2 + \varrho e^{-\beta} e_3. \quad (219)$$

Из этого соотношения следует: для того, чтобы поличисло X можно было представить в экспоненциальном виде, все a^i должны быть больше нуля, $a^i > 0$. При этом величины ϱ, α, β выражаются через координаты X в изотропном базисе по формулам:

$$\varrho = \sqrt[3]{a^1 a^2 a^3}, \quad \alpha = -\ln \frac{a^2}{\varrho}, \quad \beta = -\ln \frac{a^3}{\varrho}. \quad (220)$$

Если два числа $X, Y \in H_3$ представимы в экспоненциальном виде, то при их перемножении

$$X_1 X_2 = Z \Rightarrow Z = \varrho_1 \varrho_2 \cdot \exp[(\alpha_1 + \alpha_2)j + (\beta_1 + \beta_2)k]. \quad (221)$$

Таким образом, для величины ϱ в экспоненциальном представлении выполняются все требования, сформулированные выше, для нормы поличисла, поэтому для $X \in H_3$

$$|X| = \sqrt[3]{a^1 a^2 a^3}, \quad a^1, a^2, a^3 > 0. \quad (222)$$

Норма поличисла неотъемлема от экспоненциального представления поличисла, поэтому область определения нормы задаётся именно тремя неравенствами $a^1 > 0, a^2 > 0, a^3 > 0$, а не одним неравенством $a^1 a^2 a^3 > 0$.

Пусть для числа $X \in H_3$ определена норма (222), умножим его на унимодулярное число $g \in H_3$, то есть число, которое представимо в экспоненциальном виде и которое имеет единичную норму $|g| = 1$. Тогда

$$|Xg| = |X|, \quad (223)$$

а значит в пространстве H_3 имеется двухпараметрическая абелева группа Ли $G_1(H_3)$, сохраняющая норму поличисел, причём элементами этой группы являются унимодулярные поличисла H_3 , а групповое умножение суть поличисловая бинарная операция.

Пространство H_3 является линейным пространством, но если мы будем считать векторы (213) радиус-векторами, то мы перейдём от аффинного к центроаффинному пространству и в этом случае координаты радиус-векторов являются точками координатного пространства (a^1, a^2, a^3) . Если в этом пространстве мы захотим вычислить длину кривой, то нам необходим элемент длины ds , вид которого однозначно следует из полученной выше нормы поличисла:

$$ds = \sqrt[3]{da^1 da^2 da^3}, \quad da^1 > 0, \quad da^2 > 0, \quad da^3 > 0. \quad (224)$$

Тем самым пространство H_3 является неквадратичным финслеровым пространством с изометрической непрерывной группой симметрии $G_1(H_3)$.

Нельзя ли рассматриваемое пространство H_3 , хотя бы в нулевом или первом приближении, считать пространством событий для физических задач, когда зависимость физических величин от какой-нибудь одной пространственной координаты отсутствует? Для положительного ответа на этот вопрос следует искать такой базис $1, J, K$,

$$X = x^0 \cdot 1 + x \cdot J + y \cdot K, \quad (225)$$

где $x^0 = ct$, c - скорость света, t - время, x, y - пространственные координаты, в котором (базисе) при нерелятивистских скоростях ($dx^0 \gg |dx|, |dy|$) с точностью до $\left|\frac{v}{c}\right|^2$ включительно элемент длины принимает вид:

$$ds = \sqrt[3]{da^1 da^2 da^3} \simeq dx^0 - \frac{dx^2 + dy^2}{2dx^0}. \quad (226)$$

При этом базисный элемент 1 менять не следует, а весь произвол в выборе нового базиса заключается в невырожденном линейном преобразовании базисных векторов j, k (четыре действительных параметра), что позволяет сохранить экспоненциальное представление в базисе $1, J, K$. Прежде всего перейдём к координатам x^0, x^1, x^2 в базисе $1, j, k$,

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ j \\ k \end{pmatrix}, \quad (227)$$

$$\begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \end{pmatrix}, \quad (228)$$

при тех же предположениях о малости скоростей $dx^0 \gg |dx^1|, |dx^2|$:

$$\begin{aligned} ds &= \sqrt[3]{(x^0 + dx^1 + dx^2)(dx^0 - dx^1)(dx^0 - dx^2)} \simeq \\ &\simeq dx^0 - \frac{(dx^1)^2 + dx^1 dx^2 + (dx^2)^2}{3 dx^0}. \end{aligned} \quad (229)$$

Надо найти такое преобразование координат

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (230)$$

a, b, c, d - искомые действительные числа, чтобы последняя правая дифференциальная форма (229) превратилась в последнюю правую дифференциальную форму (226). Потребовав это, получим систему трёх уравнений с четырьмя неизвестными:

$$\left. \begin{aligned} 2ab + ad + bc + 2cd &= 0, \\ a^2 + ac + c^2 &= \frac{3}{2}, \\ b^2 + bd + d^2 &= \frac{3}{2}. \end{aligned} \right\} \quad (231)$$

Проще, однако, привести дифференциальную форму (229) к каноническому виду классическим способом: найти собственные векторы и собственные значения соответствующей билинейной формы, взять эти собственные векторы с некоторыми коэффициентами в качестве базисных векторов и учесть, что следует из формулы (226), неоднозначность выбора координат x, y , которые всегда можно повернуть на некоторый эллиптический угол φ . В результате получим

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \\ \frac{1}{\sqrt{2}} & -\sqrt{\frac{3}{2}} \end{pmatrix} \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (232)$$

или

$$\begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} 1 & \sqrt{2} & 0 \\ 1 & -\frac{1}{\sqrt{2}} & -\sqrt{\frac{3}{2}} \\ 1 & -\frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x^0 \\ x \\ y \end{pmatrix}. \quad (233)$$

Эти формулы не дают общего решения системы алгебраических уравнений (231), так как возможно, например, ещё поменять местами координаты x и y и/или независимо изменить знаки у этих координат, но главное получено: координатное пространство x^0, x, y можно считать трёхмерным пространством-временем, трёхмерным пространством событий.

Эллиптический поворот на угол φ в плоскости (x, y) можно выполнить дополнительно всегда, но хотелось бы иметь для матрицы преобразования (233) наиболее простое выражение, одно из таких выражений получается, если положить $\varphi = \frac{\pi}{4}$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{\sqrt{3}-1}{2} & -\frac{\sqrt{3}+1}{2} \\ 1 & -\frac{\sqrt{3}+1}{2} & \frac{\sqrt{3}-1}{2} \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 & 1 \\ 1 & \sqrt{2} \sin \frac{\pi}{12} & -\sqrt{2} \cos \frac{\pi}{12} \\ 1 & -\sqrt{2} \cos \frac{\pi}{12} & \sqrt{2} \sin \frac{\pi}{12} \end{pmatrix}. \quad (234)$$

Матрица (234) является симметрической.

Таким образом, непрерывные однопараметрические переходы в пространстве H_3 от одного физического базиса, в котором имеет место экспоненциальное представление поличисел

и нерелятивистский переход к трёхмерному пространству Галилей, к другому физическому базису образуют группу $SO(2)$.

Итак, поличисловое пространство H_3 есть финслерово пространство с метрической функцией

$$L(da) = \sqrt[3]{da^1 da^2 da^3}. \quad (235)$$

Компоненты обобщённого импульса определяются формулами

$$p_i = \frac{1}{3} \frac{\sqrt[3]{da^1 da^2 da^3}}{da^i} \quad (236)$$

и связаны тангенциальным уравнением индикатрисы

$$p_1 p_2 p_3 - \frac{1}{3^3} = 0. \quad (237)$$

Пространство, конформно связанное с поличисловым пространством H_3 , есть финслерово пространство с метрической функцией

$$\tilde{L}(da; a) = \kappa(a) \sqrt[3]{da^1 da^2 da^3}. \quad (238)$$

Компоненты обобщённого импульса в таком пространстве определяются формулами

$$\tilde{p}_i = \frac{\kappa(a)}{3} \frac{\sqrt[3]{da^1 da^2 da^3}}{da^i} \quad (239)$$

и связаны тангенциальным уравнением индикатрисы

$$\tilde{p}_1 \tilde{p}_2 \tilde{p}_3 - \frac{\kappa^3(a)}{3^3} = 0. \quad (240)$$

Откуда следует выражение коэффициента растяжения-сжатия $\kappa(a)$ через Мировую функцию $S_W(a)$

$$\kappa(a) = 3 \sqrt[3]{\frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^2} \frac{\partial S_W}{\partial a^3}}, \quad \frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^2} \frac{\partial S_W}{\partial a^3} > 0 \quad (241)$$

и выражение для лагранжиана

$$\mathfrak{L} = \text{const} \frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^2} \frac{\partial S_W}{\partial a^3}. \quad (242)$$

Соответственно уравнение поля для Мировой функции запишется как

$$\frac{\partial}{\partial a^1} \left(\frac{\partial S_W}{\partial a^2} \frac{\partial S_W}{\partial a^3} \right) + \frac{\partial}{\partial a^2} \left(\frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^3} \right) + \frac{\partial}{\partial a^3} \left(\frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^2} \right) = 0, \quad (243)$$

при дополнительном условии

$$\frac{\partial S_W}{\partial a^1} \frac{\partial S_W}{\partial a^2} \frac{\partial S_W}{\partial a^3} > 0. \quad (244)$$

Построить некоторое множество простейших аналитических функций поличисловой H_3 переменной не составляет труда: это полиномы и сходящиеся ряды переменной $X \in H_3$, причём в качестве параметров можно использовать как действительные числа, так и поличисла H_3 .

Так как в изотропном базисе умножение чисел H_3 происходит покомпонентно, то все такие аналитические функции поличисловой переменной, если параметрами являются только действительные числа, в том же изотропном базисе имеют вид:

$$F(X) = f(a^1)e_1 + f(a^2)e_2 + f(a^3)e_3, \quad (245)$$

где f - гладкая функция одной действительной переменной. Учитывая это и непосредственно обратившись к определению аналитической функции гиперкомплексной переменной (210), получим, что для поличисел H_3 в изотропном базисе любая функция вида

$$F(X) = f^1(a^1)e_1 + f^2(a^2)e_2 + f^3(a^3)e_3, \quad (246)$$

где f^i - три, вообще говоря, разные гладкие функции одной действительной переменной, является аналитической. Аналитические функции вида (246) получаются, если брать полиномы или сходящиеся ряды поличисловой H_3 переменной с параметрами, которые также являются поличислами H_3 .

Определим функцию $S_W(a)$ как произвольную линейную комбинацию компонент аналитической функции $F(X)$ (246) с действительными коэффициентами $\alpha_1, \alpha_2, \alpha_3$, то есть

$$S_W(a) = \alpha_1 f^1(a^1) + \alpha_2 f^2(a^2) + \alpha_3 f^3(a^3). \quad (247)$$

Прямой подстановкой этой функции в уравнение (243), проверяем, что данная функция является решением уравнения (243). Условие (244) превращается в следующее неравенство:

$$\alpha_1 \alpha_2 \alpha_3 \frac{df^1}{da^2} \frac{df^2}{da^2} \frac{df^3}{da^3} > 0. \quad (248)$$

9 Полилинейные формы

Идеи обобщения понятия скалярного произведения до скалярного полипроизведения, то есть скалярного произведения с числом "сомножителей" (аргументов) более двух, и изучения пространств с полилинейной формой для дальнейшего применения в теоретической физике принадлежат Д.Г. Павлову [18]. Такие пространства при выполнении ряда условий являются финслеровыми пространствами. С другой стороны финслеровы пространства невырожденных поличисел размерности больше двух относятся именно к пространствам со скалярным полипроизведением.

Пусть $A \ni X, Y, Z, \dots$ - некоторое линейное пространство над полем действительных чисел $R \ni x, y, z, a, b, c, \dots$, то есть в этом пространстве определено сложение двух произвольных элементов и умножение произвольного элемента на любое действительное число, причем

$$(a + b)(xA + yB) = (ax)A + (bx)A + (ay)B + (by)B. \quad (249)$$

Элементы линейного пространства будем называть векторами. Если для элементов $X, Y, \dots, Z$ найдутся такие действительные числа $a, b, \dots, c$, не все из которых равны нулю, что

$$aX + bY + \dots + cZ = 0, \quad (250)$$

то говорят, что элементы $X, Y, \dots, Z$ пространства A *линейно зависимы*. Если таких чисел найти невозможно, то элементы $X, Y, \dots, Z$ называются *линейно независимыми*.

Если в пространстве A имеется n линейно независимых элементов, а любые $(n + 1)$ элементов всегда линейно зависимы, то пространство называют n -мерным. Такое пространство будем обозначать A_n . В n -мерном линейном пространстве всегда можно фиксировать n линейно независимых элементов $e_1, e_2, \dots, e_n$, которые называются базисными, а сам такой набор - *базисом*, и выражать произвольный элемент X в виде линейной комбинации базисных элементов:

$$X = x^1 e_1 + x^2 e_2 + \dots + x^n e_n, \quad (251)$$

где $x^1, x^2, \dots, x^n$ - координаты элемента X в базисе $e_1, e_2, \dots, e_n$. Такое разложение однозначно, поэтому вместо того, чтобы работать непосредственно с элементами самого линейного n -мерного пространства, можно работать с наборами n действительных чисел $(x^1, x^2, \dots, x^n)$ - координатами элементов в данном базисе.

Будем говорить, что на пространстве A_n задана m -линейная ($m \geq 2$) симметрическая форма, или *скалярное произведение* m векторов, если для любого набора m элементов $X_{(1)}, X_{(2)}, \dots, X_{(m)}$ этого пространства определено действительное число $(X_{(1)}, X_{(2)}, \dots, X_{(m)})$, которое обладает следующими двумя свойствами. Во-первых, при любой перестановке аргументов это число не меняется

$$(X_{(i_1)}, X_{(i_2)}, \dots, X_{(i_m)}) = (X_{(1)}, X_{(2)}, \dots, X_{(m)}), \quad (252)$$

где $(i_1, i_2, \dots, i_m)$ - подстановка $(1, 2, \dots, m)$. Во-вторых, по каждому аргументу имеет место линейное свойство, например, по первому аргументу:

$$(aX + bX_{(1)}, X_{(2)}, \dots, X_{(m)}) = a(X, X_{(2)}, \dots) + b(X_{(1)}, X_{(2)}, \dots), \quad (253)$$

где a, b - произвольные действительные числа.

Если не нужно акцентировать внимание на числе аргументов m -линейной симметрической формы, то кратко такую форму будем называть полилинейной, подразумевая при этом, что она к тому же обязательно и симметрическая.

Если выбран некоторый базис $e_1, e_2, \dots, e_n$, то в этом базисе полилинейная форма принимает вид

$$(X_{(1)}, X_{(2)}, \dots, X_{(m)}) = \omega_{i_1 i_2 \dots i_m} x_{(1)}^{i_1} x_{(2)}^{i_2} \dots x_{(m)}^{i_m}, \quad (254)$$

где суммирование по дважды встречающимся индексам i_k в правой части идет от 1 до n , а

$$\omega_{i_1 i_2 \dots i_m} = (e_{i_1}, e_{i_2}, \dots, e_{i_m}). \quad (255)$$

Из определения m -линейной симметрической формы в линейном пространстве A_n следует, что её значение не зависит от выбора базиса, то есть её значение является инвариантом, поэтому при переходе от одного базиса к другому компоненты $\omega_{i_1 i_2 \dots i_m}$ формы преобразуются как компоненты m раз ковариантного тензора, так как компоненты векторов преобразуются как компоненты один раз контравариантных тензоров.

Из условия симметричности и формулы (255) следует, что при любой перестановке индексов у компонент $\omega_{i_1 i_2 \dots i_m}$ симметрической формы их значение не меняется, например:

$$\omega_{2,1,3,4,\dots,m} = \omega_{1,2,3,4,\dots,m}, \quad \omega_{1,2,\dots,m-1,m} = \omega_{m,m-1,\dots,2,1}. \quad (256)$$

Если число аргументов полилинейной формы равно двум, форма называется билинейной.

Если число аргументов m -линейной формы равно размерности пространства ($m = n$), будем называть такую полилинейную форму *полиформой*.

Рассмотрим m -линейную форму от m одинаковых аргументов, то есть $(X, X, \dots, X)$. Формально будем обозначать такую форму $|X|^m$,

$$|X|^m = \omega_{i_1 i_2 \dots i_m} x^{i_1} x^{i_2} \dots x^{i_m}. \quad (257)$$

Таким образом, $|X|^m$ - это однородная форма (полином) m -го порядка от n действительных аргументов $x^1, x^2, \dots, x^n$, координат вектора X . Если существует базис, в котором форма $|X|^m$ имеет такой вид, что при перестановке двух любых координат $x^i \leftrightarrow x^j$ вектора X , значение формы не меняется, то форму $|X|^m$, как и исходную m -линейную симметрическую форму $(X_{(1)}, X_{(2)}, \dots, X_{(m)})$ в этом базисе, будем называть *сверхсимметрической*, или сверхсимметрической записью исходной формы в специальном базисе.

В качестве примера рассмотрим в аффинном пространстве $A_4 \ni X(x^0, x^1, x^2, x^3)$ билинейную форму вида

$$(X, Y) = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3. \quad (258)$$

Очевидно, что она симметрическая

$$(Y, X) = (X, Y) \quad (259)$$

и линейная по каждому аргументу, так для первого аргумента справедлива следующая формула:

$$(aX_{(1)} + bX_{(2)}, Y) = a(X_{(1)}, Y) + b(X_{(2)}, Y). \quad (260)$$

Форма (258) не является сверхсимметрической, так как

$$|X|^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2, \quad (261)$$

и, например, при $x^0 \leftrightarrow x^1$ величина $|X|^2$, вообще говоря, меняет своё значение.

Выберем в качестве новых базисных векторов векторы:

$$\left. \begin{aligned} e_0 &= \left(\frac{3+\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2}, \frac{1+\sqrt{3}}{2} \right), \\ e_1 &= \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}, \\ e_2 &= \begin{pmatrix} 1 & 0 & 1 & 0 \end{pmatrix}, \\ e_3 &= \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \right\} \quad (262)$$

Так как

$$\begin{vmatrix} \frac{3+\sqrt{3}}{2} & \frac{1+\sqrt{3}}{2} & \frac{1+\sqrt{3}}{2} & \frac{1+\sqrt{3}}{2} \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{vmatrix} = \mp \sqrt{3}, \quad (263)$$

эти векторы линейно независимы, и их можно выбрать в качестве базисных. Векторы e_i обладают следующими свойствами:

$$(e_0, e_0) = 0, \quad (e_1, e_1) = 0, \quad (e_2, e_2) = 0, \quad (e_3, e_3) = 0, \quad (264)$$

$$(e_i, e_j) = 1, \quad \text{если } i \neq j. \quad (265)$$

Так как выполняются свойства (264), то векторы e_1, e_2, e_3, e_4 по определению являются изотропными, и базис, состоящий из таких векторов, естественно назвать изотропным.

Используя эти свойства базисных векторов и оставив для координат прежние обозначения, запишем форму (258) в новых координатах:

$$(X, Y) = x^0 y^1 + x^0 y^2 + x^0 y^3 + x^1 y^2 + x^1 y^3 + x^2 y^3 + \\ + x^1 y^0 + x^2 y^0 + x^3 y^0 + x^2 y^1 + x^3 y^1 + x^3 y^2. \quad (266)$$

Форма (261) тогда принимает вид:

$$(X, X) = 2x^0 x^1 + 2x^0 x^2 + 2x^0 x^3 + 2x^1 x^2 + 2x^1 x^3 + 2x^2 x^3. \quad (267)$$

При любой перестановке $x^i \leftrightarrow x^j$ форма (267) не меняется по форме и по значению, поэтому данная форма и форма (266) по определению являются сверхсимметрическими, или можно сказать, что они являются сверхсимметрическими записями форм (258) и (261).

Если от системы координат, в которых билинейная форма (258) принимает сверхсимметрический вид (266), перейти к новой системе координат с помощью преобразований группы Лоренца, то опять получим систему координат в которых билинейная форма (258) имеет тот же сверхсимметрический вид (266). Новые базисные векторы $e_{0'}, e_{1'}, e_{2'}, e_{3'}$ будут обладать теми же свойствами (264) и (265), то есть, в частности, будут изотропными. Таким образом, существует шести параметрическое семейство базисов со свойствами (264) и (265).

Всякая ли m -линейная симметрическая форма допускает сверхсимметрическую запись? Автор не знает ответа на этот вопрос. Если да, то хорошо бы иметь алгоритм построения такого базиса, в котором это явно просматривается. Если не всякая, то необходимо иметь критерий, с помощью которого это было бы можно выяснить.

Полилинейные формы вида (266) и аналогичные для $m > 2$ представляют собой интересный объект для изучения, особенно интересны соответствующие полиформы ($m = n$), последние возникают при изучение гиперкомплексных чисел H_n . Пусть $e_1, e_2, \dots, e_n$ - базис, а $x^1, x^2, \dots, x^n$ - координаты, выпишем несколько таких полиформ:

$m = n = 2$,

$$(X, Y) = \frac{1}{2}(x^1 y^2 + x^2 y^1), \quad (268)$$

соответственно

$$(X, X) = x^1 x^2; \quad (269)$$

$m = n = 3$,

$$(X, Y, Z) = \frac{1}{6} \times \quad (270)$$

$$\times (x^1 y^2 z^3 + x^2 y^3 z^1 + x^3 y^1 z^2 + x^2 y^1 z^3 + x^3 y^2 z^1 + x^1 y^3 z^2),$$

соответственно

$$(X, X, X) = x^1 x^2 x^3; \quad (271)$$

$\mathbf{m} = \mathbf{n} = 4$,

$$(X, Y, Z, U) = \frac{1}{24} \times$$

$$\begin{aligned} & \times (x^1 y^2 z^3 u^4 + x^2 y^3 z^4 u^1 + x^3 y^4 z^1 u^2 + x^4 y^1 z^2 u^3 + \\ & + x^2 y^1 z^3 u^4 + x^3 y^2 z^4 u^1 + x^4 y^3 z^1 u^2 + x^1 y^4 z^2 u^3 + \\ & + x^3 y^2 z^1 u^4 + x^4 y^3 z^2 u^1 + x^1 y^4 z^3 u^2 + x^2 y^1 z^4 u^3 + \\ & + x^4 y^2 z^3 u^1 + x^1 y^3 z^4 u^2 + x^2 y^4 z^1 u^3 + x^3 y^1 z^2 u^4 + \\ & + x^1 y^3 z^2 u^4 + x^2 y^4 z^3 u^1 + x^3 y^1 z^4 u^2 + x^4 y^2 z^1 u^3 + \\ & + x^1 y^2 z^4 u^3 + x^2 y^3 z^1 u^4 + x^3 y^4 z^2 u^1 + x^4 y^1 z^3 u^2), \end{aligned} \quad (272)$$

соответственно

$$(X, X, X, X) = x^1 x^2 x^3 x^4; \quad (273)$$

$\mathbf{m} = \mathbf{n}$,

$$(X_{(1)}, X_{(2)}, \dots, X_{(n)}) = \frac{1}{n!} \sum_{(1,2,\dots,n)} x_{(1)}^{i_1} x_{(2)}^{i_2} \dots x_{(n)}^{i_n}, \quad (274)$$

где суммирование идёт по наборам индексов $(i_1, i_2, \dots, i_n)$, которые являются всевозможными подстановками набора n натуральных чисел $(1, 2, \dots, n)$, соответственно

$$|X|^n = x^1 x^2 \dots x^n. \quad (275)$$

Сверхсимметрические полиформы (274) при $n \geq 2$ будем называть h_n -формами, а компоненты этих форм в специальном базисе, в котором они имеют вид (274), будем обозначать $\chi_{i_1 i_2 \dots i_n}$. Из формулы (274) следует, что

$$\chi_{i_1 i_2 \dots i_n} = \begin{cases} \frac{1}{n!}, & \text{если все индексы разные,} \\ 0, & \text{во всех остальных случаях.} \end{cases} \quad (276)$$

Такой специальный базис будем называть ψ -базисом, или *изотропным базисом* и обозначать $\psi_1, \psi_2, \dots, \psi_n$, или

$e_1, e_2, \dots, e_n$. Полиформы (274) связаны с поличислами H_n , именно по формуле (275) n -я степень нормы числа $X \in H_n$, если у этого числа норма определена, выражается через координаты в изотропном базисе.

Если существует такой базис, в котором форма $|X|^m$ содержит не все координаты вектора X , то такую форму и соответствующую ей форму $(X, Y, \dots, Z)$ будем называть *вырожденными*. В дальнейшем нас будут интересовать только *невырожденные полиформы*.

Среди всех полилинейных форм выделяются формы, которые обладают особым свойством, которое мы будем называть *свойством разрешимости*.

Будем говорить, что полилинейная форма (254) обладает *свойством разрешимости*, или *разрешима*, если существует такой тензор $\omega^{j_1 j_2 \dots j_m}$, что

$$\begin{aligned} \omega^{j_1 j_2 \dots j_m} \omega_{j_1 i_2 \dots i_m} x^{i_2} \dots x^{i_m} \omega_{j_2 k_2 \dots k_m} x^{k_2} \dots x^{k_m} \dots \omega_{j_m l_2 \dots l_m} x^{l_2} \dots x^{l_m} = \\ = \alpha (\omega_{i_1 i_2 \dots i_m} x^{i_1} x^{i_2} \dots x^{i_m})^{m-1}, \end{aligned} \quad (277)$$

где α - некоторое действительное число, причём в правой части вначале производится суммирование по индексам $i_1, i_2, \dots, i_m$, внутри круглых скобок, а затем полученная сумма возводится в $(m-1)$ -ую степень.

Естественно, что при этом компоненты матрицы $(\omega^{j_1 j_2 \dots j_m})$ выражаются неким образом через компоненты матрицы $(\omega_{i_1 i_2 \dots i_m})$ разрешимой формы. Если существует такой базис, в котором

$$(\omega^{j_1 j_2 \dots j_m}) = \beta \cdot (\omega_{i_1 i_2 \dots i_m}), \quad (278)$$

где β - некоторое действительное число, то будем называть такой базис *собственным базисом полилинейной формы*. О форме же будем говорить, что она имеет собственный базис.

Любая невырожденная билинейная форма является разрешимой. Полиформы (274), связанные с гиперкомплексными числами H_n , разрешимы, а базис, в котором эти формы имеют вид (274), есть собственный базис этих форм.

Для билинейной формы, если $\det(\omega_{ij}) \neq 0$, всегда можно построить дважды контравариантный тензор ω^{ij} , компоненты которого образуют обратную матрицу к матрице (ω_{ij}) , поэтому выполняется свойство разрешимости (277):

$$\omega^{j_1 j_2} \cdot \omega_{j_1 i_1} x^{i_1} \cdot \omega_{j_2 i_2} x^{i_2} = \omega_{i_1 i_2} x^{i_1} x^{i_2}. \quad (279)$$

Таким образом, невырожденные билинейные формы всегда разрешимы. Более того, для них всегда существует собственный базис, составленный из собственных векторов матрицы (ω_{ij}) , делённых на $\sqrt{|\lambda_i|}$, где λ_i - собственное значение соответствующего собственного вектора. В таком базисе матрицы (ω_{ij}) , (ω^{ij}) совпадают

$$(\omega^{ij}) = (\omega_{ij}) = \text{diag}(a_1, a_2, \dots, a_n), \quad (280)$$

где $a_i = \pm 1$.

10 Антиопределители

Если $n > 3$, вычислить значение h_n -формы даже в ψ -базисе (274) не просто. Для того чтобы это понять, достаточно взглянуть на формулу (272), по которой вычисляется h_4 -форма. В данном параграфе мы опишем удобный алгоритм вычисления h_n -форм.

Имеет место некоторое формальное сходство формулы (274) с формулой вычисления определителя n -го порядка. Отличие состоит в присутствии в формуле (274) общего множителя $\frac{1}{n!}$ и отсутствии знаковых множителей перед аддитивными членами под знаком суммы.

В силу этого представим h_n -форму в виде множителя $\frac{1}{n!}$ и некой величины, которую будем

называть *антиопределителем*, или *перманентом* n -го порядка, то есть по определению:

$$(X_{(1)}, X_{(2)}, \dots, X_{(n)}) = \frac{1}{n!} \left| \begin{array}{cccc} x_{(1)}^1 & x_{(2)}^1 & \dots & x_{(n)}^1 \\ x_{(1)}^2 & x_{(2)}^2 & \dots & x_{(n)}^2 \\ \dots & \dots & \dots & \dots \\ x_{(1)}^n & x_{(2)}^n & \dots & x_{(n)}^n \end{array} \right|_+, \quad (281)$$

$$adet(x_{(j)}^i) \equiv \left| \begin{array}{cccc} x_{(1)}^1 & x_{(2)}^1 & \dots & x_{(n)}^1 \\ x_{(1)}^2 & x_{(2)}^2 & \dots & x_{(n)}^2 \\ \dots & \dots & \dots & \dots \\ x_{(1)}^n & x_{(2)}^n & \dots & x_{(n)}^n \end{array} \right|_+ \equiv \sum_{(1,2,\dots,n)} x_{(1)}^{i_1} x_{(2)}^{i_2} \dots x_{(n)}^{i_n}, \quad (282)$$

где суммирование идёт по всевозможным подстановкам $(1, 2, \dots, n)$.

Алгоритм вычисления антиопределителей и их свойства сформулируем аналогично алгоритму вычисления определителей. Словами определение (3.2.2) можно записать следующим образом.

1). Антиопределитель квадратной $n \times n$ матрицы $\hat{X} = (x_{ij})$ есть сумма $n!$ произведений элементов матрицы $\hat{X}$, причём в каждом таком произведении присутствует один и только один элемент из каждой строки и каждого столбца.

Из определения непосредственно следуют ряд свойств антиопределителей.

2). Антиопределитель от транспонированной матрица равен антиопределителю от исходной матрицы,

$$adet(\hat{X}^T) = adet(\hat{X}). \quad (283)$$

Это же свойство можно переформулировать по-другому.

3). Если какое-то утверждение или свойство справедливо для столбцов антиопределителя, то оно справедливо и для его строк, и наоборот.

Это позволяет формулировать многие утверждения или свойства только в одном экземпляре - только для столбцов, что мы и сделаем.

4). Если у матрицы имеется столбец состоящий из одних нулей, то антиопределитель этой матрицы равен нулю.

5). Если какой-нибудь столбец матрицы имеет общий множитель, то его можно вынести за знак антиопределителя.

6). Если какой-нибудь столбец с номером j матрицы $\hat{X}$ можно представить как сумму двух столбцов: столбца a и столбца b - то антиопределитель матрицы есть сумма двух антиопределителей:

$$adet(\hat{X}) = adet(\hat{X}_a) + adet(\hat{X}_b), \quad (284)$$

где матрица $\hat{X}_a$ получается из матрицы $\hat{X}$ заменой j -го столбца на столбец a , а матрица $\hat{X}_b$ получается из матрицы $\hat{X}$ заменой j -го столбца на столбец b .

Из последних двух свойств следует более общее утверждение.

7). Пусть столбец с номером j матрицы $\hat{X}$ есть линейная комбинация любого количества столбцов:

$$x_i = \alpha \cdot a_i + \beta \cdot b_i + \dots + \gamma \cdot c_i + \dots, \quad (285)$$

где $\alpha, \beta, \dots, \gamma, \dots$ - числа. Тогда

$$\text{adet}(\hat{X}) = \alpha \cdot \text{adet}(\hat{X}_a) + \beta \cdot \text{adet}(\hat{X}_b) + \dots + \gamma \cdot \text{adet}(\hat{X}_c) + \dots \quad (286)$$

Матрицы $\hat{X}_a, \hat{X}_b, \dots, \hat{X}_c, \dots$ получаются из матрицы $\hat{X}$ заменой j -го столбца на столбцы $a, b, \dots, c, \dots$ соответственно.

Обратимся опять к определению антиопределителя:

$$\text{adet}(\hat{X}) \equiv \sum_{(1,2,\dots,n)} x_{i_1 1} x_{i_2 2} \dots x_{i_n n} . \quad (287)$$

Выделим в сумме, стоящей справа, все слагаемые, содержащие некоторый элемент x_{ij} матрицы $\hat{X} = x_{ij}$, и вынесем его за скобки. В скобках останется сумма из $(n-1)!$ слагаемых, каждое из которых будет содержать по одному элементу из каждой строки, кроме строки i , и каждого столбца, кроме столбца j , матрицы $\hat{X}$, то есть это будет антиопределитель, обозначим его X_{ij} , матрицы $(n-1) \times (n-1)$, которая получается из матрицы $\hat{X}$ вычёркиванием i -ой строки и j -го столбца. Будем называть X_{ij} алгебраическим дополнением элемента x_{ij} . Проведем описанную выше процедуру для каждого элемента некоторого выделенного столбца и получим алгоритм вычисления антиопределителя n -го порядка через антиопределители порядка $(n-1)$.

8). Формула разложения антиопределителя по фиксированному столбцу:

$$\text{adet}(\hat{X}) = x_{1j} X_{1j-} + x_{2j} X_{2j-} + \dots + x_{nj} X_{nj-} . \quad (288)$$

9). С помощью процедуры разложения антиопределителя по столбцу вычисление любого антиопределителя n -го порядка всегда можно свести к вычислению антиопределителей второго или третьего порядка, которые вычисляются аналогично тому, как вычисляются обыкновенные определители, но без знаков минус:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}_+ = ad + cb, \quad (289)$$

$$\begin{vmatrix} a & \alpha & d \\ \beta & b & \gamma \\ e & \delta & c \end{vmatrix}_+ = abc + \alpha\gamma e + \beta\delta d + ebd + \beta\alpha c + \delta\gamma a, \quad (290)$$

где $a, b, c, d, e, \alpha, \beta, \gamma, \delta$ - действительные числа.

10). Антиопределитель от диагональной матрицы равен произведению её элементов на главной диагонали.

11). Антиопределитель от матрицы, у которой отличные от нуля элементы стоят на побочной диагонали, равен произведению этих элементов.

12). Антиопределитель от треугольной матрицы, то есть матрицы, у которой выше или ниже главной диагонали (или побочной диагонали) все элементы равны нулю, равен произведению элементов, стоящих на главной диагонали (или побочной диагонали).

13). Если у антиопределителя поменять местами два столбца, то антиопределитель не изменит своё значение.

Из этого свойства следует более общее правило.

14). При любой перестановке столбцов антиопределителя его значение не меняется.

15). Антиопределитель от матрицы $n \times n$, у которой все элементы равны единице, равен $n!$.

16). Антиопределитель от матрицы $n \times n$, у которой все столбцы одинаковые, равен произведению $n!$ на произведение всех элементов одного столбца.

17). Антиопределитель от матрицы $\hat{X}$ ($n \times n$), которая состоит из $(n - 1)$ одинаковых столбцов с элементами $x_1, x_2, \dots, x_n$ и одного, вообще говоря, другого столбца с элементами $x'_1, x'_2, \dots, x'_n$, вычисляется по формуле

$$adet(\hat{X}) = (n - 1)! \cdot (x'_1 x_2 \dots x_n + x_1 x'_2 \dots x_n + \dots + x_1 \dots x_{n-1} x'_n), \quad (291)$$

или, если все $x_i \neq 0$, эту формулу можно переписать следующим образом:

$$adet(\hat{X}) = (n - 1)! \cdot x_1 x_2 \dots x_n \left(\frac{x'_1}{x_1} + \frac{x'_2}{x_2} + \dots + \frac{x'_n}{x_n} \right). \quad (292)$$

18). Если n - нечётно, то антиопределитель от антисимметрической матрицы равен нулю.

19). Если матрица $\hat{X}$ является блочной матрицей, у которой на главной или побочной диагонали стоят матрицы $\hat{A}, \hat{B}, \dots, \hat{C}$, а выше или ниже стоят нули, то антиопределитель матрицы $\hat{X}$ равен произведению антиопределителей матриц $\hat{A}, \hat{B}, \dots, \hat{C}$:

$$adet(\hat{X}) = adet(\hat{A}) \cdot adet(\hat{B}) \cdot \dots \cdot adet(\hat{C}). \quad (293)$$

11 Линейные пространства со скалярным полипроизведением

Рассмотрим линейное пространство $A \ni X, Y, Z, \dots$ над полем действительных чисел $R \ni x, y, z, a, b, c, \dots$, в котором определена m -линейная симметрическая форма, то есть для любых m векторов определено скалярное полипроизведение $(X_1, X_2, \dots, X_m)$. Термин "скалярное полипроизведение" подчёркивает, что число "сомножителей" (аргументов) в таком математическом объекте, вообще говоря, больше двух, но, конечно, обычное скалярное произведение с двумя аргументами является частным случаем скалярного полипроизведения.

Будем говорить, что вектор Y трансверсален ("ортогонален") вектору X , когда

$$(X, X, \dots, X, Y) = 0. \quad (294)$$

Если не только вектор Y трансверсален вектору X , но X трансверсален вектору Y , то назовём такие векторы взаимно трансверсальными (взаимно "ортогональными").

Пусть имеется базис $e_1, e_2, \dots, e_n$, состоящий из попарно взаимно трансверсальных векторов, причём

$$|e_i|^m \neq 0. \quad (295)$$

Будем называть такие базисы трансверсальными ("ортогональными"). Коэффициенты разложения произвольного вектора

$$X = x^1 e_1 + x^2 e_2 + \dots + x^n e_n \quad (296)$$

в этом базисе вычисляются по формулам

$$x^1 = \frac{(e_1, \dots, e_1, X)}{|e_1|^m}, \dots, x^n = \frac{(e_n, \dots, e_n, X)}{|e_n|^m}. \quad (297)$$

Если к тому же для всех базисных векторов

$$|e_i|^m = 1, \quad (298)$$

то будем называть такой базис "ортонормированным". Ниоткуда не следует, что такой базис, хотя бы один, существует в конкретном линейном пространстве со скалярным полипроизведением. Предположим, что в данном полилинейном пространстве такой базис $e_1, e_2, \dots, e_n$ существует. Тогда коэффициенты разложения вектора X по базисным векторам e_i равны скалярным произведениям $(m-1)$ -го вектора e_i и вектора X :

$$x^1 = (e_1, \dots, e_1, X), \dots, x^n = (e_n, \dots, e_n, X). \quad (299)$$

Эти формулы являются аналогами (обобщениями) формул для вычисления коэффициентов разложения вектора в ортонормированном базисе евклидова пространства. Само разложение запишется при этом в нашем случае следующим образом:

$$X = (e_i, \dots, e_i, X) \cdot e_i.$$

Для того, чтобы показать возможность существования подобных базисов в неквадратичных пространствах, рассмотрим в качестве примера линейное пространство с h_4 -формой. Перейдём от базиса $\psi_1, \psi_2, \psi_3, \psi_4$, в котором скалярное произведение четырёх векторов X, Y, Z, U имеет вид (272) к базису $1, j, k, jk$ по формулам:

$$\left. \begin{aligned} 1 &= \psi_1 + \psi_2 + \psi_3 + \psi_4, \\ j &= \psi_1 + \psi_2 - \psi_3 - \psi_4, \\ k &= \psi_1 - \psi_2 + \psi_3 - \psi_4, \\ jk &= \psi_1 - \psi_2 - \psi_3 + \psi_4. \end{aligned} \right\} \quad (300)$$

Для того чтобы векторы $1, j, k, jk$ составляли базис, они должны быть линейно независимы. Это выполняется, так как

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{vmatrix} = -16 \neq 0. \quad (301)$$

Этот базис является "ортонормированным". Чтобы это проверить надо воспользоваться формулами (281), (282), (292) вычисления скалярных произведений в пространстве с h_4 -формой через антиопределители. Всего при этом нам придётся вычислить 16 антиопределителей, например:

$$(e_2, e_2, e_2, e_2) = \frac{1}{24} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{vmatrix}_+ = 1, \quad (302)$$

$$(e_2, e_2, e_2, e_4) = \frac{1}{24} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \end{vmatrix}_+ = 0. \quad (303)$$

Результат: в пространстве поличисел H_4 базис $1, i, j, jk$ (300) является "ортонормированным" и в нём для каждого числа X имеет место разложение

$$X = (1, 1, 1, X) \cdot 1 + (i, i, i, X) \cdot i + (j, j, j, X) \cdot j + (jk, jk, jk, X) \cdot jk.$$

Если невырожденная полилинейная форма на некотором множестве векторов принимает положительные значения, то соответствующее полилинейное пространство является метрическим, а именно - финслеровым.

Рассмотрим m -линейную форму, определённую в данном полилинейном пространстве, от m одинаковых аргументов:

$$|X|^m \equiv (X, X, \dots, X). \quad (304)$$

Всё множество векторов представим себе как объединение четырёх непересекающихся подмножеств:

- 1) 0 - нулевой вектор;
- 2) множество изотропных векторов, это такие векторы $X \neq 0$, что

$$|X|^m = 0; \quad (305)$$

- 3) множество измеримых векторов, то есть таких векторов X , что

$$|X|^m > 0; \quad (306)$$

- 4) множество неизмеримых векторов, для которых

$$|X|^m < 0. \quad (307)$$

Естественно считать длины нулевого и изотропных векторов равными нулю, а длину l измеримого вектора X определить формулой

$$l = \sqrt[m]{(X, X, \dots, X)}. \quad (308)$$

Будем считать наше линейное пространство центроаффинным, то есть существует некая точка O (центр), и начала всех радиус-векторов совпадают с этой точкой. Тогда мы получаем координатное пространство, у которого центроаффинные касательные пространства $\mathcal{A}_n(M)$ в каждой точке M основного пространства изоморфно исходному центроаффинному пространству. Элемент длины в основном пространстве определим формулой:

$$dl = \sqrt[m]{(dX, dX, \dots, dX)} \quad (309)$$

или

$$dl = \sqrt[m]{\omega_{i_1 i_2 \dots i_m} dx^{i_1} dx^{i_2} \dots dx^{i_m}}. \quad (310)$$

Таким образом, центроаффинное полилинейное пространство с метрической функцией

$$L(dx) = \sqrt[m]{\omega_{i_1 i_2 \dots i_m} dx^{i_1} dx^{i_2} \dots dx^{i_m}} \quad (311)$$

является финслеровым пространством, если для m -линейной формы существуют измеримые векторы и если

$$\text{rang} \left(\frac{\partial^2 L(dx; x)}{\partial (dx^i) \partial (dx^j)} \right) = n - 1. \quad (312)$$

Пространства со скалярным полипроизведением, у которых форма $|X|^m$ вырожденная, то есть в каком-то базисе содержит не все n координат векторной переменной X , не удовлетворяют условию (312).

То как была определена метрическая функция в линейном пространстве со скалярным полипроизведением, формула (311), и соответственно длина вектора (308), дает нам "максимальную" геометрию в том смысле, что все измеримые векторы (306) пространства остаются измеримыми векторами в определённой выше финслеровой геометрии. В общем случае можно построить и другие финслеровы геометрии на основе фиксированного линейного пространства с фиксированной m -линейной формой. Поясним это на примерах.

Рассмотрим пространство с билинейной формой (258). Измеримые векторы (x^0, x^1, x^2, x^3) , в смысле (261), в этом случае это векторы, удовлетворяющие условию

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 > 0. \quad (313)$$

Если определена максимальная финслерова геометрия, то метрическая функция определена на всех измеримых, в смысле (261), векторах, то есть каждый измеримый вектор порождает луч, пересекающий индикатрису, уравнение которой есть

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = 1. \quad (314)$$

Если вектор X_{ind} - это радиус-вектор с координатами какой-то произвольной точки индикатрисы, то множество векторов вида αX_{ind} , где $\alpha > 0$ - произвольное положительное действительное число, совпадает с множеством измеримых векторов, в смысле (261). Индикатриса (269) - поверхность второго порядка, двуполостной гиперболоид - состоит из двух несвязанных поверхностей, одна из которых K_+ находится в конусе будущего (её можно выделить неравенством $x^0 > 0$), а другая K_- - в конусе прошлого (её можно выделить неравенством $x^0 < 0$). Если мы не хотим ограничивать непрерывную группу изометрической симметрии, свойственную билинейной форме, то на основе билинейной формы (258) можно определить три финслеровы геометрии:

1) максимальная финслерова геометрия, область определения метрической функции - $K_+ \cup K_-$;

2) финслерова геометрия с областью определения K_+ , координата x^0 всех измеримых векторов при этом удовлетворят условию $x^0 > 0$;

3) финслерова геометрия с областью определения K_- , координата x^0 всех измеримых векторов при этом удовлетворят условию $x^0 < 0$.

В математическом плане финслеровы геометрии 2) и 3) не отличаются друг от друга, так как при преобразовании координат $x^0 \rightarrow -x^0$, переходят друг в друга, а финслерова геометрия 1) качественно отличается от геометрии 2) и геометрии 3).

Рассмотрим пространство с h_4 -формой (272). Измеримые радиус-векторы (x^1, x^2, x^3, x^4) в смысле (306) в данном случае удовлетворяют условию:

$$x^1 x^2 x^3 x^4 > 0. \quad (315)$$

Если в координатном пространстве (x^1, x^2, x^3, x^4) вводится максимальная геометрия, уравнение индикатрисы в касательном центроаффинном пространстве в каждой точке $(x_0^1, x_0^2, x_0^3, x_0^4)$ основного пространства будет иметь вид

$$(x^1 - x_0^1)(x^2 - x_0^2)(x^3 - x_0^3)(x^4 - x_0^4) = 1. \quad (316)$$

Индикатриса (316) - поверхность четвёртого порядка, восьмиполостный гиперboloид - состоит из восьми несвязанных поверхностей, одна из которых K_+ находится в конусе будущего (её можно выделить неравенствами $x^i - x_0^i > 0$), а другая K_- - в конусе прошлого (её можно выделить неравенством $x^i - x_0^i < 0$).

Если в линейном пространстве со скалярным полипроизведением размерности n возможно определить несколько качественно различных финслеровых геометрий, то обязательно будем выделять две из них: максимальную геометрию, которую будем обозначать $\mathfrak{F}_n^{\text{max}}$, и финслерову геометрию, у которой все измеримые векторы лежат в конусе будущего и которую будем обозначать $\mathfrak{F}_n^+$.

Если m -форма обладает свойством разрешимости (277), то для финслеровой геометрии с метрической функцией (311) можно легко найти тангенциальное уравнение индикатрисы. Так по определению

$$p_i = \frac{\omega_{i i_2 \dots i_m} dx^{i_2} \dots dx^{i_m}}{(\omega_{j_1 j_2 \dots j_m} dx^{j_1} dx^{j_2} \dots dx^{j_m})^{\frac{m-1}{m}}}. \quad (317)$$

Используя формулу (277), получим тангенциальное уравнение индикатрисы:

$$\omega^{j_1 j_2 \dots j_m} p_{j_1} p_{j_2} \dots p_{j_m} - \alpha = 0. \quad (318)$$

12 Норма и группа симметрии поличисел

Непосредственно из теоремы Вейерштрасса следует, что любая система невырожденных поличисел P_n изоморфна алгебре квадратных диагональных матриц $(k + m) \times (k + m)$, у которых первые k элементов - произвольные действительные числа, а остальные m элементов - комплексные числа, причем алгебра P_n рассматривается над полем действительных чисел, то есть такие матрицы складываются и перемножаются обычным образом, а умножаются только на действительные числа. Размерность таких гиперкомплексных чисел равна

$$k + 2m = n. \quad (319)$$

В силу этого неизоморфные системы невырожденных поличисел однозначно определяются (классифицируются) парой действительных чисел, например, k и m , размерность пространства определяется формулой (319). Чтобы не изобретать новое обозначение для невырожденных поличисел предлагается использовать старое обозначение для поличисел, но записывать размерность пространства в виде формулы P_{k+2m} , например, $P_{3+2 \cdot 4}$, где $k = 3$, $m = 4$, а размерность пространства $n = 11$.

Таким образом, для любой системы невырожденных поличисел существует базис $e_1, e_2, \dots, e_n$, в котором:

$$p_{ij}^l = \begin{cases} 1, & i = j = l = 1, 2, \dots, k, k+1, k+3, \dots, n-1, \\ 1, & j = l = i+1 = k+2, k+4, \dots, n, \\ 1, & i = l = j+1 = k+2, k+4, \dots, n, \\ -1, & i = j = l+1 = k+2, k+4, \dots, n, \\ 0, & \text{во всех остальных случаях.} \end{cases} \quad (320)$$

Такой базис будем называть изотропным. При алгебраических вычислениях (преобразованиях) удобнее пользоваться именно этим базисом.

В матричном представлении на множестве квадратных действительных матриц $n \times n$ изотропный базис можно реализовать из матриц трёх видов: диагональных матриц, у которых на главной диагонали имеется единственный отличный от нуля элемент, равный единице, а все остальные элементы равны нулю; а также блочных матриц двух видов, у которых на главной диагонали стоит одна из двух матриц

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

а все остальные элементы равны нулю.

Используя числовой тензор p_{ij}^l , можно построить много других числовых тензоров, например,

$$q_{ij} = p_{is}^l \cdot p_{lj}^s. \quad (321)$$

Для поличисел тензор q_{ij} является симметрическим, так как

$$q_{ij} = q_{ji}. \quad (322)$$

Построим этот тензор в изотропном базисе (320):

$$(q_{ij}) = \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 2 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & -2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & 2 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & -2 \end{pmatrix}, \quad (323)$$

то есть в этом базисе компоненты тензора q_{ij} образуют диагональную матрицу, у которой первые k элементов равны 1, а остальные $2m$ элементов заполняют главную диагональ парами $\{2, -2\}$, поэтому в изотропном базисе

$$\det(q_{ij}) = (-4)^m.$$

Отсюда следует, что в любой системе координат для невырожденных поличисел $P_{k+2\cdot m}$ выполняется условие

$$\det(q_{ij}) \neq 0, \quad (324)$$

а значит можно построить дважды контравариантный тензор q^{ij} такой, что

$$q^{il}q_{lj} = q_{jl}q^{li} = \delta_j^i. \quad (325)$$

Невырожденные поличисла $P_{r+2\cdot 0}$ занимают особое положение среди всех невырожденных поличисел, поэтому введём для них специальное обозначение

$$H_r \equiv P_{r+2\cdot 0}. \quad (326)$$

Алгебра поличисел H_r изоморфна алгебре действительных квадратных диагональных матриц $r \times r$ над полем действительных чисел. Удобно также выделить поличисла $P_{0+2\cdot m} \equiv H_m^C$. Можно представить, что элементами такой алгебры являются комплексные квадратные диагональные матрицы $m \times m$, их можно складывать, перемножать и умножать на действительные числа, то есть это алгебра над полем действительных чисел. Тогда справедлива формула

$$P_{k+2m} = H_k \oplus H_m^C. \quad (327)$$

Очевидно, что норма на множестве поличисел P_{k+2m} в изотропном базисе определяется формулой

$$|X| = \sqrt[n]{x^1 \dots x^k [(x^{k+1})^2 + (x^{k+2})^2] \dots [(x^{n-1})^2 + (x^n)^2]}, \quad (328)$$

а область, где норма определена, задается неравенствами

$$x^1 x^2 \dots x^k \geq 0. \quad (329)$$

Все три требования, сформулированные в начале данного раздела, которым должна удовлетворять норма поличисел, выполняются.

Область

$$x^1 > 0, x^2 > 0, \dots, x^k > 0,$$

будем называть конусом будущего. Если заменить неравенство (329) этим более сильным условием принадлежности поличисел конусу будущего, то тем более все три аксиомы в определении нормы поличисел выполняются.

В ковариантной форме норма поличисла (328) записывается следующим образом:

$$|X| = \sqrt[n]{g_{i_1 i_2 \dots i_n} x^{i_1} x^{i_2} \dots x^{i_n}}, \quad (330)$$

где $g_{i_1 i_2 \dots i_n}$ - метрический тензор.

Учитывая, что в изотропном базисе тензор p_{ij}^l определяется формулой (320), получим в этом базисе

$$|X|^n = \det(x^i p_{ij}^l), \quad (331)$$

но справа в этой формуле стоит величина, которая не изменяется при переходе от одного базиса к другому, поэтому формула (331) верна в любом базисе, то есть не зависит от выбора базиса. Формула (331) определяет норму числа X , если правая часть больше или равна нулю,

$\det(x^i p_{ij}^l) \geq 0$, второе и третье требования в определении нормы поличисла выполняются автоматически.

Группой симметрии $G_1(P_{k+2m})$ пространства P_{k+2m} будем называть непрерывную группу линейных преобразований координат векторов этого пространства в фиксированном базисе, которые (преобразования) не меняют норму любого числа, для которого норма существует. Индекс "1" имеет важный смысл, ниже будет показано, что группа симметрии изоморфна множеству унимодулярных поличисел с групповым умножением, совпадающим с поличисловым.

Непосредственно из формулы (328) следует, что группа $G_1(P_{k+2m})$ состоит из двух видов преобразований: эллиптических поворотов и гиперболических поворотов, под последними мы понимаем одновременное растяжение по одним координатам и обязательно сжатие по другим, при выполнении некоторого специального условия.

Эллиптические повороты осуществляются матрицами вида:

$$\hat{A} = \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & \cos \varphi_i & -\sin \varphi_i & 0 & \dots & 0 \\ 0 & \dots & 0 & \sin \varphi_i & \cos \varphi_i & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \quad (332)$$

Здесь матрица поворота

$$\begin{pmatrix} \cos \varphi_i & -\sin \varphi_i \\ \sin \varphi_i & \cos \varphi_i \end{pmatrix} \quad (333)$$

осуществляет правый эллиптический поворот в плоскости (x^i, x^{i+1}) , где $i = k+1, k+3, \dots, n-1$, на действительный угол φ_i . Всего таких коммутирующих между собой поворотов равно m .

Гиперболические повороты осуществляются матрицами вида

$$\hat{B} = \begin{pmatrix} e^{\varepsilon_1} & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & e^{\varepsilon_i} & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & e^{\varepsilon_{i+1}} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & e^{\varepsilon_{n-1}} & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & e^{\varepsilon_n} \end{pmatrix}. \quad (334)$$

где ε_j - действительные числа, причем

$$\left. \begin{aligned} \varepsilon_{k+1} = \varepsilon_{k+2}, \varepsilon_{k+3} = \varepsilon_{k+4}, \dots, \varepsilon_{n-1} = \varepsilon_n, \\ \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_k + \varepsilon_{k+1} + \varepsilon_{k+2} + \dots + \varepsilon_{n-1} + \varepsilon_n = 0. \end{aligned} \right\} \quad (335)$$

Число независимых гиперболических поворотов равно $(k+m-1)$, все они коммутируют между собой и с эллиптическими поворотами (332).

Таким образом, непрерывная группа симметрии $G_1(P_{k+2m})$ является абелевой непрерывной $(n-1)$ параметрической группой Ли, так как

$$m + (k + m - 1) = k + 2m - 1 = n - 1.$$

13 Длина отрезка кривой в пространстве P_{k+2m}

Зададим в координатном пространстве P_{k+2m} в изотропном базисе $e_1, e_2, \dots, e_n$ со структурным тензором (320) некоторую кривую

$$x^i = x^i(\tau), \quad \tau \in [\tau_1, \tau_2], \quad (336)$$

при $\tau = \tau_1$ кривая проходит через точку A , а при $\tau = \tau_2$ - через точку B , причем все бесконечно малые векторы с координатами

$$dx^i = \dot{x}^i(\tau) d\tau \quad (337)$$

измеримы, то есть для них определена норма. Тогда естественно определить длину отрезка кривой между двумя точками A и B как интеграл

$$l_{AB} = \int_A^B \sqrt[n]{dx^1 \dots dx^k [(dx^{k+1})^2 + (dx^{k+2})^2] \dots [(dx^{n-1})^2 + (dx^n)^2]}, \quad (338)$$

или

$$l_{AB} = \int_{\tau_1}^{\tau_2} \sqrt[n]{\dot{x}^1 \dots \dot{x}^k [(\dot{x}^{k+1})^2 + (\dot{x}^{k+2})^2] \dots [(\dot{x}^{n-1})^2 + (\dot{x}^n)^2]} \cdot d\tau. \quad (339)$$

Таким образом, если $n = k + 2 \cdot m > 2$, то метрическое пространство P_{k+2m} не является ни евклидовым, ни псевдоевклидовым. Это плоское финслерово пространство с метрической функцией

$$L(dx) = \sqrt[n]{dx^1 \dots dx^k [(dx^{k+1})^2 + (dx^{k+2})^2] \dots [(dx^{n-1})^2 + (dx^n)^2]}, \quad (340)$$

инвариантной относительно группы $G_1(P_{k+2m})$. Компоненты p_i обобщенного импульса вычисляются следующим образом:

$$p_i = \frac{\partial L(dx)}{\partial (dx^i)} \equiv \frac{\partial L(\dot{x})}{\partial (\dot{x}^i)}. \quad (341)$$

Эти величины связаны функциональным соотношением

$$\Phi(p_1, p_2, \dots, p_n) = 0, \quad \text{или} \quad \Phi(p) = 0, \quad (342)$$

которое принято называть тангенциальным уравнением индикатрисы. Для метрической функции (340) тангенциальное уравнение индикатрисы можно записать в виде:

$$p_1 p_2 \dots p_k (p_{k+1}^2 + p_{k+2}^2) (p_{k+3}^2 + p_{k+4}^2) \dots (p_{n-1}^2 + p_n^2) - \frac{4^m}{n^n} = 0. \quad (343)$$

Пусть в координатном пространстве $x^1, x^2, \dots, x^n$ заданы две финслеровы геометрии, причём между метрическими функциями $L(dx; x)$ и $L'(dx; x)$ этих геометрий имеет место соотношение

$$L'(dx; x) = \kappa(x) \cdot L(dx; x), \quad (344)$$

где $\kappa(x) > 0$ - некоторая функция, зависящая только от точки основного пространства. Тогда такие геометрии называются *конформно связанными* [2].

Финслерова геометрия конформно связанная с геометрией (340) в каждой точке имеет касательное пространство изоморфное поличисловому пространству $P_{k+2 \cdot m}$, а тангенциальное уравнение индикатрисы в таком пространстве имеет вид:

$$p_1 p_2 \dots p_k (p_{k+1}^2 + p_{k+2}^2) (p_{k+3}^2 + p_{k+4}^2) \dots (p_{n-1}^2 + p_n^2) - \frac{4^m}{n^n} \kappa^n(x) = 0. \quad (345)$$

Напомним, что если известно тангенциальное уравнение индикатрисы

$$\Phi(p; x) = 0, \quad (346)$$

то каноническая система дифференциальных уравнений для определения экстремалей (геодезических, мировых линий) в финслеровом пространстве записывается следующим образом:

$$\dot{x}^i = \frac{\partial \Phi}{\partial p_i} \cdot \lambda(p; x), \quad \dot{p}_i = -\frac{\partial \Phi}{\partial x^i} \cdot \lambda(p; x), \quad (347)$$

где $\lambda(p; x) > 0$ - произвольная скалярная функция.

Для финслерова пространства конформно связанного с пространством $P_{k+2 \cdot m}$ с тангенциальным уравнением индикатрисы (345) частные производные функции $\Phi(p; x)$ определяются формулами:

$$\frac{\partial \Phi}{\partial p_i} = \begin{cases} \frac{4^m}{n^n} \cdot \frac{\kappa^n}{p_i}, & i = 1, 2, \dots, k, \\ \frac{4^m}{n^n} \cdot \kappa^n \frac{2p_i}{p_i^2 + p_{i+1}^2}, & i = k+1, k+3, \dots, n-1, \\ \frac{4^m}{n^n} \cdot \kappa^n \frac{2p_i}{p_{i-1}^2 + p_i^2}, & i = k+2, k+4, \dots, n; \end{cases} \quad (348)$$

$$\frac{\partial \Phi}{\partial x^i} = -\frac{4^m}{n^{n-1}} \cdot \kappa^{n-1} \frac{\partial \kappa}{\partial x^i}. \quad (349)$$

Подставив эти производные в уравнения (347), получим систему дифференциальных уравнений для определения экстремалей в пространстве, конформно связанном с пространством невырожденных поличисел $P_{k+2 \cdot m}$.

14 Экспоненциальное представление невырожденных поличисел

Под экспоненциальной функцией $\exp(X)$ от поличисла $X \in P_{k+2 \cdot m}$ будем понимать ряд

$$\exp(X) = 1 + X + \frac{1}{2!} X^2 + \dots + \frac{1}{s!} X^s + \dots, \quad (350)$$

а под *экспоненциальным представлением* числа $X \in P_{k+2,m}$, если у этого числа имеется норма $|X|$, - однозначную запись этого числа в виде

$$X = |X| \cdot \exp(\varphi_2 E_2 + \varphi_3 E_3 + \dots + \varphi_n E_n), \quad (351)$$

где $\varphi_2, \varphi_3, \dots, \varphi_n$ - действительные числа, а $E_1 \equiv 1, E_2, E_3, \dots, E_n$ - некий специальный базис, в котором такая однозначная запись возможна. Величина $\varphi_2, \varphi_3, \dots, \varphi_n$ будем называть угловыми переменными (эллиптическими или гиперболическими).

Говорить о экспоненциальной записи нулевого элемента в $P_{k+2,m}$ или делителей нуля не имеет смысла, так как для них не существует однозначной записи (351).

Только в единственной алгебре невырожденных поличисел, алгебре комплексных чисел любое не равное нулю число $z = x + iy$ имеет однозначное экспоненциальное представление $z = e^{i\alpha}$. В общем случае для алгебры $P_{k+2,m}$ экспоненциальное представление возможно только при выполнении некоторых условий, из которых обязательно должно следовать условие

$$(X, X, \dots, X) > 0. \quad (352)$$

Докажем это утверждение, предположив, что мы уже построили такой базис, а затем явно его построим. Рассмотрим множество унимодулярных чисел, имеющих экспоненциальное представление, то есть чисел $A \in P_{k+2,m}$, у которых $|A| = 1$ и которые можно записать как

$$A = \exp(\alpha_2 E_2 + \alpha_3 E_3 + \dots + \alpha_n E_n). \quad (353)$$

Тогда множество преобразований пространства $P_{k+2,m}$ вида

$$X \rightarrow X', \quad X' = A \cdot X \quad (354)$$

образуют группу линейных преобразований пространства $P_{k+2,m}$, сохраняющих модули всех чисел, у которых модуль определен, так как

$$|X'| = |A| \cdot |X| \Rightarrow |X'| = |X|. \quad (355)$$

Кроме того, эта группа - абелева $(n-1)$ параметрическая непрерывная группа Ли. Выше мы построили преобразования этой группы в координатном пространстве и обозначили ее символом $G_1(P_{k+2,m})$. Таким образом, эта группа есть группа унимодулярных поличисел.

Все эти рассуждения дают нам алгоритм построения базисных элементов $E_2, E_3, \dots, E_n$: они должны быть генераторами данной абелевой группы Ли. Это позволяет нам построить базис $E_1 \equiv 1, E_2, E_3, \dots, E_n$ из изотропного базиса $e_1, e_2, \dots, e_n$ со структурным тензором (320), если $k \neq 0$, например, так:

$$\left. \begin{aligned} E_1 &= e_1 + e_2 + \dots + e_k + e_{k+1} + e_{k+3} + \dots + e_{n-1}, \\ E_2 &= e_1 - e_2, \\ E_3 &= e_1 - e_3, \\ &\vdots \\ E_k &= e_1 - e_k, \\ E_{k+1} &= 2e_1 - e_{k+1}, \\ E_{k+2} &= e_{k+2}, \\ E_{k+3} &= 2e_1 - e_{k+3}, \\ E_{k+4} &= e_{k+4}, \\ &\vdots \\ E_{n-1} &= 2e_1 - e_{n-1}, \\ E_{k+2} &= e_n. \end{aligned} \right\} \quad (356)$$

Если $k = 0$, то, например, так:

$$\left. \begin{aligned} E_1 &= e_1 + e_3 + \dots + e_{n-1}, \\ E_2 &= e_2, \\ E_3 &= e_1 - e_3, \\ E_4 &= e_4, \\ &\vdots \\ E_{n-1} &= e_1 - e_{n-1}, \\ E_n &= e_n. \end{aligned} \right\} \quad (357)$$

Используя числа $E_2, E_3, \dots, E_n$ в качестве генераторов группы, получим все преобразования (332), (334).

Базис $E_1 \equiv 1, E_2, E_3, \dots, E_n$ (356) или (357) не единственный, в котором имеет место экспоненциальное представление (351). Любой другой такой базис получается из базиса (356) или (357) произвольным линейным невырожденным преобразованием базисных элементов $E_2, E_3, \dots, E_n$.

Таким образом, при выборе базиса, который допускает экспоненциальное представление числа, имеется $(n - 1)^2$ параметрический произвол, который можно использовать, например, чтобы сделать базис "ортонормированным" (если это возможно) или чтобы добиться выполнения какого-либо другого необходимого свойства.

Выясним, в каком случае поличисло X , имеющее координаты x^i в изотропном базисе $e_1, e_2, \dots, e_n$,

$$X = x^1 e_1 + x^2 e_2 + \dots + x^n e_n,$$

при выполнении условия (352) может быть представлено в экспоненциальном виде (351). Для этого внесём $|X|$ под экспоненту, тогда под экспонентой окажется некоторое поличисло Y . Запишем поличисло Y в изотропном базисе, координаты этого числа в изотропном базисе обозначим y^i . Тогда, используя свойства изотропного базиса, получим

$$\begin{aligned} X &= \exp(Y) = \\ &= \exp(y_1) e_1 + \exp(y_2) e_2 + \dots + \exp(y_k) e_k + \\ &+ \exp(y_{k+1}) \cos(y_{k+2}) e_{k+1} + \exp(y_{k+1}) \sin(y_{k+2}) e_{k+2} + \dots \\ &+ \exp(y_{n-1}) \cos(y_n) e_{n-1} + \exp(y_{n-1}) \sin(y_n) e_n. \end{aligned} \quad (358)$$

Таким образом, для того чтобы поличисло $X \in P_{k+2,m}$, имело экспоненциальное представление, необходимо и достаточно, чтобы его координаты x^i в изотропном базисе удовлетворяли следующим условиям:

$$\left. \begin{aligned} x^1 &> 0, \quad x^2 > 0, \quad \dots, \quad x^k > 0, \\ (x^{k+1})^2 + (x^{k+2})^2 &\neq 0, \quad \dots, \quad (x^{n-1})^2 + (x^n)^2 \neq 0. \end{aligned} \right\} \quad (359)$$

Заметим, что при выполнении условий (359) поличисло $X \in P_{k+2,m}$ находится в конусе будущего, а условие (352) выполняется автоматически.

Итак, для того, чтобы невырожденное поличисло $X \in P_{k+2,m}$ имело экспоненциальное представление (351), оно должно лежать в конусе будущего.

15 Функции поличисловой переменной

Пусть пока $e_1, e_2, \dots, e_n$ произвольный базис в пространстве $P_{k+2.m}$. Функцией переменной $X = x^1 e_1 + \dots + x^n e_n \in P_{k+2.m}$ называется функция $F(X) \in P_{k+2.m}$, такая, что

$$F(X) = f^1(x^1, \dots, x^n) e_1 + f^2(x^1, \dots, x^n) e_2 + \dots + f^n(x^1, \dots, x^n) e_n, \quad (360)$$

где $f^i(x) \in R$ - n произвольных функций (в дальнейшем мы будем считать их гладкими) от n действительных аргументов. Рассмотрим дифференциал

$$dF(X) \equiv F(X + dX) - F(X),$$

где $dX \in P_{k+2.m}$ - произвольное бесконечно малое поличисло в том смысле, что

$$dX = \varepsilon Y, \quad Y \in P_{k+2.m}, \quad \varepsilon \rightarrow 0.$$

Тогда

$$dF(X) = \frac{\partial f^l}{\partial x^i} \cdot dx^i \cdot e_l. \quad (361)$$

Если дифференциал $dF(X)$ функции $F(X)$ можно представить в виде

$$dF(X) = F'(X) \cdot dX + o(dx^i), \quad (362)$$

где $F'(X)$ некоторая функция переменной $P_{k+2.m}$, то функцию $F(X)$ принято называть аналитической функцией поличисловой $P_{k+2.m}$ переменной, а $F'(X)$ - производной этой аналитической функции по той же переменной.

Из двух последних формул получим

$$\frac{\partial f^l}{\partial x^i} = p_{ji}^l f'^j. \quad (363)$$

Воспользуемся одним из соотношений (206), тогда из предыдущей формулы следует

$$f'^l = \epsilon^i \frac{\partial f^l}{\partial x^i}. \quad (364)$$

Подставив полученное выражение в (363), имеем n^2 соотношений

$$\frac{\partial f^l}{\partial x^i} = \epsilon^s p_{ji}^l \frac{\partial f^j}{\partial x^s}, \quad (365)$$

из которых независимых не более $n^2 - n$, так как свертка левой и правой части с ϵ^i (компоненты единичного элемента в данном базисе) приводит к n тождествам. Соотношения (365) для функций комплексной переменной принято называть соотношениями Коши - Римана. Сохраним этот термин для функций поличисловой переменной. Если эти соотношения выполняются, то, определив функцию $F'(X)$ с помощью формул (364), приходим к выполнению формулы (362).

Таким образом, мы пришли к следующему результату: для того, чтобы функция $F(X)$ поличисловой переменной была аналитической, необходимо и достаточно, чтобы выполнялись соотношения Коши - Римана (365).

Непосредственно из определения аналитической функции поличисловой переменной следует:

1) линейная комбинация аналитических функций поличисловой переменной с действительными коэффициентами α и β или поличисловыми переменными A и B есть аналитическая функция той же переменной, причем

$$[\alpha F_{(1)}(X) + \beta F_{(2)}(X)]' = \alpha F'_{(1)}(X) + \beta F'_{(2)}(X), \quad (366)$$

$$[AF_{(1)}(X) + BF_{(2)}(X)]' = AF'_{(1)}(X) + BF'_{(2)}(X);$$

2) поличисловое произведение двух аналитических функций поличисловой переменной, есть аналитическая функция той же переменной, причем

$$[F_{(1)}(X) \cdot F_{(2)}(X)]' = F'_{(1)}(X)F_{(2)}(X) + F_{(1)}(X)F'_{(2)}(X); \quad (367)$$

3) аналитическая функция поличисловой переменной от аналитической функции той же поличисловой переменной есть аналитическая функция все той же поличисловой переменной, причем

$$[F_{(1)}(F_{(2)}(X))] = F'_{(1)}(Y)|_{Y=F_{(2)}(X)} \cdot F'_{(2)}(X). \quad (368)$$

Общий вид аналитических функций переменной $P_{k+2 \cdot m}$ можно найти, работая в изотропном базисе $e_1, e_2, \dots, e_n$ со структурным тензором p_{ij}^l (320), но выполнить эту задачу еще проще, если вспомнить, что невырожденные поличисла $P_{k+2 \cdot m}$ изоморфны прямой сумме (327) алгебр H_k и H_m^C , то есть алгебре квадратных диагональных частично комплексных матриц над полем действительных чисел. В этой алгебре любой элемент $\hat{X}$ может быть представлен в виде

$$\hat{X} = x^1 \hat{\Psi}_1 + \dots + x^k \hat{\Psi}_2 + \dots + x^{k+1} \hat{\Psi}_{k+1} + \dots + x^{m+k} \hat{\Psi}_{k+m}, \quad (369)$$

где $\hat{\Psi}_i$ - действительная квадратная диагональная матрица, у которой единственным отличным от нуля элементом, является i -й, равный единице; x^i ($i = 1, 2, \dots, k$) - k действительных чисел, а x^j ($j = k+1, k+2, \dots, k+m$) - m комплексных чисел. Можно рассматривать матрицы $\hat{\Psi}_1, \hat{\Psi}_2, \dots, \hat{\Psi}_{k+m}$ как базис. Тогда

$$\hat{\Psi}_i \hat{\Psi}_j = \tilde{p}_{ij}^l \hat{\Psi}_l, \quad \tilde{p}_{ij}^l = \begin{cases} 1, & \text{если } i = j = l, \\ 0, & \text{во всех остальных случаях,} \end{cases} \quad (370)$$

а функции поличисловой переменной в таком представлении имеют вид

$$F(\hat{X}) = f^1(x^1, \dots, x^{k+m}) \hat{\Psi}_1 + \dots + f^{k+m}(x^1, \dots, x^{k+m}) \hat{\Psi}_{k+m}, \quad (371)$$

$f^i \in R$, если $i = 1, 2, \dots, k$, и $f^j \in C$, если $j = k+1, k+2, \dots, k+m$. Подставим $\tilde{p}_{ij}^l$ в соотношения Коши - Римана (363), получим

$$\frac{\partial f^i}{\partial x^{i-}} = f^{i-}, \quad (372)$$

а все остальные частные производные $\frac{\partial f^l}{\partial x^i}$ при $i \neq j$ равны нулю. С учётом этого формула (371) для аналитических функций поличисловой переменной уточняется:

$$F(\hat{X}) = f^i(x^1)\hat{\Psi}_1 + \dots + f^k(x^k)\hat{\Psi}_k + \\ + f^{k+1}(x^{k+1})\hat{\Psi}_{k+1} + \dots + f^{k+m}(x^{k+m})\hat{\Psi}_{k+m},$$

$x^i, f^i \in R$, если $i = 1, 2, \dots, k$, и $x^j, f^j \in C$, если $j = k+1, k+2, \dots, k+m$.

Перейдем опять изотропному базису $e_1, e_2, \dots, e_n$ со структурным тензором (320), в котором поличисловая переменная X имеет n действительных координат $x^1, x^2, \dots, x^n$. Тогда с учётом выше изложенного произвольная аналитическая функция переменной $P_{k+2 \cdot m}$ в изотропном базисе имеет вид

$$F(X) = f^1(x^1)e_1 + \dots + f^k(x^k)e_k + f^{k+1}(x^{k+1}, x^{k+2})e_{k+1} + \\ + f^{k+2}(x^{k+1}, x^{k+2})e_{k+2} + \dots + f^{n-1}(x^{n-1}, x^n)e_{n-1} + f^n(x^{n-1}, x^n)e_n, \quad (373)$$

где $f^1(x^1), \dots, f^k(x^k)$ - произвольные гладкие функции одной действительной переменной, а пары функций

$$\{f^{j-}(x^{j-}, x^{j+1}), f^{(j-+1)}(x^{j-}, x^{j+1})\}, \quad j \equiv j_- = k+1, k+3, \dots, n-1 \quad (374)$$

являются компонентами аналитических функций комплексных переменных $(x^j; x^{j+1})$.

Так как аналитические функции комплексной переменной бесконечное число раз дифференцируемы, то аналитическая функция $F(X)$ переменной $P_{k+2 \cdot m}$ столько раз дифференцируема, сколько раз дифференцируемы все компоненты $f^1(x^1), \dots, f^k(x^k)$.

Итак, в общем случае аналитическая функция $P_{k+2 \cdot m}$ переменной однозначно определяется, если заданы k гладких функций одной действительной переменной и m аналитических функций комплексной переменной. Как известно [4], аналитическая функция комплексной переменной однозначно определяется заданием на некотором множестве точек в области ее определения, например, на отрезке некоторой кривой. Для всех элементарных функций одной действительной переменной можно с помощью аналитического продолжения построить однозначно аналитическую функцию комплексной переменной. Последнее свойство выполняется и для аналитических функций $P_{k+2 \cdot m}$ переменной.

Используем компоненты аналитической функции $F(X)$ для перехода от системы координат $x^{i'}$, в которой элемент длины в пространстве $P_{k+2 \cdot m}$ не зависит от точки пространства, к некоторой криволинейной системе координат x^i , в которой элемент длины финслерова пространства будет зависеть не только от дифференциалов координат, но и от самих координат. Пусть

$$x^{i'} = f^i(x^1, x^2, \dots, x^n), \quad (375)$$

где f^i компоненты аналитической функции $F(X)$, штрих у индекса указывает на принадлежность к другой системе координат. Необходимо, чтобы якобиан такого преобразования был конечен и отличен от нуля. Достаточно это проверить в каком-то одном базисе, например, в изотропном. Тогда согласно формулам (363) и (331), получим:

$$\frac{D(f^1, f^2, \dots, f^n)}{D(x^1, x^2, \dots, x^n)} = (F', F', \dots, F') \neq 0, \quad (376)$$

где

$$F'(X) = \frac{df^1}{dx^1}e_1 + \dots + \frac{df^k}{dx^k}e_k + \\ + \frac{\partial f^{k+1}}{\partial x^{k+1}}e_{k+1} + \frac{\partial f^{k+2}}{\partial x^{k+1}}e_{k+2} + \dots + \frac{\partial f^{n-1}}{\partial x^{n-1}}e_{n-1} + \frac{\partial f^n}{\partial x^{n-1}}e_n$$

- производная аналитической функции $F(X)$ переменной $X \in P_{k+2,m}$, а

$$(F', F', \dots, F') = \frac{df^1}{dx^1} \dots \frac{df^k}{dx^k} \times \\ \times \left[\left(\frac{\partial f^{k+1}}{\partial x^{k+1}} \right)^2 + \left(\frac{\partial f^{k+1}}{\partial x^{k+2}} \right)^2 \right] \dots \left[\left(\frac{\partial f^{n-1}}{\partial x^{n-1}} \right)^2 + \left(\frac{\partial f^{n-1}}{\partial x^n} \right)^2 \right].$$

Здесь мы не использовали обозначение нормы в n -ой степени $|F'|^n$ функции $F'(X)$ вместо формы $(F', F', \dots, F')$, так как функция $F'(X)$ при этом может не иметь нормы.

Таким образом, преобразование координат (375) или отображение одной области пространства $P_{k+2,m}$ на другую область того же пространства, осуществляемое с помощью компонент аналитической функции $F(X)$, взаимно однозначно в некоторой области, если в этой области скалярное полипроизведение с одним и тем же аргументом $F'(X)$ имеет конечное значение, не равное нулю.

Последнее утверждение и формула (376) показывают, что понятие скалярного полипроизведения, впервые введенное в работе [18], естественным образом возникает при рассмотрении преобразований пространства $P_{k+2,m}$ с помощью аналитических функций $P_{k+2,m}$ переменной и не сводится к понятию нормы и метрики.

Посмотрим как изменится метрическая функция $L(dx)$ (340) при таком преобразовании координат или преобразовании самого пространства. Для этого вначале потребуем выполнения условия

$$(F', F', \dots, F') > 0, \quad (377)$$

тогда у функции F' существует норма

$$|F'| = \sqrt[n]{(F', F', \dots, F')}.$$

Подставим (375) в (340) и получим ту же метрическую функцию, но уже в координатах x^i ,

$$L(dx^{1'}, dx^{2'}, \dots, dx^{n'}) = |F'(X)| \cdot L(dx^1, dx^2, \dots, dx^n), \quad (378)$$

умноженную на коэффициент растяжения-сжатия

$$\kappa(x) = |F'(X)|.$$

Подчеркнём ещё раз, что слева и справа в формуле (378) стоит одна и та же функция $L(\xi)$.

Таким образом, аналитическая функция невырожденной поличисловой переменной осуществляет в некоторой области D финслерова поличислового пространства конформное преобразование координат или конформное преобразование самой области D этого пространства, если ее производная имеет отличную от нуля норму, причём норма производной суть коэффициент растяжения-сжатия.

Это же утверждение можно сформулировать несколько иначе.

Пусть финслерова геометрия $L(dx; x)$ конформно связана с финслеровой геометрией $L(dx)$ плоского поличислового пространства $P_{k+2\cdot m}$, то есть

$$L(dx; x) = \kappa(x)L(dx), \quad (379)$$

причем

$$\kappa(x) = |F'(X)| \neq 0, \quad (380)$$

где $F'(X)$ - производная, имеющая конечную норму, аналитической функции $F(X)$ поличисловой переменной $P_{k+2\cdot m}$, тогда геометрия $L(dx; x)$ получается из плоской геометрии $L(dx)$ пространства $P_{k+2\cdot m}$ введением криволинейных координат с помощью компонент аналитической функции $F(X)$, а значит, геометрия $L(dx; x)$ также является той же самой плоской геометрией, что и исходная геометрия.

Итак, в любом невырожденном поличисловом пространстве P_n размерности $n \geq 3$ имеется бесконечно параметрическая непрерывная группа конформных преобразований, в то время как в евклидовых и псевдоевклидовых пространствах размерности $n \geq 3$ непрерывная конформная группа всегда конечно параметрическая.

Аналогично тому как конформные преобразования на евклидовой плоскости связаны не только с аналитическими функциями комплексной переменной, но и с функциями, комплексно сопряжёнными к аналитическим, (последние приводят к конформным преобразованиям второго рода), конформные преобразования в пространстве поличисел $P_{k+2\cdot m}$ также не сводятся только к аналитическим функциям поличисловой переменной. Изменение знака у $(k+2)$ -ой, $(k+4)$ -ой, ..., n -ой компонент аналитической функции поличисловой переменной даёт 2^m возможностей; произвольная подстановка 1-ой, 2-ой, ..., k -ой компонент аналитической функции также приводит к функции, осуществляющей конформное преобразование, что даёт $k!$ возможностей; произвольная подстановка $(k-1)$ -ой, $(k-3)$ -ой, ..., $(n-1)$ -ой пар компонент аналитической функции поличисловой переменной также приводит к функции, осуществляющей конформное преобразование, что даёт $m!$ возможностей. Таким образом, конформные преобразования в пространстве поличисел $P_{k+2\cdot m}$ связаны с аналитическими функциями поличисловой переменной и имеют, если их различать по родам, не менее $2^m \cdot k! \cdot m!$ родов.

16 Пространство гиперкомплексных чисел H_4

Поличисла $H_4 \equiv P_{4+2\cdot 0}$ изоморфны алгебре действительных квадратных диагональных матриц 4×4 . Координаты в изотропном базисе пространства H_4 будем обозначать $\xi^1, \xi^2, \xi^3, \xi^4$, а сам базис - $\psi_1, \psi_2, \psi_3, \psi_4$, то есть любое поличисло можно представить в виде разложения

$$X = \xi^1 \psi_1 + \xi^2 \psi_2 + \xi^3 \psi_3 + \xi^4 \psi_4, \quad (381)$$

или четверки действительных чисел $(\xi^1, \xi^2, \xi^3, \xi^4)$. Бинарная операция поличислового умножения на множестве базисных векторов $\psi_1, \psi_2, \psi_3, \psi_4$ определяется следующей формулой:

$$\psi_i \cdot \psi_j = \delta_{ij} \psi_{j-}.$$

Если $\xi^1 \xi^2 \xi^3 \xi^4 > 0$, то для числа $X \in H_4$ определена норма

$$|X| = \sqrt[4]{\xi^1 \xi^2 \xi^3 \xi^4} > 0. \quad (382)$$

Пространство H_4 является финслеровым пространством с метрической функцией

$$L(d\xi) = \sqrt[4]{d\xi^1 d\xi^2 d\xi^3 d\xi^4}. \quad (383)$$

Никакими преобразованиями координат подкоренное выражение в правой части нельзя привести к квадратичному виду или квадрату однородной дифференциальной формы второго порядка, то есть такое финслерово пространство качественно отличается от евклидова и псевдоевклидовых пространств размерности четыре. Метрику, определяемую метрической функцией (383), называют иногда метрикой Бервальда - Моора.

В пространстве H_4 существует базис $1, j, k, jk$,

$$\left. \begin{aligned} 1 &= \psi_1 + \psi_2 + \psi_3 + \psi_4, \\ j &= \psi_1 + \psi_2 - \psi_3 - \psi_4, \\ k &= \psi_1 - \psi_2 + \psi_3 - \psi_4, \\ jk &= \psi_1 - \psi_2 - \psi_3 + \psi_4, \end{aligned} \right\} \quad (384)$$

закон умножения базисных элементов которого определен формулами:

$$\begin{aligned} 1 \cdot 1 &= 1, & 1 \cdot j &= j, & 1 \cdot k &= k, & 1 \cdot jk &= jk, \\ j \cdot 1 &= j, & j \cdot j &= 1, & j \cdot k &= jk, & j \cdot jk &= k, \\ k \cdot 1 &= k, & k \cdot j &= jk, & k \cdot k &= 1, & k \cdot jk &= j, \\ jk \cdot 1 &= jk, & jk \cdot j &= k, & jk \cdot k &= j, & jk \cdot jk &= 1. \end{aligned}$$

Координаты x^0, x^1, x^2, x^3 числа X в этом базисе связаны с координатами $\xi^1, \xi^2, \xi^3, \xi^4$ того же числа в изотропном базисе формулами:

$$\left. \begin{aligned} \xi^1 &= x^0 + x^1 + x^2 + x^3, \\ \xi^2 &= x^0 + x^1 - x^2 - x^3, \\ \xi^3 &= x^0 - x^1 + x^2 - x^3, \\ \xi^4 &= x^0 - x^1 - x^2 + x^3. \end{aligned} \right\} \quad (385)$$

Базис $1, j, k, jk$ является "ортонормированным" и в нем имеет место экспоненциальное представление чисел, у которых

$$\xi^1 > 0, \xi^2 > 0, \xi^3 > 0, \xi^4 > 0.$$

Для того, чтобы проверить первое утверждение, достаточно вычислить всевозможные скалярные полипроизведения вида (A, B, B, B) , где A, B - любые базисные элементы $1, j$,

k, jk . Покажем, как происходит вычисление таких скалярных полипроизведений на одном примере, а затем приведем результаты вычислений.

Выразим скалярное полипроизведение (j, k, k, k) через антиопределитель:

$$(j, k, k, k) = \frac{1}{4!} \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{vmatrix}_+ . \quad (386)$$

Так как из любой строки можно вынести общий числовой множитель за знак перманента, вынесем (-1) из второй и третьей строки

$$(j, k, k, k) = \frac{1}{4!} \begin{vmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{vmatrix}_+ \quad (387)$$

и воспользуемся выше полученной формулой вычисления антиопределителя, у которого все столбцы одинаковые, кроме одного:

$$(j, k, k, k) = 0 . \quad (388)$$

Действуя аналогичным образом, вычислим скалярные полипроизведения (A, B, B, B) , где A, B любые базисные элементы $1, j, k, jk$. Приведем результаты вычислений:

$$\left. \begin{aligned} (1, 1, 1, 1) &= (j, j, j, j) = (k, k, k, k) = (jk, jk, jk, jk) = 1 , \\ (1, j, j, j) &= (1, k, k, k) = (1, jk, jk, jk) = 0 , \\ (j, 1, 1, 1) &= (k, 1, 1, 1) = (jk, 1, 1, 1) = 0 , \\ (j, k, k, k) &= (k, j, j, j) = (j, jk, jk, jk) = 0 , \\ (jk, j, j, j) &= (k, jk, jk, jk) = (jk, k, k, k) = 0 . \end{aligned} \right\} \quad (389)$$

Таким образом, базис $1, j, k, jk$ является "ортонормированным" а значит коэффициенты разложения числа X в этом базисе могут быть выражены через скалярные полипроизведения следующим образом:

$$X = (X, 1, 1, 1) \cdot 1 + (X, j, j, j) \cdot j + (X, k, k, k) \cdot k + (X, jk, jk, jk) \cdot jk . \quad (390)$$

Базисные элементы j, k, jk получаются из базисных элементов E_2, E_3, E_4 (356) с помощью невырожденного линейного преобразования:

$$j = -E_2 + E_3 + E_4, \quad k = E_2 - E_3 + E_4, \quad jk = E_2 + E_3 - E_4, \quad (391)$$

так как

$$\begin{vmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{vmatrix} = 4 \neq 0 . \quad (392)$$

Поэтому в базисе $1, j, k, jk$ для чисел $X \in H_4$, имеющих отличную от нуля норму $|X| \neq 0$ и удовлетворяющих ещё некоторым требованиям, возможно экспоненциальное представление:

$$X = |X| \cdot \exp(\alpha j + \beta k + \gamma jk), \quad (393)$$

где α, β, γ - действительные числа, угловые переменные.

Если функция $F(x)$ одной действительной переменной x представима как полином или сходящийся ряд переменной x , то функция $F(X)$ переменной $X \in H_4$ в изотропном базисе также определена,

$$F(X) = F(\xi^1)\psi_1 + F(\xi^2)\psi_2 + F(\xi^3)\psi_3 + F(\xi^4)\psi_4 \quad (394)$$

и является аналитической функцией H_4 переменной. Применяя эту формулу к экспоненциальному представлению (393), имеем

$$\begin{aligned} X &= \xi^1\psi_1 + \xi^2\psi_2 + \xi^3\psi_3 + \xi^4\psi_4 = \\ &= |X| \left(e^{\alpha+\beta+\gamma}\psi_1 + e^{\alpha-\beta-\gamma}\psi_2 + e^{-\alpha+\beta-\gamma}\psi_3 + e^{-\alpha-\beta+\gamma}\psi_4 \right), \end{aligned} \quad (395)$$

откуда находим выражения угловых переменных через координаты числа X в изотропном базисе:

$$\alpha = \ln \left(\frac{\xi^1\xi^2}{\xi^3\xi^4} \right), \quad \beta = \ln \left(\frac{\xi^1\xi^3}{\xi^2\xi^4} \right), \quad \gamma = \ln \left(\frac{\xi^1\xi^4}{\xi^2\xi^3} \right). \quad (396)$$

Формула (395) дает и условия, при выполнении которых поличисло $X \in H_4$ может быть представлено в экспоненциальном виде:

$$\xi^1 > 0, \quad \xi^2 > 0, \quad \xi^3 > 0, \quad \xi^4 > 0.$$

Область, выделяемая в пространстве H_4 этими неравенствами, отождествляется с конусом будущего, если интерпретировать пространство H_4 как четырехмерное пространство-время.

Итак, все поличисла $X \in H_4$, лежащие в конусе будущего, имеют в ортонормированном базисе $1, j, k, jk$ экспоненциальное представление.

Из формулы (394) следует, что произвольная элементарная функция одной действительной переменной однозначно определяет аналитическую функцию переменной H_4 .

Любая аналитическая функция переменной H_4 в изотропном базисе $\psi_1, \psi_2, \psi_3, \psi_4$ имеет вид

$$F(X) = f^1(\xi^1)\psi_1 + f^2(\xi^2)\psi_2 + f^3(\xi^3)\psi_3 + f^4(\xi^4)\psi_4, \quad (397)$$

где f^i - произвольные дифференцируемые функции одной действительной переменной. Если все функции f^i дифференцируемы $r+1$ раз, то все производные функции $F(X)$ переменной H_4 вплоть до r -ой также являются аналитическими функциями. Запись произвольной аналитической функции (397) в "ортонормированном" базисе $1, j, k, jk$ (384) имеет более

громоздкий вид:

$$\begin{aligned}
F(X) &= \\
&= \frac{1}{4}[f^1(x^0 + x^1 + x^2 + x^3) + f^2(x^0 + x^1 - x^2 - x^3) + \\
&\quad + f^3(x^0 - x^1 + x^2 - x^3) + f^4(x^0 - x^1 - x^2 + x^3)] \cdot \mathbf{1} + \\
&\quad + \frac{1}{4}[f^1(x^0 + x^1 + x^2 + x^3) + f^2(x^0 + x^1 - x^2 - x^3) - \\
&\quad - f^3(x^0 - x^1 + x^2 - x^3) - f^4(x^0 - x^1 - x^2 + x^3)] \cdot \mathbf{j} + \\
&\quad + \frac{1}{4}[f^1(x^0 + x^1 + x^2 + x^3) - f^2(x^0 + x^1 - x^2 - x^3) + \\
&\quad + f^3(x^0 - x^1 + x^2 - x^3) - f^4(x^0 - x^1 - x^2 + x^3)] \cdot \mathbf{k} + \\
&\quad + \frac{1}{4}[f^1(x^0 + x^1 + x^2 + x^3) - f^2(x^0 + x^1 - x^2 - x^3) - \\
&\quad - f^3(x^0 - x^1 + x^2 - x^3) + f^4(x^0 - x^1 - x^2 + x^3)] \cdot \mathbf{jk}.
\end{aligned} \tag{398}$$

В "ортонормированных" координатах x^0, x^1, x^2, x^3 метрическая функция (383) финслерова пространства H_4 записывается следующим образом:

$$\begin{aligned}
L(dx) &= \\
&= [(dx^0 + dx^1 + dx^2 + dx^3)(dx^0 + dx^1 - dx^2 - dx^3) \times \\
&\quad \times (dx^0 - dx^1 + dx^2 - dx^3)(dx^0 - dx^1 - dx^2 + dx^3)]^{1/4}.
\end{aligned} \tag{399}$$

Разложим выражение в квадратных скобках по степеням dx^0 , получим

$$\begin{aligned}
L^4(dx) &= (dx^0)^4 - 2(dx^0)^2 \left[(dx^1)^2 + (dx^2)^2 + (dx^3)^2 \right] + \\
&\quad + 8 dx^0 dx^1 dx^2 dx^3 + (dx^1)^4 + (dx^2)^4 + (dx^3)^4 - \\
&\quad - 2 \left[(dx^1)^2 (dx^2)^2 + (dx^1)^2 (dx^3)^2 + (dx^2)^2 (dx^3)^2 \right].
\end{aligned} \tag{400}$$

Предположим, что

$$dx^\alpha = \varepsilon^\alpha dx^0, \quad 0 < \varepsilon^\alpha \ll 1, \tag{401}$$

тогда, если пренебречь членами ε^3 по сравнению с 1, что соответствует в механике пренебрежению членами $\left(\frac{v}{c}\right)^3$, где v - скорость частицы, а c - скорость света, по сравнению с

единицей (нерелятивистское приближение), для элемента длины $L(dx)$ (399) в пространстве H_4 получим приближенную формулу:

$$L_{H_4}(dx) \simeq \sqrt{(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2}. \quad (402)$$

Справа в этой формуле стоит элемент длины в пространстве Минковского, поэтому справедливо следующее утверждение.

Координаты x^0, x^1, x^2, x^3 в "ортонормированном" базисе пространства H_4 в нерелятивистском приближении в геометрическом (метрическом) плане ведут себя также как общепринятые координаты четырехмерного пространства-времени Минковского, которые в СТО отождествляются с координатами пространства событий.

Непрерывная группа $G_1(H_4)$ изометрической симметрии в пространстве H_4 является трёхпараметрической абелевой группой Ли, элементами которой являются матрицы

$$\begin{pmatrix} e^{\varepsilon_1} & 0 & 0 & 0 \\ 0 & e^{\varepsilon_2} & 0 & 0 \\ 0 & 0 & e^{\varepsilon_3} & 0 \\ 0 & 0 & 0 & e^{\varepsilon_4} \end{pmatrix}, \quad \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 = 0,$$

действующие в координатном пространстве $\xi_1, \xi_2, \xi_3, \xi_4$. Эта группа не содержит в качестве подгруппы никакой группы, даже отдалённо напоминающей группу трёхмерных пространственных поворотов. И всё же группа $G_1(H_4)$ содержит три однопараметрических преобразования, с помощью которых можно переходить в другие системы отсчёта, в которых наблюдаемые объекты будут иметь другую скорость относительно наблюдателя. Для того, чтобы это показать, введём вначале несколько обозначений, причём некоторое время для упрощения формул все индексы будем писать внизу:

$$\xi_i = A_{ij}x_j, \quad (A_{ij}) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}, \quad A_{ik}A_{kj} = 4\delta_{ij}, \quad (403)$$

$$W = (1 + v_1 + v_2 + v_3)(1 + v_1 - v_2 - v_3)(1 - v_1 + v_2 - v_3)(1 - v_1 - v_2 + v_3), \quad (404)$$

$$\sqrt{1 - v^2} = \sqrt[4]{W(v_1, v_2, v_3)}, \quad (405)$$

или

$$v = \sqrt{1 - \sqrt{W(v_1, v_2, v_3)}}. \quad (406)$$

В нерелятивистском приближении

$$v \simeq \sqrt{v_1^2 + v_2^2 + v_3^2}. \quad (407)$$

Если же вектор $(1, v_1, v_2, v_3)$ стремится к изотропному направлению, то $v \rightarrow 1$. Отметим также, что в общем случае

$$W(-v_1, -v_2, -v_3) \neq W(v_1, v_2, v_3).$$

Сложение скоростей.

Непрерывная группа изометрической симметрии $G_1(H_4)$, оставляющая инвариантным элемент длины (399) в пространстве H_4 и состоящая из непрерывных линейных преобразований вида

$$x'_i = \frac{1}{4} A_{ik} D_{km} A_{mj} x_j, \quad (408)$$

где

$$(D_{km}) = \text{diag}(\exp \varepsilon_0, \exp \varepsilon_1, \exp \varepsilon_2, \exp \varepsilon_3), \quad (409)$$

действительные параметры ε_i изменяются в пределах $(-\infty, \infty)$ и связаны соотношением

$$\varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0, \quad (410)$$

может быть параметризована тремя действительными величинами, V_1, V_2, V_3 , имеющими смысл компонент скорости, которую приобретает покоящийся объект после преобразования (408):

$$\exp \varepsilon_i = \frac{A_{ij} V_j}{\sqrt{1 - V^2}}, \quad (411)$$

где $i, j = 0, 1, 2, 3$; $V_0 = 1$. Если в исходной системе отсчёта объект имел скорость (v_1, v_2, v_3) , то в новой системе отсчёта этот же объект будет иметь скорость

$$\left. \begin{aligned} v'_1 &= \frac{v_1 + V_1 + v_2 V_3 + v_3 V_2}{1 + v_1 V_1 + v_2 V_2 + v_3 V_3}, \\ v'_2 &= \frac{v_2 + V_2 + v_1 V_3 + v_3 V_1}{1 + v_1 V_1 + v_2 V_2 + v_3 V_3}, \\ v'_3 &= \frac{v_3 + V_3 + v_1 V_2 + v_2 V_1}{1 + v_1 V_1 + v_2 V_2 + v_3 V_3}. \end{aligned} \right\} \quad (412)$$

Из определения группы $G_1(H_4)$ имеем

$$(x'_{(2)0} - x'_{(1)0}) \sqrt{1 - (v')^2} = (x_{(2)0} - x_{(1)0}) \sqrt{1 - v^2}, \quad (413)$$

что даёт формулу для модуля трёхмерной скорости в новой системе координат

$$v' = \sqrt{1 - \frac{(1 - v^2)(1 - V^2)}{(1 + v_1 V_1 + v_2 V_2 + v_3 V_3)^2}}, \quad (414)$$

так как при преобразованиях (408) - (411)

$$x'_{(2)0} - x'_{(1)0} = \frac{1 + v_1 V_1 + v_2 V_2 + v_3 V_3}{\sqrt{1 - V^2}} (x_{(2)0} - x_{(1)0}). \quad (415)$$

Если среди компонент v_α и V_α отличны от нуля только по одной компоненте вдоль одного и того же специального направления, например, $(v_1, 0, 0)$ и $(V_1, 0, 0)$, то формулы (412) совпадают с соответствующими формулами сложения скоростей в СТО.

**Переход из покоящейся инерциальной системы
в движущуюся систему отсчёта.**

Рассмотрим переход из старой (нештрихованной) системы отсчета в новую (штрихованную) систему отсчета, инерциально движущуюся со скоростью (V_1, V_2, V_3) относительно старой системы отсчета, то есть точка, движущаяся в старой системе отсчета со скоростью (V_1, V_2, V_3) , в новой системе отсчета будет двигаться со скоростью $(0, 0, 0)$. Формулы (408) - (410) останутся теми же, а вместо формулы (411) получим

$$\exp(-\varepsilon_i) = \frac{A_{ij}V_j}{\sqrt{1-V^2}}, \quad (416)$$

то есть переходы от одной системы отсчета к другой, рассмотренные выше и в данном случае, являются обратными друг к другу. Заметим, что замена в преобразовании (408) - (411) (V_1, V_2, V_3) на $(-V_1, -V_2, -V_3)$ в общем случае не дает обратное преобразование к преобразованию (408) - (411).

Итак, при переходе к системе координат движущейся со скоростью (V_1, V_2, V_3) старые координаты, будут выражаться через новые по формулам

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{4\sqrt{1-V^2}} \cdot \hat{A} \cdot \begin{pmatrix} (1+V_1+V_2+V_3)(x'_0+x'_1+x'_2+x'_3) \\ (1+V_1-V_2-V_3)(x'_0+x'_1-x'_2-x'_3) \\ (1-V_1+V_2-V_3)(x'_0-x'_1+x'_2-x'_3) \\ (1-V_1-V_2+V_3)(x'_0-x'_1-x'_2+x'_3) \end{pmatrix}, \quad (417)$$

где матрица $\hat{A}$ имеет компоненты A_{ij} (403).

Рассмотрим это преобразование, когда компоненты скорости новой системы отсчета в старой по трем специальным направлениям равны нулю, кроме одной компоненты, например, $V_1 \neq 0$, а $V_2 = 0$ и $V_3 = 0$, тогда

$$V = |V_1|, \quad (418)$$

а формулы (417) принимают вид

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-V_1^2}} & \frac{V_1}{\sqrt{1-V_1^2}} & 0 & 0 \\ \frac{V_1}{\sqrt{1-V_1^2}} & \frac{1}{\sqrt{1-V_1^2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-V_1^2}} & \frac{V_1}{\sqrt{1-V_1^2}} \\ 0 & 0 & \frac{V_1}{\sqrt{1-V_1^2}} & \frac{1}{\sqrt{1-V_1^2}} \end{pmatrix} \cdot \begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}, \quad (419)$$

или

$$\left. \begin{aligned} x_0 &= \frac{x'_0 + V_1 x'_1}{\sqrt{1-V_1^2}} & x_1 &= \frac{V_1 x'_0 + x'_1}{\sqrt{1-V_1^2}} \\ x_2 &= \frac{x'_2 + V_1 x'_3}{\sqrt{1-V_1^2}} & x_3 &= \frac{V_1 x'_2 + x'_3}{\sqrt{1-V_1^2}} \end{aligned} \right\}. \quad (420)$$

Такое преобразование координат $(x'_0, x'_1) \leftrightarrow (x_0, x_1)$ совпадает с соответствующим преобразованием в СТО, а преобразование координат $(x'_2, x'_3) \leftrightarrow (x_2, x_3)$ отличается от соответствующих преобразования в СТО, где $x_2 = x'_2, x_3 = x'_3$.

В пространстве событий всегда возникает вопрос о том, что такое истинное пространственное трёхмерное расстояние. Если истинное пространственное расстояние r_{H_4} в H_4 искать тем же способом, как это обычно делается в СТО и ОТО [11], то в результате получим [15]:

$$r_{H_4} \equiv r_d = |x^\alpha| + |x^\beta|; \quad \text{причем} \quad |x^\gamma| \leq |x^\alpha|, |x^\beta|. \quad (421)$$

Здесь $\alpha, \beta, \gamma = 1, 2, 3$; $\alpha \neq \beta, \alpha \neq \gamma, \beta \neq \gamma$. В евклидовом пространстве гиперповерхность, определяемая уравнением

$$r_d = \text{const}, \quad (422)$$

является ромбододекаэдром [16]. Так как такое трёхмерное пространство, элементами которого являются параллельные прямые мировые линии, есть плоское пространство, то все индикатрисы такого пространства являются ромбододекаэдрами, а метрическая функция определяется формулой:

$$L(dx^1, dx^2, dx^3) = |dx^\alpha| + |dx^\beta|, \quad |dx^\gamma| \leq |dx^\alpha|, |dx^\beta|, \quad (423)$$

где $\alpha, \beta, \gamma = 1, 2, 3$; $\alpha \neq \beta, \alpha \neq \gamma, \beta \neq \gamma$. Но пространство с такой метрической функцией не является финслеровым.

17 Конформный потенциал

Успехи приложения теории функций комплексной переменной к решению задач физики, особенно решению уравнения Лапласа на евклидовой плоскости, подвигли многих математиков и физиков к попыткам построения аналогичных теорий для других пространств. Особое значение, конечно, здесь занимает четырёхмерное пространство-время.

Теория конформного потенциала строится для пространств, конформно связанных с финслеровыми пространствами, у которых метрические функции однозначно определены. При этом выделяется множество плоских финслеровых пространств, взятых в качестве исходных, а среди последних - класс поличисловых пространств.

Если в качестве исходных пространств берутся поличисловые пространства, то тогда будем говорить о гиперкомплексном потенциале, так как получающаяся теория наиболее близка к теории комплексного потенциала на евклидовой плоскости.

Пусть $x^1, x^2, \dots, x^n$ - финслерово пространство с метрической функцией

$$L(\xi; x) \equiv L(\xi^1, \xi^2, \dots, \xi^n; x^1, x^2, \dots, x^n), \quad (424)$$

где $\xi^1, \xi^2, \dots, \xi^n$ - касательное пространство в точке $M(x^1, x^2, \dots, x^n)$ основного пространства. Тогда элемент длины определяется формулой

$$ds = L(dx^1, dx^2, \dots, dx^n; x^1, x^2, \dots, x^n), \quad (425)$$

а компоненты обобщенного импульса

$$p_i = \frac{\partial L(dx; x)}{\partial (dx^i)} \quad (426)$$

связаны тангенциальным уравнением индикатрисы

$$\Phi(p; x) = 0. \quad (427)$$

В финслеровых пространствах тангенциальному уравнению индикатрисы всегда можно придать некоторую специальную форму

$$\Phi_m(p; x) - 1 = 0, \quad (428)$$

где $\Phi_m(p; x)$ - однородная функция m -го порядка ($m > 0$) по первым n аргументам. Очевидно, что функция Финслера в этом случае суть функция

$$\Phi(p; x) = \Phi_m(p; x) - 1.$$

Если функция $S(x)$ определяет в рассматриваемом финслеровом пространстве нормальную конгруэнцию геодезических, то она должна удовлетворять дифференциальному уравнению

$$\Phi \left(\frac{\partial S}{\partial x^1}, \dots, \frac{\partial S}{\partial x^n}; x \right) = 0. \quad (429)$$

Функцию $S(x)$ в классической механике называют действием как функцией координат, а уравнение, соответствующее уравнению (429), - уравнением Гамильтона - Якоби. Если функция $S(x)$ известна, то поле обобщенных импульсов находится как

$$p_i = \frac{\partial S}{\partial x^i}, \quad (430)$$

а экстремали (мировые линии, траектории движения, линии тока) находятся из системы уравнений

$$\dot{x}^i = \left. \frac{\partial \Phi(p; x)}{\partial p_i} \right|_{p_k = \frac{\partial S}{\partial x^k}} \cdot \lambda(x),$$

или

$$\dot{x}^i = \left. \frac{\partial \Phi_m(p; x)}{\partial p_i} \right|_{p_k = \frac{\partial S}{\partial x^k}} \cdot \lambda(x), \quad (431)$$

где $\lambda(x) \neq 0$ - некоторая функция, $\dot{x}^i \equiv \frac{dx^i}{d\tau}$, а τ - параметр вдоль мировой линии, параметр эволюции.

Финслерово пространство, конформно связанное с исходным, имеет метрическую функцию вида

$$\tilde{L}(\xi; x) = \kappa(x)L(\xi; x), \quad (432)$$

где $\kappa(x) > 0$ - некая скалярная функция. Элемент длины в таком пространстве определяется формулой

$$d\tilde{s} = \kappa(x)L(dx; x). \quad (433)$$

Обобщенные импульсы

$$\tilde{p}_i = \kappa(x) \frac{\partial L(dx; x)}{\partial (dx^i)} \quad (434)$$

связаны соотношением

$$\Phi_m(\tilde{p}; x) - \kappa^m = 0, \quad (435)$$

то есть функция Финслера пространства, конформно связанного с основным пространством, может быть определена следующим образом:

$$\tilde{\Phi}(\tilde{p}; x) = \Phi_m(\tilde{p}; x) - \kappa^m. \quad (436)$$

Пусть $S_W(x)$ - произвольная скалярная функция, тогда в области, где выполняется неравенство

$$\Phi_m \left(\frac{\partial S_W}{\partial x^1}, \dots, \frac{\partial S_W}{\partial x^n}; x \right) > 0, \quad (437)$$

определена финслерова геометрия, конформно связанная с исходной и полем коэффициента растяжения-сжатия

$$\kappa(x) = \sqrt[m]{\Phi_m \left(\frac{\partial S_W}{\partial x^1}, \dots, \frac{\partial S_W}{\partial x^n}; x \right)}, \quad (438)$$

причем в этой же области функция $S_W(x)$ определяет нормальную конгруэнцию геодезических

$$\dot{x}^i = \left. \frac{\partial \Phi_m(\tilde{p}; x)}{\partial \tilde{p}_i} \right|_{\tilde{p}_k = \frac{\partial S_W}{\partial x^k}} \cdot \tilde{\lambda}(x), \quad (439)$$

где $\tilde{\lambda}(x) \neq 0$ - некоторая функция, $\dot{x}^i \equiv \frac{dx^i}{d\tilde{\tau}}$, а $\tilde{\tau}$ - параметр эволюции. Таким образом, функция $S_W(x)$ автоматически является действием как функцией координат в финслеровом пространстве с метрической функцией $\tilde{L}(\xi; x)$ (432).

Отметим, что при замене функции $S(x)$ и параметра эволюции τ на функцию $S_W(x)$ и параметр эволюции $\tilde{\tau}$ формула (431) формально совпадает с формулой (439), хотя геометрии разные.

Предположим, что инвариантный элемент объёма в исходном финслеровом пространстве с однозначно заданной метрической функцией $L(\xi; x)$ определён и имеет вид:

$$dV = \rho(x) \cdot dx^1 dx^2 \dots dx^n. \quad (440)$$

Тогда в конформно связанном пространстве с метрической функцией $\tilde{L}(\xi; x)$ (432) также определён инвариантный элемент объёма:

$$d\tilde{V} = \rho(x) \left[\Phi_m \left(\frac{\partial S_W}{\partial x^1}, \dots, \frac{\partial S_W}{\partial x^n}; x \right) \right]^{\frac{n}{m}} dx^1 dx^2 \dots dx^n. \quad (441)$$

Согласно принципу самодостаточности финслеровой геометрии, если скалярное поле не задано, оно определяется из уравнения поля Лагранжа - Эйлера - Остроградского [5] с лагранжианом

$$\mathcal{L} = const \cdot \rho(x) \left[\Phi_m \left(\frac{\partial S_W}{\partial x^1}, \dots, \frac{\partial S_W}{\partial x^n}; x \right) \right]^{\frac{n}{m}}. \quad (442)$$

Так как лагранжиан зависит только от производных поля S_W , уравнение поля принимает вид:

$$\frac{\partial}{\partial x^i} \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial S_W}{\partial x^i} \right)} = 0,$$

или

$$\frac{\partial}{\partial x^i} \left[\varrho(x) \Phi_m^{\frac{n-m}{m}} \frac{\partial \Phi}{\partial \left(\frac{\partial S_W}{\partial x^i} \right)} \right] = 0. \quad (443)$$

Функция $S_W(x)$, удовлетворяющую уравнению поля (443), будем называть *Мировой функцией*, а само уравнение - *фундаментальным уравнением* исходного и конформно связанного с ним финслеровых пространств.

Любое невырожденное поличисловое пространство $P_n \ni X$ является финслеровым, причем любая компонента произвольной аналитической функции $F(X)$ в нем удовлетворяет уравнению поля для Мировой функции $S_W(x)$. Аналитические функции можно перемножать, брать их линейные комбинации, строить функцию от функции, при этом опять будут получаться аналитическая функция, каждая компонента которых будет удовлетворять уравнению поля для Мировой функции. Если не оговорено другое, в качестве Мировой функции в поличисловом пространстве будем брать компоненту $U(x^1, \dots, x^2)$ при единице

$$F(X) = U(x) \cdot 1 + V^1(x) \cdot E_2 + \dots + V^{n-1}(x) \cdot E_n \quad (444)$$

в базисе $1, E_2, \dots, E_n$, в котором имеет место экспоненциальное представление поличисел, то есть

$$X = |X| \exp(\alpha^1 \cdot E_2 + \dots + \alpha^{n-1} \cdot E_n), \quad (445)$$

тогда в этом базисе

$$\ln \left(\frac{X}{b} \right) = \ln \left(\frac{|X|}{|b|} \right) \cdot 1 + \alpha^1 \cdot E_2 + \dots + \alpha^{n-1} \cdot E_n, \quad (446)$$

где b - действительное число, а $|X|$ - модуль поличисла.

Теорию конформного потенциала в пространствах, конформно связанных с невырожденными поличисловыми пространствами, или построенную на основе поличисловых пространств, будем называть *теорией гиперкомплексного потенциала*. Именно в последней присутствует наиболее много черт, аналогичных свойствам, которые имеют место в теории комплексного потенциала на евклидовой плоскости, при этом ряд свойств и формул включают в себя не только компоненту $U(x)$, но и остальные компоненты: $V^1, V^2, \dots, V^{n-1}$ - соответствующей аналитической функции поличисловой переменной, как это имеет место и в теории комплексного потенциала на евклидовой плоскости.

Замечание. Пространства невырожденных поличисел являются частным случаем более широкого класса плоских финслеровых пространств, каждой точке которых, вообще говоря, нельзя сопоставить поличисло, но для этого более широкого класса пространств понятие конформного потенциала работает, хотя в общем случае оно не связано с понятием аналитических функций и лишь частично связано с конформными преобразованиями этих пространств.

18 Пространство двойных чисел

Двойные, или гиперболические числа H_2 можно определить как двумерную линейную алгебру, в которой существует базис $1, j$ со следующими свойствами:

$$H_2 \ni X = x^0 + x^1 \cdot j, \quad 1 \cdot 1 = 1, \quad 1 \cdot j = j \cdot 1 = j, \quad j^2 = 1, \quad (447)$$

где x^0, x^1 - действительные числа.

В качестве потенциала выберем действительную компоненту U аналитической функции

$$\begin{aligned} F(X) &\equiv U + j \cdot V = \\ &= a \cdot \ln \left(\frac{X}{b} \right) \equiv \frac{a}{2} \ln \left[\frac{(x^0)^2 - (x^1)^2}{b} \right] + j \cdot \frac{a}{2} \ln \left[\frac{x^0 + x^1}{x^0 - x^1} \right], \end{aligned} \quad (448)$$

a, b - действительные числа, то есть

$$U(x^0, x^1) = \frac{a}{2} \ln \left[\frac{(x^0)^2 - (x^1)^2}{b} \right]. \quad (449)$$

Ограничимся конусом будущего:

$$x^0 > |x^1|, \quad (450)$$

где аналитическая функция (448), а значит и потенциал (449), определены.

Для того, чтобы записать уравнения (439), которым подчиняются мировые линии материальных объектов, надо определиться с выбором функции Финслера и функции $\tilde{\lambda}(x)$. И хотя наблюдаемые величины не зависят от этого выбора, от него зависит конкретный вид уравнений (439).

Если мы хотим, чтобы формулы (439) напоминали формулы теории комплексного потенциала, то следует выбрать

$$\tilde{\Phi}(p; x) = p_0^2 - p_1^2 - \kappa^2(x^0, x^1), \quad \tilde{\lambda}(x^0, x^1) = \frac{1}{2}, \quad (451)$$

тогда

$$\dot{x}^0 = \frac{\partial U}{\partial x^0}, \quad \dot{x}^1 = -\frac{\partial U}{\partial x^1}, \quad (452)$$

где $\dot{x}^i \equiv \frac{dx^i}{d\tau}$, а τ - некий параметр вдоль мировой линии, параметр эволюции. В этом случае

$$\sqrt{(\dot{x}^0)^2 - (\dot{x}^1)^2} = \sqrt{\left(\frac{\partial U}{\partial x^0} \right)^2 - \left(\frac{\partial U}{\partial x^1} \right)^2} \equiv \kappa(x^0, x^1), \quad (453)$$

где $\kappa(x^0, x^1)$ - коэффициент растяжения-сжатия конформного преобразования в пространстве H_2 , которое (преобразование) определяется аналитической функцией (448). Таким образом, вдоль мировой линии в данном случае

$$\sqrt{(dx^0)^2 - (dx^1)^2} = \kappa(x^0, x^1) d\tau.$$

Если мы хотим, чтобы уравнения (439) по форме и физической интерпретации соответствовали формулам СТО, то следует выбрать

$$\tilde{\Phi}(p; x) = \sqrt{p_0^2 - p_1^2} - \kappa(x^0, x^1), \quad \tilde{\lambda}(x^0, x^1) = 1, \quad (454)$$

тогда

$$\dot{x}^0 = \frac{\frac{\partial U}{\partial x^0}}{\sqrt{\left(\frac{\partial U}{\partial x^0}\right)^2 - \left(\frac{\partial U}{\partial x^1}\right)^2}}, \quad \dot{x}^1 = -\frac{\frac{\partial U}{\partial x^1}}{\sqrt{\left(\frac{\partial U}{\partial x^0}\right)^2 - \left(\frac{\partial U}{\partial x^1}\right)^2}}, \quad (455)$$

где $\dot{x}^i \equiv \frac{dx^i}{d\tau}$, а τ - некий параметр вдоль мировой линии. В этом случае

$$\sqrt{(\dot{x}^0)^2 - (\dot{x}^1)^2} = 1, \quad (456)$$

то есть в этом случае параметр эволюции суть собственное время, соответствующее мировой линии в пространстве H_2 , умноженное на скорость света. Таким образом, вдоль мировой линии в данном случае

$$\sqrt{(dx^0)^2 - (dx^1)^2} = d\tau.$$

Выбирая формулы (454), для потенциала (449) получим дифференциальные уравнения для определения мировых линий, по которым двигаются материальные объекты, составляющие рассматриваемую замкнутую физическую систему:

$$\frac{dx^0}{d\tau} = \frac{x^0}{\sqrt{(x^0)^2 - (x^1)^2}}, \quad \frac{dx^1}{d\tau} = \frac{x^1}{\sqrt{(x^0)^2 - (x^1)^2}} \Rightarrow \frac{dx^1}{dx^0} = \frac{x^1}{x^0}. \quad (457)$$

Решениями этой системы уравнений являются мировые линии:

$$(x^0, x^1) = \left(\frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \frac{\frac{v}{c}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \right) \cdot \tau, \quad x^1 = \frac{v}{c} \cdot x^0, \quad (458)$$

где c - скорость света, а v - скорость материального объекта, соответствующего данной мировой линии, конечно, $|v| < c$. Все такие мировые линии являются лучами, исходят из начала отсчета, особой точки потенциала, и расположены в конусе будущего.

Совместим себя с наблюдателем (мировой линией) $v = 0$ и будем вести отсчет времени начиная с некоторого значения t_0 . Тогда мы будем наблюдать *удаление* материальных объектов от нас с различными скоростями, зависящими от расстояния до этих объектов. Модуль координаты x^1 будет характеризовать нам истинное пространственное расстояние r от наблюдателя до материальных объектов. Из формул (457) и (458) получим

$$|v| \equiv \frac{dr}{dt} = \frac{r}{t_0 + t}, \quad (459)$$

а при $t \ll t_0$ имеем

$$|v| \simeq \frac{1}{t_0} \cdot r. \quad (460)$$

Таким образом, мы получили закон Хаббла с коэффициентом Хаббла $H = \frac{1}{t_0}$, что вполне согласуется с формулой, которая получается для этого закона в космологической модели Эйнштейна - де Ситтера [13], где $H = \frac{2}{3t_0}$.

19 Поличисловое пространство H_4

Гиперкомплексные числа H_4 изоморфны алгебре действительных квадратных диагональных матриц 4×4 . Координаты в изотропном базисе пространства H_4 будем обозначать $\xi^1, \xi^2, \xi^3, \xi^4$, а сам базис - $\psi_1, \psi_2, \psi_3, \psi_4$, то есть любое число можно представить в виде разложения

$$X = \xi^1 \psi_1 + \xi^2 \psi_2 + \xi^3 \psi_3 + \xi^4 \psi_4, \quad (461)$$

или четверки действительных чисел $(\xi^1, \xi^2, \xi^3, \xi^4)$. Бинарная операция поличислового умножения определяется заданием правила умножения для базисных векторов:

$$\psi_i \psi_j = \begin{cases} \psi_i, & \text{если } i = j \\ 0, & \text{если } i \neq j. \end{cases} \quad (462)$$

Если $\xi^1 \xi^2 \xi^3 \xi^4 > 0$, то для числа $X \in H_4$ определена норма

$$|X| = \sqrt[4]{\xi^1 \xi^2 \xi^3 \xi^4} > 0. \quad (463)$$

Пространство H_4 является метрическим финслеровым пространством с метрической функцией

$$L(d\xi) = \sqrt[4]{d\xi^1 d\xi^2 d\xi^3 d\xi^4}, \quad (464)$$

которую никакими преобразованиями координат нельзя привести к квадратному корню из дифференциальной квадратичной формы, то есть такое финслерово пространство качественно отличается от евклидова и псевдоевклидовых пространств размерности четыре. Метрику, которая определяется метрической функцией (464), называют иногда метрикой Бервальда - Моора [17].

В пространстве H_4 существует базис $1, j, k, jk$,

$$\left. \begin{aligned} 1 &= \psi_1 + \psi_2 + \psi_3 + \psi_4, \\ j &= \psi_1 + \psi_2 - \psi_3 - \psi_4, \\ k &= \psi_1 - \psi_2 + \psi_3 - \psi_4, \\ jk &= \psi_1 - \psi_2 - \psi_3 + \psi_4, \end{aligned} \right\} \quad (465)$$

закон умножения базисных элементов которого определен Таб. 5.1.

$\times$	1	j	k	jk
1	1	j	k	jk
j	j	1	jk	0
k	k	jk	1	j
jk	jk	k	j	1

Таб. 5.1

Координаты x^0, x^1, x^2, x^3 числа X в этом базисе связаны с координатами $\xi^1, \xi^2, \xi^3, \xi^4$ того же числа в изотропном базисе формулами:

$$\left. \begin{aligned} \xi^1 &= x^0 + x^1 + x^2 + x^3, \\ \xi^2 &= x^0 + x^1 - x^2 - x^3, \\ \xi^3 &= x^0 - x^1 + x^2 - x^3, \\ \xi^4 &= x^0 - x^1 - x^2 + x^3. \end{aligned} \right\} \quad (466)$$

Базис $1, j, k, jk$ является "ортонормированным" именно в нем в конусе $\xi^1, \xi^2, \xi^3, \xi^4 > 0$ определено экспоненциальное представление чисел $X \in H_4$.

Если гладкая функция $F(x)$ одной действительной переменной x представима как полином или сходящийся ряд переменной x , то функция $F(X)$ переменной $X \in H_4$ также определена и является аналитической, в изотропном базисе она запишется следующим образом:

$$F(X) = F(\xi^1)\psi_1 + F(\xi^2)\psi_2 + F(\xi^3)\psi_3 + F(\xi^4)\psi_4. \quad (467)$$

В "ортонормированном" базисе $1, j, k, jk$ та же функция имеет вид:

$$\begin{aligned} F(X) = & \\ &= \frac{1}{4}[F(x^0 + x^1 + x^2 + x^3) + F(x^0 + x^1 - x^2 - x^3) + \\ &+ F(x^0 - x^1 + x^2 - x^3) + F(x^0 - x^1 - x^2 + x^3)] \cdot 1 + \\ &+ \frac{1}{4}[F(x^0 + x^1 + x^2 + x^3) + F(x^0 + x^1 - x^2 - x^3) - \\ &- F(x^0 - x^1 + x^2 - x^3) - F(x^0 - x^1 - x^2 + x^3)] \cdot j + \\ &+ \frac{1}{4}[F(x^0 + x^1 + x^2 + x^3) - F(x^0 + x^1 - x^2 - x^3) + \\ &+ F(x^0 - x^1 + x^2 - x^3) - F(x^0 - x^1 - x^2 + x^3)] \cdot k + \\ &+ \frac{1}{4}[F(x^0 + x^1 + x^2 + x^3) - F(x^0 + x^1 - x^2 - x^3) - \\ &- F(x^0 - x^1 + x^2 - x^3) + F(x^0 - x^1 - x^2 + x^3)] \cdot jk. \end{aligned} \quad (468)$$

Элемент длины в пространстве, конформно связанном с пространством H_4 , в изотропном базисе имеет вид

$$ds = \kappa(\xi^1, \xi^2, \xi^3, \xi^4) \sqrt[4]{d\xi^1 d\xi^2 d\xi^3 d\xi^4}. \quad (469)$$

Обобщенные импульсы находятся по формулам

$$p_i = \frac{1}{4}\kappa(\xi) \frac{\sqrt[4]{d\xi^1 d\xi^2 d\xi^3 d\xi^4}}{d\xi^i}, \quad (470)$$

а тангенциальное уравнение индикатрисы можно записать, например, так

$$p_1 p_2 p_3 p_4 - \frac{\kappa^4(\xi)}{4^4} = 0. \quad (471)$$

Выберем в качестве функции $F(x)$ в формуле (467) функцию

$$F(x) = a \ln(x/b), \quad (472)$$

где $a, b > 0$ - действительные параметры. Потенциал - это действительная часть аналитической функции $F(X)$ (468), поэтому

$$U(\xi^1, \xi^2, \xi^3, \xi^4) = \frac{a}{4} \ln \left(\frac{\xi^1 \xi^2 \xi^3 \xi^4}{b^4} \right). \quad (473)$$

Для нахождения мировых линий, которые порождает потенциал (473), необходимо решить систему уравнений (439), которая в нашем конкретном случае принимает вид

$$\xi^i = \frac{\frac{\partial U}{\partial \xi^1} \frac{\partial U}{\partial \xi^2} \frac{\partial U}{\partial \xi^3} \frac{\partial U}{\partial \xi^4}}{\frac{\partial U}{\partial \xi^i}} \lambda(\xi), \quad (474)$$

где $\lambda(\xi) \neq 0$ - некоторая скалярная функция. При соответствующем выборе $\lambda(\xi)$, уравнения принимают более простой вид

$$\dot{\xi}^i = \xi^i. \quad (475)$$

Введем переменную

$$x^0 = \frac{1}{4}(\xi^1 + \xi^2 + \xi^3 + \xi^4), \quad (476)$$

которая в четырехмерном пространстве Бервальда - Моора играет ту же роль, что и координата x^0 в пространстве Минковского, тогда

$$\dot{x}^0 = x^0 \quad \Rightarrow \quad \frac{d\xi^i}{dx^0} = \frac{\xi^i}{x^0} \quad \Rightarrow \quad \xi^i = \xi_0^i \cdot x^0, \quad (477)$$

где ξ_0^i - постоянные. Таким образом, все мировые линии - лучи, исходящие из начала координат (особой точки потенциала), причем движение материальных тел является равномерным и прямолинейным, поэтому для координат в "ортонормированном" базисе имеют место следующие формулы:

$$x^\mu = v^\mu x^0, \quad \frac{dx^\mu}{dx^0} = \frac{x^\mu}{x^0}, \quad (478)$$

где $\mu = 1, 2, 3$, а v^μ - действительные постоянные.

Совместим наблюдателя с мировой линией $v^\mu = 0$ и будем вести отсчет времени начиная с некоторого момента t_0 . Считая, что скорость удаления материальных объектов $|v| = \frac{dr}{dt}$, где r - истинное пространственное расстояние, мы будем наблюдать *удаление* от нас материальных объектов с различными скоростями, зависящими от истинного пространственного расстояния до объекта. Предположим, что наблюдатели при обработке экспериментальных данных используют евклидовы расстояния в качестве истинных пространственных расстояний, то есть

$$r \simeq r_{ev} \equiv \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}, \quad (479)$$

и тогда из формул (477) следует обычный закон Хаббла

$$\frac{dr}{dt} \simeq \frac{1}{t_0} \cdot r, \quad (480)$$

и этот закон не зависит от пространственных направлений. Таким образом, закон Хаббла получен с коэффициентом Хаббла $H = \frac{1}{t_0}$, что вполне согласуется с формулой, которая получается для этого закона в космологической модели Эйнштейна - де Ситтера [13], где $H = \frac{2}{3t_0}$.

Если истинное пространственное расстояние в H_4 искать тем же способом, как это обычно делается в СТО и ОТО [11], то в результате получим [15]:

$$r \equiv r_d = |x^\alpha| + |x^\beta|; \quad \text{причем} \quad |x^\gamma| \leq |x^\alpha|, |x^\beta|. \quad (481)$$

Здесь $\alpha, \beta, \gamma = 1, 2, 3$; $\alpha \neq \beta$, $\alpha \neq \gamma$, $\beta \neq \gamma$. В евклидовом пространстве гиперповерхность, определяемая уравнением

$$r_d = \text{const}, \quad (482)$$

является ромбододекаэдром [16]. Если предположить, что истинное трёхмерное пространственное расстояние определяется формулой (481), в этом случае закон Хаббла принимает вид

$$|v| \simeq \frac{1}{t_0} \cdot r_d. \quad (483)$$

Если наши представления о пространстве остались евклидовыми, теоретически полученную формулу для закона Хаббла следует переписать следующим образом:

$$|v| \simeq \frac{r_d}{t_0 r_{ev}} \cdot r_{ev}. \quad (484)$$

Тогда коэффициент Хаббла определяется формулой

$$H = \frac{r_d}{t_0 r_{ev}} \quad (485)$$

и зависит от пространственных направлений, имея 12 максимумов, 6 минимумов и 8 локальных минимумов, причём

$$\frac{H_{max}}{H_{min}} = \sqrt{2}, \quad \frac{H_{max}^{loc}}{H_{min}} = \frac{1 + \sqrt{2}}{2}. \quad (486)$$

Каждый минимум окружен четырьмя максимумами, а каждый локальный минимум - тремя максимумами.

Обе формулы (480) и (484) не будут противоречить друг другу, если предположить, что использовать геометрию Бервальда - Моора следует только на расстояниях "близким к размерам" Вселенной. Тогда при изменении расстояния до галактик от самых малых до самых больших, коэффициент Хаббла в логарифмической модели претерпевает изменения:

$$\frac{1}{t_0} \xrightarrow{H} \frac{r_d}{t_0 r_{ev}}, \quad (487)$$

причём появляется зависимость этого коэффициента от пространственных направлений.

Таким образом, дополнительная анизотропия в четырёхмерном пространстве-времени приводит к анизотропии в истинном трёхмерном пространстве. Астрономические данные [22] указывают на наличие анизотропии коэффициента Хаббла и позволяют иначе взглянуть и на другое устоявшееся и, по-видимому, ошибочное убеждение, а именно, на уверенность большинства физиков в применимости к релятивистским моделям реального физического Мира исключительно квадратичных метрик. Пространства же невырожденных поличисел P_n при $n > 2$ всегда обладают дополнительной анизотропией по сравнению с квадратичными пространствами той же размерности.

Список литературы

- [1] Рашевский П.К. Геометрическая теория уравнений с частными производными. М. - Л., ОГИЗ, 1947.
- [2] Рашевский П.К. Риманова геометрия и тензорный анализ. М., "Наука" 1967.

- [3] Кантор И.Л., Солодовников А.С., Гиперкомплексные числа. М., "Наука 1973.
- [4] Свешников А.Г., Тихонов А.Н. Теория функций комплексной переменной. М., 1967.
- [5] Эльсольц Л.Э. Дифференциальные уравнения и вариационное исчисление. М., "Наука 1969.
- [6] Гельфанд И.М. Лекции по линейной алгебре. М., "Наука 1971.
- [7] Нетер Э. Инвариантные вариационные задачи: Вариационные принципы механики. Сб. статей под ред. Полака Л.С., М., "ГИФМЛ 1959, с.611 - 630.
- [8] Гельфанд И.М., Фомин С.В. Вариационное исчисление. М., "Госиздат. физико-математической литературы 1961.
- [9] Будаков Б.М., Фомин С.В. Кратные интегралы и ряды. М., "Наука 1967.
- [10] Ландау Л.Д., Лифшиц Е.М. Механика. М., "Наука 1965.
- [11] Ландау Л.Д., Лифшиц Е.М. Теория поля. М., "Наука 1967.
- [12] Боголюбов Н.Н, Ширков Д.В. Введение в теорию квантованных полей. М., "Наука 1973.
- [13] Мёллер К. Теория относительности. М., АТОМИЗДАТ, 1975.
- [14] Богословский Г.Ю. Статус и перспективы теории локально анизотропного пространства-времени, Физика ядра и частиц. Под ред. Б.С.Ишханова и др. М., Изд-во МГУ, 1997.
- [15] Bogoslovsky G.Yu., Goenner H.F. On the possibility of phase transitions in the geometric structure of space-time // Phys. Lett. A 244, 1998, pp.222–228.
- [16] Bogoslovsky G.Yu., Goenner H.F. Finslerian spaces possessing local relativistic symmetry // Gen. Relativ. Gravit. 31, 1999, pp.1565–1603.
- [17] Богословский Г.Ю. 4-импульс частицы и уравнение массовой поверхности в полностью анизотропном пространстве-времени // Гиперкомплексные числа в геометрии и физике, 2(4), 2005, с.27 - 43.
- [18] Павлов Д.Г. Обобщение аксиом скалярного произведения // Гиперкомплексные числа в геометрии и физике, 1(1), 2004, с.5 - 19.
- [19] Павлов Д.Г. Хронометрия трёхмерного времени // Гиперкомплексные числа в геометрии и физике, 1(1), 2004, с.20 - 32.
- [20] Павлов Д.Г. Четырёхмерное время // Гиперкомплексные числа в геометрии и физике, 1(1), 2004, с.33 - 42.
- [21] Гарасько Г.И. Теория поля и финслеровы пространства // Гиперкомплексные числа в геометрии и физике, 2(6), Т.3, 2006, с.6 - 20.
- [22] McClure M.L., Dyer C.C. Anisotropy in the Hubble constant as observed in the HST Extragalactic Distance Scale Key Project results. arXiv:astro-ph/0703556v1

Элементы геометрии гладких многообразий (I): производные Ли и их приложения

С.С. Кокарев

Институт гиперкомплексных систем в геометрии и физике, Фрязино, Россия

РНОЦ “Логос”, Ярославль, Россия.

E-mail: logos-center@mail.ru

В лекциях систематически излагаются основы теории гладких многообразий, необходимые для обоснования аппарата производных Ли. В качестве приложений этого аппарата рассмотрены иллюстративные примеры из геометрической теории дифференциальных уравнений и техника отыскания изометрий и конформных симметрий на примере псевдоевклидовых и некоторых простых римановых метрик. Особое внимание уделяется подробному исследованию изометрий однородных финслеровых кубических метрик и их инвариантной классификации.

1 Введение и обозначения

Со времени создания общей теории относительности (1915) идея геометрического описания законов природы получила мощный импульс для своего развития и послужила основанием для формулировки и разработки подавляющего большинства современных физико-геометрических концепций и теорий [1, 2]. В современной геометрии общей основой всех геометрических построений является понятие *многообразия*, которое представляет собой глубокое обобщение наших представлений о поверхностях, пространстве и времени [3]. В физике многообразия наделяются дополнительными структурами, отвечающими за основные свойства пространства, времени и материи. При этом физические величины описываются *тензорными полями* на многообразии.

Одним из фундаментальных свойств окружающего нас мира является явная или скрытая *симметрия* его законов. С одной стороны, принципы симметрии помогают нам формулировать эти законы, с другой — они являются ключом к точному аналитическому исследованию уравнений, выражающих эти законы. Общепринятым (хотя, конечно, далеко не исчерпывающим) математическим понятием, отвечающим за описание симметрии, является понятие *группы* [4]. Элементами группы симметрий в физике являются такие преобразования физической системы, которые оставляют ее, в определенном смысле, неизменной. К примеру все воображаемые движения твердого стержня, при которых остается неизменной его длина, образуют *группу 3-мерных вращений*, состоящую из параллельных переносов, собственных вращений и инверсий. При этом первые два типа преобразований относятся к числу *непрерывных симметрий*, поскольку такие преобразования могут быть осуществлены путем непрерывного изменения параметров преобразования (углов и параметров трансляций), начиная от их нулевых значений (т.е. тождественного преобразования). Преобразования инверсии представляют пример *дискретных преобразований*. Соединение идеи непрерывной симметрии с идеей многообразия приводит к одному из центральных понятий современной физики и математики — *группы Ли*. Многообразие группы Ли состоит из точек, которые

можно интерпретировать как преобразования, а кривые на этом многообразии являются геометрическим образом однопараметрических семейств преобразований.

Когда мы хотим математически выразить факт постоянства некоторой физической величины F во времени, мы записываем условие этого постоянства в виде:

$$\frac{dF}{dt} = 0 \quad (1)$$

— равенства нулю производной по времени от этой величины. Когда мы хотим выразить постоянство этой величины в пространстве, мы записываем условие этого постоянства в аналогичном виде:

$$\frac{dF}{dx} = 0 \quad (2)$$

и т.д. В физике и геометрии часто возникают ситуации, когда физические или геометрические величины (тензоры) не остаются постоянными в пространстве и (или) во времени, но остаются постоянными при смещении вдоль некоторого особого семейства кривых на многообразии положений. Именно такую ситуацию мы и называем в физике *непрерывной симметрией физической* или геометрической величины. Надлежащее обобщение формул (1)-(2) реализуется с помощью общей конструкции *производной Ли*, определение которой не связано с существованием каких-либо дополнительных структур на многообразии, кроме общей структуры гладкости (см. след. раздел). При этом производная Ли согласована с тензорной алгеброй на многообразии: производная Ли от тензора является тензором того же типа.

Целью настоящих лекций является введение в аппарат производных Ли и небольшой обзор его применений в геометрии и физике, нацеленный на начальное ознакомление. Основное внимание в лекциях уделяется геометрическим аспектам производной Ли и ее конкретным приложениям в геометрии и физике, в то время как теоретико-групповые вопросы обсуждаются лишь в общих чертах. Везде, где это возможно и уместно, предпочтение отдается безкоординатным определениям и формулировкам. Иногда на первых этапах это требует несколько больших усилий в процессе освоения материала, но эти усилия "купаются" общностью формулировок и четким выделением инвариантных геометрических аспектов определяемых объектов и конструкций. Среди приложений относительно большой объем в лекциях занимает исследование изометрий финслеровых кубических метрик. Этот вопрос весьма слабо освещен в доступной литературе. Кроме того, общий контекст Школы-2009, на которой часть этих лекций озвучивалась автором у доски, подразумевал популяризацию идей финслеровой геометрии и ее приложений к физическим проблемам.

Лекции представляют достаточно замкнутое и систематическое введение и обоснование производной Ли и ряда ее приложений. Вместе с тем, ряд важных вопросов (исчисление внешних форм и теория связности на многообразии, "взаимодействие" внешнего дифференцирования и ковариантной производной с производной Ли и ряд других) остается за пределами лекций из соображений компромисса между объемом и строгостью изложения. Эти разделы выносятся в следующие части лекций. Читатель, заинтересованный в более полном изложении, может обратиться к множеству классических руководств и монографий [5–12].

На протяжении всех лекций мы принимаем следующую систему обозначений:

R^n — n -мерное вещественное евклидово пространство;

$\partial_i \equiv \frac{\partial}{\partial x^i}$ — оператор частного дифференцирования по переменной x^i ;

$f \circ g$ — композиция отображений g и f (первым действует отображение справа);

id — тождественное отображение.

$\{a_i\}_{i=1,\dots,n}$ — множество занумерованных объектов a_i , где индекс i пробегает все значения от 1 до n .

$T_p = T|_p = T(p)$ — эквивалентные способы обозначения геометрического объекта (тензора T) в точке p .

(T) — матрица компонент объекта T .

Везде, где это не оговаривается особо, принято правило сокращенного суммирования Эйнштейна по повторяющимся верхним и нижним индексам.

Остальные обозначения поясняются непосредственно в тексте.

Рубленным шрифтом набраны основные определения, *наклонным "италиком"* — основные термины и формулировки выводов, *машинописным* — формулировки доказываемых утверждений и теорем. Окончание доказательств и примеров отмечается значком $\square$.

Успешное изучение лекций предполагает уверенное знание читателем математического анализа, линейной алгебры и аналитической геометрии в объеме стандартного университетского курса [13–15].

2 Гладкие многообразия

Основная идея, лежащая в определении гладкого многообразия, проста: оно "склеено" из "лоскутков", каждый из которых "похож" на область евклидова пространства некоторого фиксированного (конечного) числа измерений. Чтобы придать точный смысл словам, заключенным в кавычки, перейдем к более строгим определениям.

Рассмотрим некоторое множество $\mathcal{M}$. Пара (U, φ) , где U — некоторое открытое* подмножество $\mathcal{M}$, а $\varphi: U \rightarrow D \subset R^m$ — биективное отображение U на некоторое открытое множество (область) D вещественного евклидова пространства R^m , называется *координатной картой* на $\mathcal{M}$. При этом U называется *координатной окрестностью*, φ — *координатным (картирующим) отображением*, а R^m — *арифметизирующим пространством*. Пара карт (U_1, φ_1) и (U_2, φ_2) называется C^r -согласованной, если выполняется одно из двух условий:

1. $U_1 \cap U_2 = \emptyset$;
2. На непустом пересечении карт $U_{12} = U_1 \cap U_2$ отображения $\psi_{12} = \varphi_2 \circ \varphi_1^{-1}: D'_1 = \varphi_1(U_{12}) \rightarrow D'_2 = \varphi_2(U_{12})$ и $\psi_{21} = \varphi_1 \circ \varphi_2^{-1}: D'_2 \rightarrow D'_1$ являются C^r -гладкими отображениями открытых множеств.

Рисунок 1 поясняет второй пункт этого определения. Поясним условие C^r гладкости на координатном языке. Пусть $(x^1, \dots, x^m) \equiv x$ — система координат на R^m для картирующего отображения φ_1 , а $(y^1, \dots, y^m) \equiv y$ — система координат на R^m для картирующего отображения φ_2 . Арифметизирующие пространства для картирующих отображений различных карт удобно считать различными. Таким образом, отображение φ_1 каждой точке $p_1 \in U_1$ ставит в соответствие точку $x_{p_1} = (x^1(p_1), \dots, x^m(p_1)) \in D_1 \subset R^m$, а отображение φ_2 каждой

*Существование открытых подмножеств на $\mathcal{M}$ подразумевает, что $\mathcal{M}$ — топологическое пространство. На самом деле, топологию на $\mathcal{M}$ всегда можно выбрать индуцированной евклидовой топологией на R^m , потребовав, чтобы отображения φ в любой карте были непрерывными. Возможно, что такая точка зрения не является общепринятой и даже логически последовательной, но в настоящих лекциях мы придерживаемся ее из соображений экономии места и времени. Такой же подход используется автором в курсе лекций [6]. Более детальное обсуждение топологических аспектов многообразий можно найти в [16].

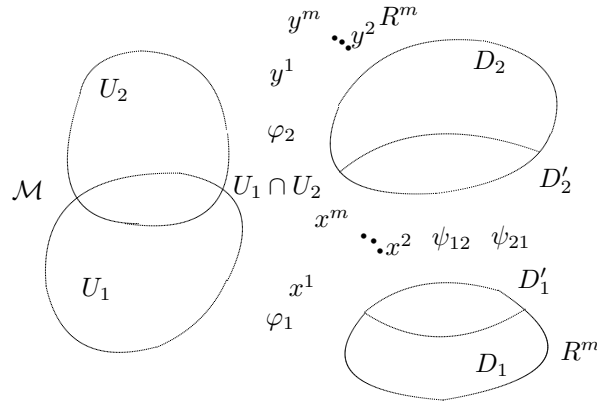


Рис. 1: К условию согласования карт.

точке $p_2 \in U_2$ ставит в соответствие точку $y_{p_2} = (y^1(p_2), \dots, y^m(p_2)) \in D_2 \subset R^m$. Числа $x^i(p_1)$ и $y^i(p_2)$ называются i -ыми координатами точек p_1 и p_2 в координатных картах (U_1, φ_1) и (U_2, φ_2) соответственно. Если пересечение U_{12} непусто, то любая точка q из него будет иметь, как минимум, два координатных представления: $x_q = (x^1(q), \dots, x^m(q))$ и $y_q = (y^1(q), \dots, y^m(q))$, которые обязательно будут взаимно-однозначно связаны друг другом, ввиду того, что оба представления порождаются одной и той же точкой и отображения φ_1 и φ_2 являются биекциями. Отображение ψ_{12} (оно называется *функцией перехода из U_1 в U_2*) как раз и является числовой функцией, переводящей координаты x в координаты y для каждой точки из U_{12} , а отображение ψ_{21} осуществляет обратное преобразование y в x . В явном виде эти отображения описываются системами:

$$y^i = f^i(x^1, \dots, x^m); \quad x^i = g^i(y^1, \dots, y^m), \quad (i = 1, \dots, m).$$

При этом $f^i \equiv \psi_{12}^i$ и $g^i \equiv \psi_{21}^i$ обозначают i -ые компоненты отображений ψ_{12} и ψ_{21} соответственно. Требование C^r гладкости означает, что все функции f^i и g^i имеют непрерывные частные производные вплоть до r -ого порядка включительно.

Атласом $\mathbf{At}(\mathcal{M})$ на множестве $\mathcal{M}$ называется совокупность попарно C^r -согласованных карт $(U_\alpha, \varphi_\alpha)$, таких, что система координатных окрестностей $\{U_\alpha\}$ образует покрытие множества $\mathcal{M}$: $\mathcal{M} = \cup_\alpha U_\alpha$. Два атласа $\mathbf{At}_1$ и $\mathbf{At}_2$ называются эквивалентными: $\mathbf{At}_1 \sim \mathbf{At}_2$, если любая пара карт из $\mathbf{At}_1$ и $\mathbf{At}_2$ является C^r -согласованной. Говорят, что класс эквивалентных атласов определяет на $\mathcal{M}$ *гладкую структуру* или *гладкость*.

Множество $\mathcal{M}$ с введенной на нем гладкой структурой называется *гладким многообразием*. При этом *размерностью* $\dim \mathcal{M}$ гладкого многообразия называется размерность m арифметирующего пространства R^m .

Размерность многообразия является его важнейшей характеристикой и иногда ее указывают явно в виде верхнего или нижнего индекса у символа многообразия: например $\mathcal{M}_m$, $\dim \mathcal{M}_m = m$. В настоящих лекциях мы всегда будем иметь дело с многообразиями, у которых гладкость имеет бесконечный порядок: $r = \infty$, и всегда $\dim \mathcal{M} = m$ и $\dim \mathcal{N} = n$, так что для сокращения записи мы не будем указывать размерности абстрактных многообразий $\mathcal{M}$ и $\mathcal{N}$, используемых в конструкциях общего характера.

Происхождение терминов "карта" и "атлас" очевидно. Поверхность Земли мы изображаем посредством совокупности плоских карт, каждая из которых, покрывает определенный участок земной поверхности. Для представления всей поверхности несколько карт объединяются в географический атлас. При этом края некоторых пар карт изображают один и тот же участок Земли и поэтому должны существовать правила, переводящие точки с одной карты на таком участке, на другую карту на нем же и обратно. Теория многообразий заимствует эти идеи из картографии и, абстрагируясь от конкретных особенностей наглядного представления, переносит их на общий случай m -измерений.

Рассмотрим примеры многообразий.

1. Евклидово пространство R^n . Самый простой способ ввести гладкость на R^n заключается в задании одной единственной карты (R^n, id) . Она называется *стандартной гладкостью*. Отметим, не углубляясь в детали, что на R^n можно ввести и другие гладкие структуры, неэквивалентные только что введенной нами [17].

2. Комплексное пространство C^n . Представление о веществе: $C^n = R^n + iR^n$ подсказывает, что простейший способ ввести гладкую структуру на C^n заключается в переходе от C^n к R^{2n} и дословному повторению конструкции предыдущего пункта. Такая гладкость, однако, "стирает" всякую информацию о комплексно-алгебраической структуре C^n . Для ее сохранения необходимо обобщить понятие вещественного многообразия на понятие *комплексного многообразия*, у которого роль арифметизирующего пространства выступает C^n , а функции перехода становятся комплексно-аналитическими [10].

3. Линейное вещественное векторное пространство L_n . В любом n -мерном линейном вещественном пространстве можно ввести базис, состоящий из n элементов $\{e_1, \dots, e_n\}$. Всякий вектор v представляется в таком базисе в виде линейной комбинации вида $v^1 e_1 + \dots + v^n e_n$, где $\{v^1, \dots, v^n\}$ — координаты v в выбранном базисе. Таким образом, при некотором фиксированном базисе L_n оказывается изоморфным R^n и, следовательно, допускает структуру n -мерного вещественного гладкого многообразия.

4. Поверхности в R^n . Стандартные определения линии и поверхности в R^3 представляют собой типичные примеры многообразий малой размерности [6]. Не представляет труда обобщить эти определения на поверхности высших размерностей в R^n . При этом, как правило в качестве единственной карты выступает параметризация линий и поверхностей. Следует отметить, что несмотря на общие теоремы о возможности реализации многообразий в виде вложенных поверхностей в объемлющее евклидово пространство достаточно большого числа измерений [6], общая формулировка понятия многообразия никак не связана с такого рода представлениями и изучение свойств многообразия не требует привлечения понятий и структур какой-либо объемлющей геометрии.

4. Сфера S^n . Сфера относится к числу простейших (но нетривиальных!) многообразий. Начнем рассмотрение сфер с простейшего случая $n = 1$ (окружность). Окружность S^1 определяется как подмножество точек плоскости R^2 , удовлетворяющих уравнению*:

$$(x^1)^2 + (x^2)^2 = 1,$$

*Выбором масштаба длины радиус окружности можно всегда сделать единичным. Сфера произвольного радиуса как многообразие принципиально ничем не отличается от единичной сферы, поэтому в дальнейшем мы для некоторого упрощения будем ограничиваться именно единичной сферой.

где (x^1, x^2) — декартовы координаты на плоскости*. Введем следующие обозначения:

$$S_{i\pm}^1 = \{p \in S^1 | x^i \gtrless 0\}, \quad i = 1, 2.$$

Атлас сферы S^1 теперь можно задать посредством четырех карт $(S_{i\pm}^1, \varphi_{i\pm})$, где $\varphi_{i\pm} = \pi_i$, а π_i — отображение проекции на прямую $x^i = 0$: $\pi_1(x^1, x^2) = (0, x^2)$, $\pi_2(x^1, x^2) = (x^1, 0)$. При этом координатные окрестности S_{i+}^1 и S_{i-}^1 не пересекаются, а на не пустых пересечениях $S_{1+}^1 \cap S_{2+}^1$, $S_{1+}^1 \cap S_{2-}^1$, $S_{1-}^1 \cap S_{2+}^1$, $S_{1-}^1 \cap S_{2-}^1$, функции перехода имеют, соответственно, вид:

$$x^2 = +\sqrt{1 - (x^1)^2}; \quad x^2 = -\sqrt{1 - (x^1)^2}; \quad x^2 = +\sqrt{1 - (x^1)^2}; \quad x^2 = -\sqrt{1 - (x^1)^2}$$

и обратные к ним (с учетом знаков у корней). Непосредственной проверкой нетрудно убедиться, что эти функции перехода на всех пересечениях гладкие с $r = \infty$. Следовательно S^1 — одномерное гладкое многообразие.

Построенная конструкция элементарно обобщается на случай 2-мерной сферы S^2 — соответствующие конкретные формулы читателю предлагается записать самостоятельно (подсказка: для S^2 мы будем иметь шесть карт). Мы сразу перейдем к построению атласа n -мерной сферы S^n , рассматриваемой как подмногообразие R^{n+1} , задаваемое уравнением:

$$(x^1)^2 + (x^2)^2 + \dots + (x^{n+1})^2 = 1.$$

Введем обозначения:

$$S_{i\pm}^n = \{p \in S^1 | x^i \gtrless 0\}, \quad i = 1, \dots, n.$$

Атлас сферы S^n теперь можно задать посредством $2(n+1)$ карт $(S_{i\pm}^n, \varphi_{i\pm})$, где $\varphi_{i\pm} = \pi_i$, а π_i — отображение проекции на гиперплоскость $x^i = 0$:

$$\pi_i(x^1, x^2, \dots, x^i, \dots, x^{n+1}) = (x^1, x^2, \dots, 0, \dots, x^n), \quad i = 1, \dots, n.$$

При этом координатные окрестности S_{i+}^n и S_{i-}^n не пересекаются, а не пустыми будут пересечения $S_{i\pm}^n \cap S_{j\pm}^n$ при $i \neq j$. Функции перехода, к примеру, для $S_{i+}^n \cap S_{j+}^n$ и $S_{i-}^n \cap S_{j+}^n$ при $1 < i < j < n+1$ имеют вид:

$$x'^1 = x^1, \dots, x'^{i-1} = x^{i-1}, \quad x'^i = \sqrt{1 - (x^1)^2 - \dots - (x^{i-1})^2 - (x^{i+1})^2 - \dots - (x^{n+1})^2},$$

$$x'^{i+1} = x^{i+1}, \dots, x'^j = x^{j+1}, \quad \dots x'^n = x^{n+1}.$$

Аналогичный вид имеют все остальные функции перехода, а также и обратные к ним. Непосредственной проверкой можно убедиться, что все функции перехода являются гладкими на пересечениях, поэтому S^n — n -мерное гладкое многообразие. Отметим, что с помощью отображения стереографической проекции можно построить более экономные атласы для всех сфер S^n , состоящие при любом n всего из двух карт [6].

Прямые произведения $\mathcal{M}_m \times \mathcal{N}_n$. Напомним, что прямым (или декартовым) произведением $\mathcal{M} \times \mathcal{N}$ двух множеств $\mathcal{M}$ и $\mathcal{N}$ называется множество всевозможных упорядоченных пар, (m, n) , где $m \in \mathcal{M}$, а $n \in \mathcal{N}$. Если $\mathcal{M}_m$ и $\mathcal{N}_n$ — гладкие многообразия,

*Для дальнейшего обобщения конструкции на многомерный случай нам удобнее перейти от привычного обозначения декартовых координат (x, y) к "безличной" нумерации (x^1, x^2) .

то на прямом произведении $\mathcal{M}_m \times \mathcal{N}_n$ гладкость порождается прямым произведением атласов $\mathbf{At}(\mathcal{M}_m) \times \mathbf{At}(\mathcal{N}_n)$, состоящим из всевозможных карт вида $(U_\alpha \times V_\beta, \varphi_\alpha \times \chi_\beta)$, где $(U_\alpha, \varphi_\alpha) \in \mathbf{At}(\mathcal{M}_m)$, $(V_\beta, \varphi_\beta) \in \mathbf{At}(\mathcal{N}_n)$. При этом прямое произведение отображений $\varphi_\alpha \times \chi_\beta$ действует по правилу:

$$(\varphi_\alpha \times \chi_\beta)(m, n) = (\varphi_\alpha(m), \chi_\beta(n)),$$

где $m \in \mathcal{M}_m$, $n \in \mathcal{N}_n$, $\varphi_\alpha(m) \in R^m$, $\chi_\beta(n) \in R^n$. Таким образом, прямое произведение двух гладких многообразий — это снова гладкое многообразие*. При этом, очевидно, $\dim(\mathcal{M}_m \times \mathcal{N}_n) = m + n$. В качестве примеров многообразий, имеющих структуру прямого произведения гладких многообразий можно привести *тор*: $T^2 = S^1 \times S^1$ и *цилиндр* $\text{Cyl}^2 = R \times S^1$.

3 Скалярные функции и векторные поля

Гладкая структура на многообразии позволяет ввести на нем элементарные геометрические объекты: гладкие скалярные и векторные поля. Поскольку при этом определение вектора на многообразии, вообще говоря, не допускает привычной интерпретации упорядоченной пары точек, нам потребуется ввести ряд новых предварительных понятий и определений, относящихся к некоторым фундаментальным структурам любого гладкого многообразия.

Скалярной функцией на многообразии $\mathcal{M}$ называется отображение $f: \mathcal{M} \rightarrow R$.

Таким образом, как обычно, скалярная функция f каждой точке $p \in \mathcal{M}$ ставит в соответствие вещественное число $f(p) \in R$. Система координат в каждой карте атласа $\mathbf{At}(\mathcal{M})$ позволяет перейти к координатному представлению функции f , которое в некоторой карте (U, φ) задается по правилу:

$$f_U(p) \equiv f \circ (\varphi^{-1} \circ \varphi)(p) = (f \circ \varphi^{-1})(x_p).$$

Таким образом, всякая скалярная функция на многообразии $\mathcal{M}_m$ задается набором своих координатных компонент $\{f_{U_\alpha}\}_\alpha$, где $f_{U_\alpha}(x_\alpha)$ — компонента в карте $(U_\alpha, \varphi_\alpha)$. При этом скалярная функция f называется *гладкой*, если ее координатные компоненты гладкие в обычном смысле вещественного анализа. Ясно, что порядок гладкости функции, определенный таким образом, не может превышать порядка гладкости многообразия, поскольку только при таком условии этот порядок сохраняется при замене координат на пересечениях карт.

Очевидно, что сумма и произведение любых двух гладких функций на $\mathcal{M}$ являются гладкой функцией на $\mathcal{M}$, поэтому совокупность всех гладких функций на многообразии образуют алгебраическую структуру *кольца*, которое мы будем обозначать $\mathfrak{F}(\mathcal{M})$.

Прежде, чем перейти к определению вектора и векторного поля на многообразии, сделаем два предварительных замечания. Даже в евклидовом пространстве вектор можно ассоциировать с упорядоченной парой точек только если ограничиться линейными (*аффинными*) преобразованиями координат. При таких преобразованиях разность декартовых (аффинных) координат преобразуется линейно и однородно, как это и требуется для компонент векторов. Нелинейные преобразования (например, переход к цилиндрической системе координат или любой другой криволинейной) нарушает линейность преобразований разности координат, и следовательно обнаруживают ограниченность ее "векторной природы". Для построения "настоящих" векторов, которые являются таковыми независимо от выбора системы координат,

*Его порядок гладкости определяется как $\min(r_1, r_2)$, где r_1, r_2 — порядки гладкости сомножителей.

мы сделаем следующее полезное наблюдение: с каждым векторным полем v в евклидовом пространстве можно ассоциировать линейный оператор "дифференцирования вдоль v который на любую скалярную функцию действует по правилу:

$$v(f) \equiv \lim_{\lambda \rightarrow 0} \frac{f(x + \lambda v) - f(x)}{\lambda} = \sum_{i=1}^3 v^i \frac{\partial f}{\partial x^i}.$$

При этом, как нетрудно видеть, соответствие между векторными полями и линейными дифференциальными операторами является взаимно-однозначным. Оказывается, именно такая точка зрения на векторные поля, как на линейные дифференциальные операторы, оказывается плодотворной для корректного перенесения определения векторных полей на многообразия.

Определим векторное поле X на M как R -линейное дифференцирование кольца $\mathfrak{F}(M)$, т.е. как отображение $X: \mathfrak{F}(M) \rightarrow \mathfrak{F}(M)$, удовлетворяющее свойствам:

1. $X(fg) = X(f)g + fX(g)$ (*правило Лейбница*);
2. $X(\lambda f + \mu g) = \lambda X(f) + \mu X(g)$, где f и g — произвольные элементы кольца $\mathfrak{F}(M)$, λ и μ — произвольные вещественные числа.
3. Суммой векторных полей $X + Y$ назовем векторное поле, которое действует на произвольную функцию f по правилу: $(X + Y)(f) = X(f) + Y(f)$.
4. Умножением векторного поля X на функцию g (слева) назовем векторное поле gX , которое на произвольную функцию f действует по правилу: $(gX)(f) = g \cdot X(f)$.

Оказывается, что перечисленных свойств и определений достаточно для того, чтобы установить следующие основные свойства векторных полей на многообразии.

Теорема.

1. $X(C) = 0$ для любой постоянной функции C ;
2. Векторные поля $(\partial_1, \dots, \partial_m)$ образуют базис векторных полей в каждой карте атласа $\text{At}(M)$.

Доказательство. Первое свойство вытекает из свойства R -линейности, правила Лейбница и следующей выкладки:

$$X(C) = CX(1) = CX(1 \cdot 1) = C(1X(1) + X(1)1), X(1) = 2X(C) \Rightarrow X(C) = 0.$$

Чтобы доказать второе свойство, мы должны доказать, что (а) векторные поля ∂_i линейно-независимы и (б) любое векторное поле X представляется в виде линейной комбинации

$$X = \sum_{i=1}^m X^i \frac{\partial}{\partial x^i},$$

где X^i — некоторые коэффициенты (функции на M), однозначно связанные с X . Отметим сначала, что векторное поле ∂_i по определению действует на функцию f в каждой точке p из координатной окрестности U по правилу:

$$\partial_i(f)(x_p) = \left. \frac{\partial f}{\partial x^i} \right|_{x=x_p},$$

где $x = (x^1, \dots, x^m)$ — система координат на U . Докажем, что поля ∂_i линейно-независимы. Пусть существует линейная комбинация $\alpha^1 \partial_1 + \dots + \alpha^m \partial_m$, которая на всех функциях обращается в нуль. Подействуем этой комбинацией поочередно на координатные функции $f^1 = x^1$, $f^2 = x^2, \dots, f^m = x^m$. Получим последовательно: $\alpha^1 = 0, \dots, \alpha^m = 0$. Следовательно векторные поля $(\partial_1, \dots, \partial_m)$ линейно-независимы. Для доказательства второй части утверждения отметим, что для всякой функции f с непрерывными частными производными имеет место следующее ее координатное представление в каждой точке некоторой карты:

$$f(x) = f(x_p) + \sum_{i=1}^m A_i(x^i - x_p^i), \quad (3)$$

где коэффициенты A_i зависят от x и x_p и при $x = x_p$ равны частным производным $\partial_i f|_p$. Выражение для A_i имеет следующий явный вид:

$$A_i = \int_0^1 \frac{\partial f}{\partial x^i}((1-t)x_p + tx) dt$$

и является, по сути, средним значением i -ой частной производной, вычисленным на прямолинейном отрезке, соединяющем точки x_p и x координатного пространства. Рассмотрим некоторое векторное поле X и положим по определению $X(x^i) \equiv X^i$. Составим комбинацию $\sum_{i=1}^m X^i \partial_i$. Для того, чтобы доказать равенство $X = X^i \partial_i$ необходимо и достаточно убедиться, что левая и правая часть этого равенства приводит к одному и тому же результату при действии на произвольную функцию f . Используя представление (3) и действуя на его левую часть векторным полем X в точке p , имеем цепочку равенств:

$$\begin{aligned} X(f)(p) &= X(f(x_p) + \sum_{i=1}^m A_i(x^i - x_p^i))|_{x=x_p} = \sum_{i=1}^m [X(A_i)(x^i - x_p^i) + A^i X(x^i - x_p^i)]|_{x=x_p} \\ &= \sum_{i=1}^m X_p^i \frac{\partial f}{\partial x^i} \Big|_{x=x_p}. \end{aligned}$$

Кроме представления (3) мы использовали последовательно R -линейность, правило действия на постоянную функцию, правило Лейбница и свойства коэффициентов A^i . Действуя на f векторным полем вида $X^i \partial_i$ в точке p очевидно получаем то же самое. Ввиду произвольности выбора точки p можно утверждать, что в каждой координатной карте семейство координатных векторных полей $\{\partial_i\}_{i=1, \dots, m}$ образует базис для векторных полей. Коэффициенты $X^i \equiv X(x^i)$ называются *координатами векторного поля в координатном базисе* $\{\partial_i\}$, связанным с некоторой картой U . $\square$

Векторное поле X на многообразии $\mathcal{M}$ называется *гладким*, если при действии на любую гладкую функцию результат будет гладкой функцией. Переходя к координатам, нетрудно убедиться, что это определение эквивалентно следующему: векторное поле X на многообразии $\mathcal{M}$ называется гладким, если в каждой карте $\mathbf{At}(\mathcal{M})$ его координаты X^i являются гладкими функциями. Операция суммы двух гладких векторных полей и умножения векторного поля на гладкую функцию, очевидно, не выводят из класса гладких векторных полей. Это означает, что гладкие векторные поля образуют $\mathfrak{F}(\mathcal{M})$ -модуль, который мы будем обозначать $\mathfrak{V}(\mathcal{M})$.

Для любой пары гладких векторных полей X и Y можно определить их *скобку Ли* $[X, Y] = XY - YX$. Чтобы получить координатное представление скобки Ли, правую часть этого определения нужно применить к произвольной гладкой функции $f \in \mathfrak{F}(\mathcal{M})$. В результате получим цепочку равенств:

$$(XY - YX)(f) = X^i \partial_i (Y^j \partial_j f) - Y^j \partial_j (X^i \partial_i f) = X^i \partial_i (Y^j \partial_j f) - (Y^j \partial_j X^i) \partial_i f = [X, Y](f),$$

где была учтена перестановочность вторых частных производных. В силу произвольности функции f , приходим к выводу о том, что скобка Ли определяет векторное поле (дифференцирование) с координатами:

$$[X, Y]^i = X^j \partial_j Y^i - Y^j \partial_j X^i. \quad (4)$$

Прямой проверкой нетрудно убедиться, что скобка Ли обладает следующими важными свойствами:

1. $[X, Y] = -[Y, X]$ (антисимметричность);
2. $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$ (R -линейность);
3. $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (тождество Якоби).

Здесь X, Y, Z — произвольные векторные поля, α и β — произвольные вещественные числа. Важные для приложений геометрические свойства скобки Ли мы рассмотрим в разделе 9.

4 Касательное и кокасательное пространства. 1-формы.

Рассмотрим множество всех векторов, отнесенных к некоторой фиксированной точке $p \in \mathcal{M}$. Это множество образует m -мерное вещественное линейное пространство, $T_p \mathcal{M}$, причем в качестве его базиса можно выбрать систему координатных полей $\{\partial_i|_p\}$, отнесенных к точке p .

Пространство $T_p \mathcal{M}$ называется касательным к многообразию $\mathcal{M}$ пространством в точке p .

Рассмотрим теперь множество всех линейных отображений $T_p \mathcal{M} \rightarrow R$, элементами которого являются линейные вещественнозначные функции на векторах из $T_p \mathcal{M}$. Каждое линейное отображение ω_p из этого множества определяется своим действием на векторы:

$$\omega_p(X_p) = \lambda \in R, \quad (5)$$

при этом число λ называется *значением ω_p на векторе X_p* . Индекс p напоминает нам, что наше рассмотрение относится к фиксированной точке p . Свойство линейности ω_p выражается равенством:

$$\omega_p(\alpha X_p + \beta Y_p) = \alpha \omega_p(X_p) + \beta \omega_p(Y_p), \quad (6)$$

которое должно выполняться для всяких векторов X_p и Y_p . Равенства

$$(\alpha(\omega_1)_p + \beta(\omega_2)_p)(X_p) = \alpha(\omega_1)_p(X_p) + \beta(\omega_2)_p(X_p) \quad (7)$$

задают на множестве таких линейных функций структуру линейного пространства.

Линейное пространство всех линейных отображений $T_p \mathcal{M} \rightarrow R$ называется кокасательным к многообразию $\mathcal{M}$ пространством в точке p и обозначается $T_p^* \mathcal{M}$, а его элементы называются 1-формами в точке p .

Каковы размерность и базис пространства $T_p^*\mathcal{M}$? Рассмотрим 1-формы $\tilde{x}_p^i$ ($i = 1, \dots, m$) которые действуют на любой вектор X_p по следующему правилу:

$$\tilde{x}_p^i(X_p) = X_p^i.$$

Другими словами 1-форма $\tilde{x}_p^i$ каждому вектору ставит в соответствие его i -ую координату. Из этого определения вытекает правило действия этих 1-форм на координатный векторный базис:

$$\tilde{x}_p^i(\partial_j|_p) = \delta_j^i. \quad (8)$$

Докажем, что совокупность $\{\tilde{x}_p^i\}_{i=1,\dots,m}$ образует базис в $T_p^*\mathcal{M}$. Линейная независимость $\{\tilde{x}_p^i\}$ вытекает из цепочки равенств:

$$\alpha_1\tilde{x}_p^1 + \dots + \alpha_m\tilde{x}_p^m = 0 \Rightarrow (\alpha_1\tilde{x}_p^1 + \dots + \alpha_m\tilde{x}_p^m)(\partial_i|_p) = 0 \Rightarrow \alpha_i = 0$$

(мы использовали равенство (8)). Обозначим для некоторой 1-формы ω_p значение $\omega_p(\partial_i|_p)$ посредством $(\omega_p)_i$. Проверьте самостоятельно, что 1-форма $(\omega_p)_i\tilde{x}_p^i$ действует на любой вектор X_p так же, как и ω_p :

$$(\omega_p)_i\tilde{x}_p^i(X_p) = (\omega_p)_i\tilde{x}_p^i(X_p^j\partial_j|_p) = (\omega_p)_iX_p^j\delta_j^i(\omega_p)_iX_p^i. \quad (9)$$

Таким образом, мы доказали, что размерность кокасательного пространства равна размерности касательного пространства и равна m , а его базис образует набор координатных 1-форм $\tilde{x}_p^i$.

Нашу локальную конструкцию $T_p^*\mathcal{M}$ можно теперь распространить на все многообразие или его часть, если разрешить 1-формам "гладко зависеть от точки", аналогично тому, как мы переходим от вектора X_p к векторному полю $X \in \mathfrak{V}(\mathcal{M})$. А именно, определим поле 1-формы ω , как $\mathfrak{F}$ -линейное отображение $\mathfrak{V}(\mathcal{M}) \rightarrow \mathfrak{F}(\mathcal{M})$. Это поле каждому векторному полю X ставит в соответствие скалярную функцию на $\mathcal{M}$:

$$\omega(X) \in \mathfrak{F}(\mathcal{M}),$$

так что в каждой точке p , имеет место формула (5). Поле 1-формы называется гладким, если его действие на любое гладкое векторное поле дает в результате гладкую функцию. $\mathfrak{F}(\mathcal{M})$ -линейность поля ω означает следующее обобщение равенства (6):

$$\omega(fX + gY) = f\omega(X) + g\omega(Y), \quad (10)$$

где f, g — произвольные функции из $\mathfrak{F}(\mathcal{M})$. Обобщение формулы (7) на поля 1-форм позволяет поточечно складывать эти поля и умножать их на функции из $\mathfrak{F}(\mathcal{M})$:

$$(f\omega_1 + g\omega_2)(X) = f\omega_1(X) + g\omega_2(X), \quad (11)$$

где ω_1, ω_2 — произвольные поля 1-форм, f, g — произвольные функции из $\mathfrak{F}(\mathcal{M})$.

Типичным и в определенном смысле основным типом 1-форм является *дифференциал* df функции f . На любом векторном поле X его действие определено по формуле:

$$df(X) \equiv X(f). \quad (12)$$

Проверьте самостоятельно, что все требуемые для 1-форм свойства в этом определении выполнены. Рассмотрим совокупность $\{dx^i\}_{i=1,\dots,m}$ дифференциалов координатных функций $x^i: \mathcal{M} \rightarrow \mathbb{R}$ в некоторой карте атласа $\mathbf{At}(\mathcal{M})$. Пользуясь определением (12), имеем:

$$dx^i(X) = X(x^i) = X^j \partial_j(x^i) = X^j \delta_j^i = X^i.$$

Другими словами, поле dx^i является продолженной на многообразие или его часть версией локальной 1-формы $\tilde{x}_p^i$. Отсюда следует, что поля $\{dx^i\}_{i=1,\dots,m}$ являются базисными в пространстве полей 1-форм и любую 1-форму можно представить в виде разложения:

$$\omega = \omega_i dx^i, \quad (13)$$

при этом совокупность $\{\omega_i\}_{i=1,\dots,m}$ называется *координатами 1-формы* ω (в данной карте с координатами x). Эти координаты теперь являются функциями координат многообразия (гладкими, если поле 1-формы гладкое). Представление (13) объясняет, почему поля 1-форм иногда называют *дифференциальными 1-формами*. Результат действия поля 1-формы ω на векторное поле X является скалярной функцией, значение которой в каждой точке p определяется формулой (9).

Как следует из формулы (11), поля 1-форм образуют $\mathfrak{F}(\mathcal{M})$ -модуль, который мы будем обозначать $\mathfrak{V}^*(\mathcal{M})$.

Сделаем несколько замечаний.

- 1-формы — это геометрические объекты в определенном смысле "равноправные" с векторами. Если наглядным образом вектора X_p является направленный отрезок, приложенный к определенной точке, то наглядным образом 1-формы ω_p является плоская площадка, проходящая через определенную точку. Эта площадка образована всеми векторами X_p , на которых выполняется равенство $\omega_p(X_p) = 0$. В координатах это равенство примет вид: $(\omega_p)_i X_p^i = 0$, что при фиксированных координатах $(\omega_p)_i$ дает уравнение плоскости, проходящей через точку p в касательном пространстве $T_p \mathcal{M}_m$. Эта плоскость называется *ядром* 1-формы в точке p и обозначается $\ker \omega_p$.
- Равноправие 1-форм и векторов проявляется в их взаимной дуальности: векторы тоже можно рассматривать как линейные функции на 1-формах, при этом $X_p(\omega_p) \equiv \omega_p(X_p)$. Формально это означает, что $(T_p)^{**}(\mathcal{M}) = T_p(\mathcal{M})$ и $\mathfrak{V}^{**}(\mathcal{M}) = \mathfrak{V}(\mathcal{M})$.
- В старой литературе по геометрии *векторы* называются *контравариантными векторами*, а 1-формы — *ковариантными векторами* и в покомпонентной записи первые записывают с индексами сверху, например, X^i , а вторые — с индексами внизу, например, ω_i . На самом деле, по существу, дело здесь не в положении индексов (хотя их расположение облегчает интерпретацию и выкладки), а в различной геометрической природе и различных законах преобразования компонент этих объектов при замене системы координат. Компоненты векторов при смене координат $x \rightarrow x' = x'(x)$ преобразуются по закону:

$$X'^i = J_j^i X^j,$$

а компоненты 1-форм по закону:

$$\omega'_i = (J^{-1})_i^j \omega_j,$$

где

$$J_j^i \equiv \frac{\partial x'^i}{\partial x^j}, \quad (J^{-1})_j^i \equiv \frac{\partial x^i}{\partial x'^j}$$

прямая и обратная матрицы Якоби. Говоря на современном языке, векторы и 1-формы реализуют различные простейшие линейные представления группы общесоординатных преобразований.

4. В качестве базисных полей необязательно рассматривать только координатные базисы $\{\partial_i\}_{i=1,\dots,m}$ и $\{dx^i\}_{i=1,\dots,m}$. Иногда в приложениях удобны базисы более общего типа: $\{e_{(i)}\}_{i=1,\dots,m}$ и $\{\theta^{(i)}\}_{i=1,\dots,m}$, где первый набор в каждой точке p определяет некоторый базис касательного пространства $T_p\mathcal{M}$, а второй набор — базис кокасательного пространства $T_p^*\mathcal{M}$. При этом эти базисы могут даже не быть дуальными друг другу, т.е. $G_j^i \equiv \theta^{(i)}(e_{(j)}) \neq \delta_j^i$ (но "матричное поле" $(G)_p$ обязано быть невырожденным, т.е. $\det(G)_p \neq 0$ для всех $p \in \mathcal{M}$). Соответствующий подход к описанию геометрических объектов на многообразиях называется *тетрадным формализмом*, а базисы такого общего типа иногда называют *неголономными*. Мы не будем использовать их в настоящих лекциях.
5. Множества всевозможных пар (p, v_p) и (p, ω_p) , где $v_p \in T_p\mathcal{M}$, а $\omega_p \in T_p^*\mathcal{M}$ называются соответственно *касательным* и *кокасательным* расслоениями многообразия $\mathcal{M}$ и обозначаются соответственно $T\mathcal{M}$ и $T^*\mathcal{M}$. Эти расслоения являются гладкими многообразиями, причем $\dim T\mathcal{M} = \dim T^*\mathcal{M} = 2m$. Наряду с самим многообразием $\mathcal{M}$ они являются важнейшими объектами при изучении глобальных топологических характеристик многообразия [5].

5 Тензоры

По аналогии с тем, как из переменных $\{x^1, \dots, x^n\}$ можно строить однородные полиномы вида:

$$P_1 = a_i x^i; \quad P_2 = a_{ij} x^i x^j; \quad \dots \quad P_n = a_{i_1 \dots i_n} x^{i_1} \dots x^{i_n},$$

где $a_i, a_{ij}, a_{i_1 \dots i_n}$ — наборы коэффициентов, из векторов и 1-форм на многообразии $\mathcal{M}$ можно строить *тензорные полиномы*, вводя операцию $\otimes$ *тензорного произведения*. В дальнейшем все конструкции являются локальными, т.е. должны определяться в точке. В настоящих лекциях, допуская для упрощения обозначений и сокращения объема некоторую вольность, мы будем сразу говорить про *тензорные поля* на многообразии, понимая под ними совокупность тензоров во всех точках многообразия и не делая различия между тензорными полями и тензорами.

Определим тензорное произведение $\omega_1 \otimes \omega_2$ двух 1-форм ω_1 и ω_2 как билинейную вещественнозначную функцию на векторных полях, действующую по правилу:

$$(\omega_1 \otimes \omega_2)(X, Y) = \omega_1(X)\omega_2(Y) \tag{14}$$

при всех $X, Y \in \mathfrak{V}(\mathcal{M})$, причем

$$(\omega_1 \otimes \omega_2)(fX + gY, Z) = f\omega_1(X)\omega_2(Z) + g\omega_1(Y)\omega_2(Z),$$

и

$$(\omega_1 \otimes \omega_2)(X, fY + gZ) = f\omega_1(X)\omega_2(Y) + g\omega_1(X)\omega_2(Z)$$

для всех $X, Y, Z \in \mathfrak{V}(\mathcal{M})$ и всех $f, g \in \mathfrak{F}(\mathcal{M})$ (свойство $\mathfrak{F}(\mathcal{M})$ -линейности по обоим аргументам).

Определенное таким образом тензорное произведение двух 1-форм является простейшим представителем — так называемым, *простым* или *разложимым* — пространства $\mathcal{T}^{(2,0)}(\mathcal{M})$ тензоров валентности $(2, 0)$, которые можно определить как $\mathfrak{F}(\mathcal{M})$ -билинейные функции на парах векторов из $\mathfrak{V}(\mathcal{M}) \times \mathfrak{V}(\mathcal{M})$. Общий элемент из этого пространства, вообще говоря, не является простым или разложимым, т.е. его, в общем случае, нельзя представить в виде тензорного произведения двух 1-форм. Оказывается, что его всегда можно представить в виде линейной комбинации простых тензоров такого вида.

Для того, чтобы в этом убедиться, рассмотрим тензорные произведения координатных 1-форм: $\omega_{ij} \equiv dx^i \otimes dx^j$. Нетрудно показать, что они образуют базис в пространстве $\mathcal{T}^{(0,2)}(\mathcal{M})$. Действительно, предполагая их линейную зависимость, имеем линейную комбинацию $a_{ij}\omega^{ij} = 0$. Это означает, что на любой паре векторов $a_{ij}\omega^{ij}(X, Y) = 0$. Выбирая в качестве X и Y всевозможные пары координатных векторов ∂_i , в силу определения (14) операции $\otimes$ и ее линейности получаем последовательно:

$$a_{11} = 0, \quad a_{12} = 0, \quad \dots, \quad a_{nn} = 0,$$

т.е. ω^{ij} — линейно независимы. Рассмотрим произвольный тензор $D \in \mathcal{T}^{(0,2)}(\mathcal{M})$. Обозначим

$$D_{ij} \equiv D(\partial_i, \partial_j).$$

Тогда этот тензор единственным образом представляется в виде разложения:

$$D = D_{ij}\omega^{ij},$$

что проверяется действием левой и правой части на произвольную пару векторов, записанных в виде их разложений в базисе ∂ . К примеру действие правой части в силу определения ω^{ij} , определения (14) и общих свойств базисов dx и ∂ приведет к результату:

$$D(X, Y) = D_{ij}X^iY^j$$

— координатному правилу вычисления действия тензора D на пару векторов X, Y . Из этого представления очевидно, что в каждой точке тензор $D \in \mathcal{T}^{(0,2)}(\mathcal{M})$ определяет обычную билинейную форму, при этом ее компоненты D_{ij} зависят от выбора точки многообразия.

Вспоминая о дуальности векторных полей и полей 1-форм и о соотношении $\mathfrak{V}^{**}(\mathcal{M}) = \mathfrak{V}(\mathcal{M})$, нетрудно построить дуальный векторный аналог пространства $\mathcal{T}^{(2,0)}(\mathcal{M})$. Простым или разложимым элементов этого пространства является тензорное произведение $X \otimes Y$ пары векторов, которое определяется своим действием на произвольную пару 1-форм:

$$(X \otimes Y)(\omega_1, \omega_2) = X(\omega_1)Y(\omega_2) \equiv \omega_1(X)\omega_2(Y)$$

со свойствами $\mathfrak{F}(\mathcal{M})$ -линейности по каждому аргументу. Тензоры такого типа являются элементами пространства тензоров $\mathcal{T}^{(0,2)}(\mathcal{M})$. Повторяя дословно все предыдущие построения и выкладки, приходим к координатному представлению любого тензора $Q \in \mathcal{T}^{(0,2)}(\mathcal{M})$:

$$Q = Q^{ij}\partial_i \otimes \partial_j,$$

где $Q^{ij} \equiv Q(dx^i, dx^j)$, а множество $\{\partial_i \otimes \partial_j\}$ образует координатный базис в $\mathcal{T}^{(0,2)}(\mathcal{M})$.

Построение общей конструкции тензора типа (r, s) теперь очевидно.

Рассмотрим $\mathfrak{F}(\mathcal{M})$ -полилинейное отображение

$$T : \underbrace{\mathfrak{V}(\mathcal{M}) \times \cdots \times \mathfrak{V}(\mathcal{M})}_r \times \underbrace{\mathfrak{V}^*(\mathcal{M}) \times \cdots \times \mathfrak{V}^*(\mathcal{M})}_s \rightarrow R,$$

где r и s — любые неотрицательные вещественные числа. Такое отображение определяет элемент пространства $\mathcal{T}^{(r,s)}(\mathcal{M})$, называемого *пространством тензоров валентности* (r, s) .

У каждого тензора такой валентности имеется r векторных аргументов и s аргументов в виде 1-форм. В координатах такой тензор представляется разложением вида:

$$T = T_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta_s} (dx^{\alpha_1} \otimes \cdots \otimes dx^{\alpha_r} \otimes \partial_{\beta_1} \otimes \cdots \otimes \partial_{\beta_s}),$$

где $T_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta_s} \equiv T(\partial_{\alpha_1}, \dots, \partial_{\alpha_r}, dx^{\beta_1}, \dots, dx^{\beta_s})$ — компоненты тензора в базисе $\{dx, \partial\}$, $(dx^{\alpha_1} \otimes \cdots \otimes dx^{\alpha_r} \otimes \partial_{\beta_1} \otimes \cdots \otimes \partial_{\beta_s})$ — базис в пространстве $\mathcal{T}^{(r,s)}(\mathcal{M})$. Действие тензора на свои аргументы в координатах записывается в виде суммы:

$$T(X_1, \dots, X_r, \omega_1, \dots, \omega_s) = T_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta_s} X_1^{\alpha_1} \cdots X_r^{\alpha_r} (\omega_1)_{\beta_1} \cdots (\omega_s)_{\beta_s}.$$

Разумеется 1-формы и векторы сами являются тензорами валентностей $(1, 0)$ и $(0, 1)$ соответственно, а скалярные функции формально являются элементами пространства тензоров нулевого ранга $\mathcal{T}^{(0,0)}(\mathcal{M})$.

Рассмотрим теперь основные операции над тензорами.

1. *Тензорное произведение тензоров.* Операция тензорного произведения определяет отображение вида:

$$\otimes : \mathcal{T}^{(r_1, s_1)}(\mathcal{M}) \times \mathcal{T}^{(r_2, s_2)}(\mathcal{M}) \rightarrow \mathcal{T}^{(r_1+r_2, s_1+s_2)}(\mathcal{M}),$$

такое, что любой паре тензоров $T_1 \in \mathcal{T}^{(r_1, s_1)}$ и $T_2 \in \mathcal{T}^{(r_2, s_2)}$ она ставит в соответствие тензор $T_1 \otimes T_2 \in \mathcal{T}^{(r_1+r_2, s_1+s_2)}$, причем его значение на любых $r_1 + r_2$ векторах X_i и $s_1 + s_2$ 1-формах определяется по правилу:

$$(T_1 \otimes T_2)(X_1, \dots, X_{r_1}, \dots, X_{r_1+r_2}, \omega_1, \dots, \omega_{s_1}, \dots, \omega_{s_1+s_2}) \equiv \\ T_1(X_1, \dots, X_{r_1}, \omega_1, \dots, \omega_{s_1}) T_2(X_{r_1+1}, \dots, X_{r_1+r_2}, \omega_{s_1+1}, \dots, \omega_{s_1+s_2}).$$

Например, если $T_1 \in \mathcal{T}^{(2,0)}(\mathcal{M})$, $T_2 \in \mathcal{T}^{(0,1)}(\mathcal{M})$, то

$$(T_1 \otimes T_2)(X, Y, \omega) = T_1(X, Y) T_2(\omega).$$

Компоненты тензорного произведения тензоров равны произведению соответствующих компонент сомножителей (докажите!). Например для приведенного выше примера:

$$(T_1 \otimes T_2)_{ij}^k = (T_1)_{ij} (T_2)^k.$$

Полагая один из сомножителей тензорного произведения скалярной функцией, приходим к операции умножения тензора на скаляр, свойства которой очевидны. Отметим, что тензорное произведение, вообще говоря, некоммукативно: $T_1 \otimes T_2 \neq T_2 \otimes T_1$.

2. *Сумма тензоров.* Суммой тензоров одинаковой валентности называется тензор той же валентности, значение которого на его аргументах равно по определению сумме значений слагаемых на тех же аргументах:

$$(T_1 + T_2)(X_1, \dots, X_r, \omega_1, \dots, \omega_s) \equiv T_1(X_1, \dots, X_r, \omega_1, \dots, \omega_s) + T_2(X_1, \dots, X_r, \omega_1, \dots, \omega_s)$$

для всех $T_1, T_2 \in T_1 \in \mathcal{T}^{(r,s)}(\mathcal{M})$ и для всех r и s . Проверьте, опираясь на это определение, что координаты суммы тензоров равны сумме соответствующих координат слагаемых. Проверьте также, что сумма тензоров дистрибутивна относительно умножения на скаляры:

$$f(T_1 + T_2) = fT_1 + fT_2$$

для всех тензоров T_1 и T_2 любой одинаковой валентности и любой скалярной функции f на $\mathcal{M}$.

3. *Свертка или внутреннее произведение тензоров.* Для любого тензора $T \in \mathcal{T}^{(r,s)}(\mathcal{M})$, у которого $r > 0$ и $s > 0$ определена операция свертки:

$$C_{r_1}^{s_1} : \mathcal{T}^{(r,s)}(\mathcal{M}) \rightarrow \mathcal{T}^{(r-1,s-1)}(\mathcal{M}),$$

действующая по правилу:

$$(C_{r_1}^{s_1} T)(X_1, \dots, \widehat{X_{r_1}}, \dots, X_r, \omega_1, \dots, \omega_1, \dots, \widehat{\omega_{s_1}}, \dots, \omega_s) \equiv \quad (15)$$

$$T(X_1, \dots, \partial_i, \dots, X_r, \omega_1, \dots, \omega_1, \dots, dx^i, \dots, \omega_s).$$

Здесь $0 < r_1 \leq r$ и $0 < s_1 \leq s$, шляпка над аргументом означает, что он отсутствует, а по индексу i производится суммирование. В координатах $C_{r_1}^{s_1}$ -свертка тензора осуществляется приравниванию r_1 -ого нижнего индекса s_1 -ому верхнему и суммированию по всем значениям этого одинакового индекса от 1 до m . Например, для разобранный в п.1 примера:

$$(C_2^1(T_1 \otimes T_2))_\alpha \equiv (T_1)_{\alpha\beta}(T_2)^\beta.$$

Ввиду согласованного (взаимно-обратного) закона преобразования базисов dx и ∂x , определение свертки не зависит от конкретного выбора базиса (проверьте!).

Сделаем в заключении этого параграфа несколько замечаний.

1. Замечательная особенность всех введенных алгебраических операций заключается в том, что тензоры при таких операциях переходят в новые тензоры. В физике тензоры используются для представления физических величин в физических уравнениях, удовлетворяющих условию их независимости от выбора системы координат. Таким образом, рассмотренные нами операции представляют собой "разрешенные правила игры" в "тензорный конструктор": например операцию суммы тензоров можно использовать для инвариантного описания принципа аддитивности или принципа суперпозиции, а тензорное произведение — для определения составных (неэлементарных) тензорных физических величин. Примером "запрещенной" операции является, например, операция, заключающаяся в суммировании всех компонент тензора. Полученное число, вообще говоря, не будет тензором (в данном случае скаляром) и его значение в одной

системе координат, вообще говоря, не будет совпадать с суммой компонент в какой-нибудь другой системе координат. Для полноты алгебру тензоров следует дополнить тензорным анализом, включающем в себя дифференциальные операции, переводящие тензоры в тензоры. К их числу относятся, например, ковариантная производная, внешний дифференциал, производная Ли. В настоящих лекциях мы рассмотрим далее лишь последнюю операцию.

- Операция $\otimes$ позволяет рассматривать множество тензоров всех валентностей $\mathcal{T}(\mathcal{M})$ как *градуированную тензорную алгебру над $\mathcal{M}$* :

$$\mathcal{T}(\mathcal{M}) \equiv \bigoplus_{r,s} \mathcal{T}^{(r,s)}(\mathcal{M}).$$

Здесь $\oplus$ — символ стандартной прямой суммы линейных пространств. Действительно, множество $\mathcal{T}(\mathcal{M})$ замкнуто относительно операции $\otimes$, при этом любой тензор относится к какому-то определенному типу, задаваемому парой чисел (r, s) , определяющей $\mathbb{Z}_+^2$ -градуировку.

- Тензорное поле является гладким, если его значение на любых гладких аргументах является гладкой скалярной функцией. Тензорное поле гладко тогда и только тогда, когда его компоненты в некоторой системе координат являются гладкими функциями на $\mathcal{M}$.
- Можно рассмотреть и другие интерпретации тензоров $\mathcal{T}^{(r,s)}(\mathcal{M})$. К примеру тензор P валентности $(1, 1)$ в соответствии с нашим определением мы должны рассматривать как отображение векторного поля и 1-формы в скалярную функцию. Но можно посмотреть на эту ситуацию иначе. Значение P на векторе X является таким, объектом, который действуя на любую 1-форму даст скаляр. Но таким свойством обладают только векторы. Значит помимо интерпретации

$$P : \mathfrak{V}(\mathcal{M}) \times \mathfrak{V}^*(\mathcal{M}) \rightarrow \mathfrak{F}(\mathcal{M}),$$

существует равноправная интерпретация:

$$P : \mathfrak{V}(\mathcal{M}) \rightarrow \mathfrak{V}(\mathcal{M}),$$

позволяющая рассматривать тензорное поле P как поле *линейного оператора* или *аффинора*. Компонентная запись делает это наблюдение явным:

$$Y^\alpha = P^\alpha_\beta X^\beta,$$

поскольку представляет собой обычное правило записи действия оператора в матричной компонентной форме. Аналогичная смена интерпретации возможна и для тензоров всех других валентностей, отличных от $(0, 0)$, $(1, 0)$, $(0, 1)$.

- Тензор $S \in \mathcal{T}^{(r,0)}(\mathcal{M})$ называется *симметричным* (*антисимметричным*), если его значение не меняется при любой (четной) перестановке его аргументов (а при любой нечетной оно меняет знак). Аналогичное определение имеет место и для тензоров валентности $(0, s)$. Тензор может быть симметричным или антисимметричным по некоторому подмножеству аргументов (или одновременно симметричным по одному подмножеству

аргументов и антисимметричным по другому). При этом тензор смешанного типа (r, s) может быть симметричным или антисимметричным по векторам или 1-формам по отдельности, но не имеет никакого смысла симметризация или антисимметризация тензора по аргументам различных типов.

6. С групповой точки зрения тензоры на многообразии являются различными, вообще говоря, приводимыми линейными представлениями группы общекординатных преобразований. При этом их разбиения на симметричные или антисимметричные соответствует разложению общего линейного представления на неприводимые относительно действия группы $P_r \times P_s$ компоненты (схемы Юнга). Здесь P_r и P_s — группы перестановок, действующие на аргументы тензора валентности (r, s) .

6 Отображения многообразий и геометрических объектов

При отображении многообразий различные геометрические объекты можно "переносить" с одних многообразий на другие.

Пусть $h: \mathcal{M} \rightarrow \mathcal{N}$ — отображение многообразий. h является *гладким отображением*, если его координатное представление является гладким в смысле обычного анализа. Более подробно, если (U, φ) карта из $\mathbf{At}(\mathcal{M})$, а (V, ψ) карта из $\mathbf{At}(\mathcal{N})$, такие что $V \cap h(U) \neq \emptyset$, то на $U \cap h^{-1}(V \cap h(U))$ определено координатное представление отображения $h_{UV}: R^m \rightarrow R^n$ по формуле:

$$h_{UV}(x_p) \equiv \psi \circ h \circ \varphi^{-1}(x_p) \in R^n. \quad (16)$$

Отображение h гладко, если функция h_{UV} — гладкая для всех пар $\{U, V\} \in \mathbf{At}(\mathcal{M}) \times \mathbf{At}(\mathcal{N})$, для которых $V \cap h(U) \neq \emptyset$. Наглядно формула (16) иллюстрируется диаграммой (17): справедливость формулы (16) эквивалентна коммутативности этой диаграммы.

$$\mathcal{M}_m @>h>> \mathcal{N}_n @V\varphi VV @VV\psi V R^m @>h_{UV}>> R^n \quad (17)$$

Конструкция координатного представления допускает некоторое обобщение. Пусть $f \in \mathfrak{F}(\mathcal{N})$ и h — гладкое отображение $\mathcal{M}_m \rightarrow \mathcal{N}_n$. Тогда на $\mathcal{M}_m$ с помощью формулы

$$f_h \equiv f \circ h \quad (18)$$

определена гладкая функция $f_h \in \mathfrak{F}(\mathcal{M})$. Другими словами, *при гладких отображении многообразий гладкие функции "переносятся" в противоположную сторону — с образа на прообраз*.

А как переносятся векторы при отображениях многообразий? Рассмотрим некоторый вектор $X_p \in T_p\mathcal{M}$.

Всякое гладкое отображение $h: \mathcal{M} \rightarrow \mathcal{N}$ определяет в каждой точке $p \in \mathcal{M}$ отображение (h_*) : $T_p\mathcal{M} \rightarrow T_{h(p)}\mathcal{N}$, которое называется *дифференциалом отображения h в точке p* и действует по правилу:

$$(h_*)(X_p)(f) \equiv X_p(f_h) = X_p(f \circ h) \quad (19)$$

для всякой $f \in \mathfrak{F}(\mathcal{N})$.

Другими словами, при отображении h всякий вектор X_p как дифференцирование на $\mathcal{M}$ переходит в вектор-дифференцирование $(h_*)_p(X_p)$ на $\mathcal{N}$, таким образом, что результат дифференцирования им произвольной функции f совпадает с результатом дифференцирования

исходного вектора перенесенной на многообразие $\mathcal{M}$ функции f_h . В силу своего определения дифференциал отображения h_* является линейным отображением касательных пространств. Чтобы установить его координатное представление, перейдем к координатам в базисах $\partial/\partial x|_p$ и $\partial/\partial y|_{h(p)}$. По определению дифференциала имеем:

$$(h_*)(X_p)(f) = [(h_*)(X_p)]^\alpha \frac{\partial f}{\partial y^\alpha} \Big|_{h(p)} X_p^\beta \frac{\partial}{\partial x^\alpha} \Big|_p (f \circ h)(p) X_p^\beta \frac{\partial f}{\partial y^\beta} \Big|_{h(p)} \frac{\partial y^\beta}{\partial x^\alpha} \Big|_p,$$

откуда, в силу произвольности f и произвольности точки p следует координатное представление h_* :

$$[(h_*)(X)]^\alpha_{h(p)} = \frac{\partial y^\alpha}{\partial x^\beta} \Big|_p X_p^\beta. \quad (20)$$

Мы видим, что дифференциал отображения h в каждой точке многообразия $\mathcal{M}$ в действительности осуществляет линейное отображение касательных пространств $T_p\mathcal{M} \rightarrow T_{h(p)}\mathcal{N}$ посредством $n \times t$ матрицы преобразования $(\partial y/\partial x)$, называемой *матрицей Якоби*, компоненты которой зависят от точки. Диаграмма (21) наглядно иллюстрирует действие дифференциала отображения: определение (19) эквивалентно коммутативности этой диаграммы.

$$\mathcal{M}_m @>h>> \mathcal{N}_n @A\pi_{\mathcal{M}}AA @A\pi_{\mathcal{N}}AT \mathcal{M} @>h_*>> T\mathcal{N} \quad (21)$$

Отображения $\pi_{\mathcal{M}}$ и $\pi_{\mathcal{N}}$ на диаграмме — это естественные проекции касательных расслоений этих многообразий, действующие по правилам:

$$\pi_{\mathcal{M}} : T\mathcal{M} \ni (p, X_p) \mapsto p; \quad \pi_{\mathcal{N}} : T\mathcal{N} \ni (q, Y_q) \mapsto q.$$

Перейдем к рассмотрению поведения 1-форм при отображениях. Нетрудно понять, что, как и функции, 1-формы "переносятся" отображениями $h : \mathcal{M}_m \rightarrow \mathcal{N}_n$ в "обратную сторону по той простой причине, что 1-формы — это и есть функции на векторах. Действительно,

если $\omega \in \mathfrak{V}^*(\mathcal{N})$, то в каждой точке $q = h(p) \in \mathcal{N}$ корректно определено отображение $h_q^* : T_q^*\mathcal{N} \rightarrow T_p^*(\mathcal{M})$, называемое *кодифференциалом отображения h в точке p* , действующее на всяком $Y_q \in T_q\mathcal{N}$ по правилу:

$$h^*(\omega_q)(X_p) \equiv \omega_p((h_*)(X_p)). \quad (22)$$

Другими словами, значение образа кодифференциала некоторой 1-формы на любом векторе пространства $T_p(\mathcal{M})$ равно значению исходной 1-формы на образе дифференциала этого вектора. В координатах согласно определению (22) будем иметь:

$$h^*(\omega_q)(X_p) = [h^*(\omega_q)]_\alpha X_p^\alpha \equiv (\omega_p)_\beta \frac{\partial y^\beta}{\partial x^\alpha} \Big|_p X_p^\alpha,$$

откуда ввиду произвольности p и X_p получаем координатное представление действия кодифференциала:

$$[h^*(\omega)]_\alpha = \frac{\partial y^\beta}{\partial x^\alpha} \omega_\beta. \quad (23)$$

Мы видим, что кодифференциал h^* отображения h в каждой точке многообразия $\mathcal{N}$ осуществляет линейное отображение кокасательных пространств $T_{h(p)}^*\mathcal{N} \rightarrow T_p\mathcal{M}$ посредством $m \times n$ матрицы преобразования $(\partial y/\partial x)^T$, которую можно назвать *сопряженной матрицей Якоби*. Диаграмма (24) наглядно иллюстрирует действие кодифференциала отображения: определение (22) эквивалентно коммутативности этой диаграммы.

$$\mathcal{M}_m @>h>> \mathcal{N}_n @A\pi_{\mathcal{M}}^*AA @A\pi_{\mathcal{N}}^*AT \mathcal{M} @<h^*<< T\mathcal{N} \quad (24)$$

Отображения $\pi_{\mathcal{M}}^*$ и $\pi_{\mathcal{N}}^*$ на диаграмме — это проекции кокасательных расслоений этих многообразий, действующие по правилам:

$$\pi_{\mathcal{M}}^* : T^*\mathcal{M} \ni (p, \omega_p) \mapsto p; \quad \pi_{\mathcal{N}}^* : T^*\mathcal{N} \ni (q, \omega_q) \mapsto q.$$

Отображения дифференциала и кодифференциала легко обобщаются на тензоры специальных типов. Пусть как и прежде $h : \mathcal{M} \rightarrow \mathcal{N}$ и пусть $T_1 \in \mathcal{T}^{(0,s)}(\mathcal{M})$ и $T_2 \in \mathcal{T}^{(r,0)}(\mathcal{N})$. Тогда определены отображения

$$(h_*)^{\otimes s} : \mathcal{T}_p^{(0,s)}(\mathcal{M}) \rightarrow \mathcal{T}_{h(p)}^{(0,s)}(\mathcal{N}), \quad (h^*)_{h(p)}^{\otimes r} : \mathcal{T}_{h(p)}^{(r,0)}(\mathcal{N}) \rightarrow \mathcal{T}_p^{(r,0)}(\mathcal{M}),$$

называемые соответственно *s-ой и r-ой тензорными степенями дифференциала и кодифференциала в точке p*, действующие на соответствующие тензоры по правилам:

$$(h_*)^{\otimes s}(T_1)(\omega_1, \dots, \omega_s) \equiv T_1(h^*(\omega_1), \dots, h_*(\omega_s)); \quad (25)$$

$$(h^*)_{h(p)}^{\otimes r}(T_2)(X_1, \dots, X_r) \equiv T_2(h_*(X_1), \dots, h_*(X_r)) \quad (26)$$

для всех $\omega_i \in T_{h(p)}^*\mathcal{N}$, и всех $X_j \in T_p\mathcal{M}$. В координатах эти отображения будут выражаться формулами:

$$(T_1)_{h(p)}^{\alpha_1 \dots \alpha_s} \frac{\partial y^{\alpha_1}}{\partial x^{\beta_1}} \Big|_{h(p)} \cdots \frac{\partial y^{\alpha_s}}{\partial x^{\beta_s}} \Big|_{h(p)} (T_1)_p^{\beta_1 \dots \beta_s}; \quad (T_2)_{h(p)}^{\alpha_1 \dots \alpha_s} \frac{\partial y^{\beta_1}}{\partial x^{\alpha_1}} \Big|_p \cdots \frac{\partial y^{\beta_r}}{\partial x^{\alpha_r}} \Big|_p (T_2)_{h(p)}^{\beta_1 \dots \beta_r}. \quad (27)$$

Тензоры смешанного типа, вообще говоря, никуда не переносятся отображениями многообразий.

Сделаем несколько замечаний к этому параграфу.

1. Гладкие функций на $\mathcal{N}$ можно понимать как элементы $\mathcal{T}^{(0,0)}(\mathcal{N})$. Таким образом, комбинаторика формул (18) и (26) дает:

$$(h^*)^{\otimes 0}(f) \equiv f_h.$$

Иногда вместо f_h используют запись h^*f для перенесенной на $\mathcal{M}$ при отображении h функции f .

2. Мы определяли и рассматривали отображения дифференциала и кодифференциала в точке. Можно ли говорить об отображениях тензорных полей? Как показывают простые примеры, в общем случае этого сделать нельзя. Предположим, что на $\mathcal{M}$ задано векторное поле X и пусть отображение $h : \mathcal{M} \rightarrow \mathcal{N}$ таково, что $h(p_1) = h(p_2)$ при $p_1 \neq p_2$. Поскольку, в общем случае, $(h_*)_{p_1}(X_{p_1}) \neq (h_*)_{p_2}(X_{p_2})$ (рассмотрите в качестве примера

отображение $x \mapsto \sin x$ в точках $x_1 = \pi/4$ и $x_2 = 3\pi/4$), то говорить об отображениях векторных полей в такой ситуации нельзя! Замечательным фактом теории отображений гладких многообразий является *существование отображения полей 1-форм и тензоров типа $\mathcal{T}^{(r,0)}(\mathcal{N})$ на многообразии $\mathcal{M}$ при любых гладких отображениях h !* (Проверьте это!)

3. Отображение h называется *диффеоморфизмом*, если оно взаимно-однозначно, гладко и имеет гладкое обратное. Диффеоморфизм часто можно понимать в активном смысле как гладкую деформацию самого многообразия, т.е. как отображение $\mathcal{M} \rightarrow \mathcal{M}$. Примером диффеоморфизма является замена координат на многообразии (при этом условие обратимости и гладкости, а значит и диффеоморфности может нарушаться в отдельных точках, на линиях или поверхностях — как говорят, на множестве точек меры нуль). Для координатного диффеоморфизма всегда существует обратное и матрица Якоби обратима. При этом формулы (20), (23) и (27) по существу переходят в законы преобразования компонент соответствующих геометрических объектов при замене координат.
4. Для диффеоморфизма $\phi : \mathcal{M} \rightarrow \mathcal{N}$ всегда существует (r, s) -дифференциал отображения $\phi_{(r,s)}$ и обратный $\phi_{(r,s)}^{-1}$, переводящие соответственно тензорные поля $T \in \mathcal{T}^{(r,s)}(\mathcal{M})$ в поля $T \in \mathcal{T}^{(r,s)}(\mathcal{N})$ и наоборот и определяемые соотношениями:

$$\phi_{(r,s)p} \equiv \underbrace{(h_*)_p \otimes \cdots \otimes (h_*)_p}_s \text{ раз} \otimes \underbrace{(h^*)_p^{-1} \otimes \cdots \otimes (h^*)_p^{-1}}_r \text{ раз} \quad (28)$$

$$\phi_{(r,s)p}^{-1} \equiv \underbrace{(h^*)_p^{-1} \otimes \cdots \otimes (h^*)_p^{-1}}_s \text{ раз} \otimes \underbrace{(h_*)_p \otimes \cdots \otimes (h_*)_p}_r \text{ раз} \quad (29)$$

5. Можно сказать, что дифференциал отображения, рассматриваемый в некоторой точке, является локальной версией самого отображения в этой точке. Более точно, *дифференциал отображения h_* является линейной аппроксимацией отображения h в каждой точке*. Таким образом, основная идея дифференциального исчисления функций вещественных переменных остается в силе и на гладких многообразиях: *"локально все отображения линейны"*. Следует подчеркнуть, что, как и в случае стандартного вещественного анализа, в особых точках отображения это свойство, вообще говоря, нарушается.

7 Интегральные кривые векторных полей и потоки на многообразиях

В этом разделе мы рассмотрим несколько полезных понятий, тесно связанных с определением и свойствами производной Ли.

Напомним, что гладкой параметризованной кривой γ на многообразии $\mathcal{M}$ называется гладкое отображение $I \rightarrow \mathcal{M}$, где I — интервал вещественной оси, который без ограничения общности всегда можно выбрать единичным: $I = (0; 1)$. Переходя к координатам $\{x^i\}$ многообразия в некоторой карте, через которую проходит кривая, будем иметь систему функций $\{x^i(t)\}$, описывающих параметризованную кривую γ в координатах этой карты. Здесь $t \in I$

— параметр на кривой. На интервале I мы имеем постоянное векторное поле d/dt — это векторный базис на I , рассматриваемом как одномерное многообразие. Его перенос на кривую γ посредством дифференциала отображения:

$$\gamma_*(d/dt) \equiv \dot{\gamma}$$

называется *векторным полем скорости на кривой* γ . Кривая γ называется регулярной, если $\dot{\gamma} \neq 0$ в каждой точке кривой (в механической интерпретации — нет точек остановки). В координатах поле скорости задается выражением (проверьте!):

$$\dot{\gamma} = \dot{x}^i(t)\partial_i,$$

где точка сверху означает дифференцирование по параметру. Рассмотрим гладкое векторное поле X на многообразии $\mathcal{M}$. Кривая γ называется *интегральной кривой векторного поля* X , если в каждой точке кривой выполняется равенство:

$$\dot{\gamma} = X|_{\gamma}, \quad (30)$$

где $X|_{\gamma}$ — ограничение векторного поля X на кривую γ , определяемое по формуле:

$$X|_{\gamma}(f) \equiv \frac{d}{dt}(\gamma^*f)$$

для всякой функции $f \in \mathfrak{F}(\mathcal{M})$. Уравнение (30) в координатах принимает вид системы обыкновенных дифференциальных уравнений:

$$\frac{dx^i}{dt} = X^i(x(t)), \quad i = 1, \dots, m. \quad (31)$$

Ее геометрический смысл заключается в коллинеарности векторов элементарных перемещений (их линейных частей) $\Delta x^i \partial_i$ вдоль кривой и векторов векторного поля в соответствующих точках кривой. Записывая это условие коллинеарности в виде*:

$$\frac{dx^1}{X^1} = \dots = \frac{dx^m}{X^m} = dt, \quad (32)$$

приходим к т.н. *системе уравнений на характеристики векторного поля* X , т.е. такие $m-1$ -мерные поверхности, которые "сотканы" из линий тока векторного поля. Уравнения характеристик появляются как совокупность интегралов системы (32):

$$F_1(x) = C_1, \dots, F_{m-1}(x) = C_{m-1}, \quad (33)$$

где $\{C_i\}$ — семейство констант интегрирования. При фиксированных значениях констант C_i каждое уравнение из (33) определяет некоторую характеристику в $\mathcal{M}$, а в совокупности пересечение $m-1$ характеристик определяет конкретную интегральную кривую поля X . К

*Мы предполагаем, что все компоненты векторного поля отличны от нуля. Если некоторые из них (например, $X^1, \dots, X^r$ $r \leq m$) равны нулю, то условие коллинеарности подразумевает, что $dx^1 = 0, \dots, dx^r = 0$, т.е. характеристики — это координатные поверхности $x^1 = \text{const}, \dots, x^r = \text{const}$.

такой конкретизации можно прийти, потребовав, чтобы эта кривая проходила через заданную точку p с координатами $(x_p^1, \dots, x_p^n)$ при $t = 0$. В таком случае в систему (33) следует подставить x_p вместо x и разрешить ее относительно C_i . После этого мы вместо (33) будем иметь систему вида:

$$F_1(x) = C_1(x_p), \dots, F_{m-1}(x) = C_{m-1}(x_p),$$

которую можно разрешить относительно каких-то $m - 1$ переменных x (без ограничения общности можно считать, что это первые $m - 1$ -переменных $x^1, \dots, x^{m-1}$):

$$x^1 = f_1(C(x_p), x^m), \dots, x^{m-1} = f_{m-1}(C(x_p), x^m). \quad (34)$$

Далее, подставляя полученные зависимости в уравнение:

$$\frac{dx^m}{X^m(x)} = dt,$$

получаем обыкновенное дифференциальное уравнение, с разделяющимися переменными. Константу интегрирования следует выбрать так, чтобы выполнялось начальное условие: $x^m(0) = f_m(C(x_p), t)|_{t=0} = x_p^m$. Теперь подставляя найденную зависимость $x^m(C(x_p), t)$ в уравнения (34), получаем полное координатное описание интегральной кривой с необходимыми свойствами:

$$x^1(t) = f_1(C(x_p), x^m(C(x_p), t)), \quad \dots, \quad x^{m-1}(t) = f_{m-1}(C(x_p), x^m(C(x_p), t)),$$

$$x^m(t) = f_m(C(x_p), t).$$

Рассмотрим следующий пример. Пусть $X = x\partial_y - y\partial_x$ — векторное поле на плоскости R^2 с координатами $\{x, y\}$. Уравнение характеристик

$$\frac{dy}{x} = -\frac{dx}{y} = dt$$

имеет интеграл $x^2 + y^2 = C = \text{const}$. Он описывает семейство окружностей различных радиусов $R = \sqrt{C}$. Если $x(0) = x_0$, $y(0) = y_0$, то $C = x_0^2 + y_0^2$ и

$$x = \sqrt{x_0^2 + y_0^2} \cos t; \quad y = \sqrt{x_0^2 + y_0^2} \sin t.$$

Очевидно, что семейство интегральных кривых векторного поля X можно рассматривать как отображение $\phi_X: \mathcal{M} \times R \rightarrow \mathcal{M}$, которое каждую точку $p \in \mathcal{M}$ смещает вдоль интегральной кривой, проходящей через эту точку, причем при $t = 0$ это отображение является тождественным:

$$\phi_X: p \mapsto \phi_X(p, t), \quad \phi_X(p, 0) = p$$

и при этом

$$(\phi_X)_*(p, d/dt) = X_p.$$

Отображение ϕ_X называется *поток* векторного поля X на многообразии $\mathcal{M}$. В силу общих теорем существования и единственности в некоторой окрестности любой неособой точки

гладкого векторного поля (т.е. точки, в которой $X \neq 0$) поток ϕ_X является диффеоморфизмом. В окрестности особых точек векторного поля диффеоморфность отображения ϕ_X может нарушаться.

Мы показали, что всякое гладкое (на самом деле достаточно непрерывности!) векторное поле определяет поток на многообразии. Покажем, что всякому потоку на многообразии, понимаемому как гладкое отображение $\phi: \mathcal{M} \times \mathbb{R} \rightarrow \mathcal{M}$, соответствует векторное поле, которое его порождает. Действительно, рассмотрим векторное поле:

$$X_\phi(p) \equiv (\phi)_*(p, d/dt)|_{t=0}.$$

В каждой точке многообразия, где определен поток, это поле определено и его линии потока являются интегральными кривыми векторного поля X_ϕ по определению (30). Таким образом, *существует взаимно-однозначное соответствие между векторными полями и потоками на гладком многообразии.*

Рассмотрим в качестве примера поток на плоскости, задаваемый в координатах соотношениями:

$$x'(x, y, t) = (t+1)x; \quad y'(x, y, t) = \frac{y}{t+1}.$$

Соответствующее векторное поле получается дифференцированием этих соотношений по t и приравниванием t к нулю. В результате будем иметь:

$$X = x\partial_x - y\partial_y.$$

Попробуйте изобразить это векторное поле стрелками на плоскости самостоятельно. Такой поток осуществляет гиперболические (псевдоевклидовы) вращения плоскости.

8 Производная Ли и ее свойства

Введенные выше конструкции дифференциала и кодифференциала отображения позволяют с каждым потоком ассоциировать операцию дифференцирования вдоль потока, определяемую, как уже говорилось, независимо от существования каких-либо геометрических структур на многообразии. Пусть на многообразии $\mathcal{M}$ задано тензорное поле $T \in \mathcal{T}^{(r,s)}(\mathcal{M})$ и пусть ϕ — поток векторного поля X .

Производная Ли $(L_X T)_p$ тензорного поля T вдоль векторного поля X в точке p определяется соотношением:

$$(L_X T)_p \equiv \lim_{t \rightarrow 0} \frac{(\phi^t)^{-1}_{(r,s)}(T_{\phi^t(p)}) - T_p}{t}, \quad (35)$$

где $(\phi^t)^{-1}_{(r,s)}$ — обратный (r, s) -дифференциал отображения ϕ (ф-ла (29)), вычисленный в точке $\phi^t(p)$.

Рисунок 2 поясняет данное определение. Через точку p проходит некоторая интегральная кривая потока ϕ , генерируемого векторным полем X . Смещаясь вдоль нее из точки p на параметрическое расстояние t , мы попадаем в точку $\phi^t(p)$. В этой точке мы наблюдаем тензор $T_{\phi^t(p)}$, который мы напрямую не можем сравнить с тензором T_p в исходной точке, поскольку эти тензоры отнесены к разным точкам многообразия и разным базисам. Но посредством обратного (r, s) -дифференциала $(\phi^t)^{-1}_{(r,s)}$ мы можем перенести тензор $T_{\phi^t(p)}$ обратно в точку p

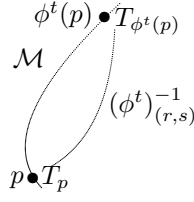


Рис. 2: К определению производной Ли.

и сравнить результат с исходным тензором T_p . Разделив разность перенесенного и исходного тензоров в точке p на t и переходя к пределу при $t \rightarrow 0$, получаем производную Ли, которая, таким образом, имеет смысл *скорости изменения тензора T вдоль потока ϕ* .

Данное безкоординатное определение производной Ли позволяет доказать ряд основных ее общих свойств.

1. Если $T \in \mathcal{T}^{(r,s)}(\mathcal{M})$, то и $L_X T \in \mathcal{T}^{(r,s)}(\mathcal{M})$.
2. Производная Ли R -линейна по обоим аргументам, т.е.

$$L_X(\lambda_1 T_1 + \lambda_2 T_2) = \lambda_1 L_X T_1 + \lambda_2 L_X T_2$$

и

$$L_{\lambda_1 X_1 + \lambda_2 X_2}(T) = \lambda_1 L_{X_1} T + \lambda_2 L_{X_2} T.$$

3. Производная Ли удовлетворяет правилу Лейбница по отношению к тензорному произведению:

$$L_X(T_1 \otimes T_2) = (L_X T_1) \otimes T_2 + T_1 \otimes L_X(T_2).$$

4. Производная Ли коммутирует со сверткой:

$$CL_X T = L_X(CT).$$

В перечисленных выше свойствах T — произвольное тензорное поле произвольной валентности, T_1, T_2 — произвольные тензорные поля произвольных одинаковых валентностей в свойстве 2 и произвольных валентностей в свойстве 3, λ_1, λ_2 — произвольные вещественные числа.

Свойство 1 вытекает непосредственно из определения (сумма и разность пары тензоров есть тензор того же типа). Свойство 2 для верхнего (тензорного) аргумента вытекает из линейности (r, s) -дифференциала и предела. Однородность производной Ли по нижнему (векторному) аргументу можно легко доказать, заметив, что замена $X \rightarrow \lambda X$ "компенсируется" переопределением параметра t потока:

$$\frac{dx}{dt} = \lambda X \Leftrightarrow \frac{dx}{d\tau} = X,$$

где $\tau = \lambda t$. Другими словами, поток векторного поля X по отношению к параметру $\tau = \lambda t$ таков же, каков поток векторного поля λX по отношению к параметру t . Таким образом, имеет место равенство:

$$\frac{(\phi_{\lambda X}^t)^{-1} T_{\phi_{\lambda X}^t(p)} - T_p}{t} \lambda = \frac{(\phi_X^{\lambda t})^{-1} T_{\phi_X^{\lambda t}(p)} - T_p}{\lambda t} \lambda = \frac{(\phi_X^\tau)^{-1} T_{\phi_X^\tau(p)} - T_p}{\tau} \lambda,$$

откуда после перехода к пределу при $\tau \rightarrow 0$ и следует доказываемое свойство:

$$L_{\lambda X}T = \lambda L_X T.$$

Для доказательства второй части свойства 2 для нижнего (векторного) аргумента производной Ли достаточно заметить, что с точностью до $o(t)$ потоки ϕ_{X_1} и ϕ_{X_2} коммутируют (см. ниже геометрическую интерпретацию скобки Ли и теорему 1 в этом параграфе), при этом $\phi_{X_1+X_2}^t = \phi_{X_1}^t \circ \phi_{X_2}^t + o(t)$. Используя это обстоятельство в определении (35), приходим к цепочке равенств

$$\begin{aligned} L_{X_1+X_2}T &\equiv \lim_{t \rightarrow 0} \frac{(\phi_{X_1+X_2}^t)^{-1}T_{\phi_{X_1+X_2}^t(p)} - T_p}{t} \\ &\lim_{t \rightarrow 0} \frac{(\phi_{X_2}^t)^{-1}((\phi_{X_1}^t)^{-1}T_{\phi_{X_1}^t(p)})_{\phi_{X_2}^t(p)} - (\phi_{X_1}^t)^{-1}(T_{\phi_{X_1}^t(p)}) + (\phi_{X_1}^t)^{-1}(T_{\phi_{X_1}^t(p)}) - X_p + o(t^2)}{t} \\ &L_{X_2} \lim_{t \rightarrow 0} (\phi_{X_1}^t)^{-1}(T_{\phi_{X_1}^t(p)}) + L_{X_1}T = L_{X_1}T + L_{X_2}T, \end{aligned}$$

что и требовалось доказать. Во второй строчке мы добавили и вычли в числителе тензор $(\phi_{X_1}^t)^{-1}T_{\phi_{X_1}^t(p)}$ в точке $\phi^t(p)$, в последней строчке мы воспользовались непрерывностью потока и равенством $\phi_X^0 = \text{id}_{\mathcal{M}}$.

Доказательство правила Лейбница основано на ступенчатой конструкции, использованной в доказательстве предыдущего свойства. Имеем цепочку равенств:

$$\begin{aligned} L_X(T_1 \otimes T_2) &\equiv \lim_{t \rightarrow 0} \frac{\phi_{(r,s)}^{-1}(T_1 \otimes T_2)_{\phi^t(p)} - (T_1 \otimes T_2)_p}{t} \\ &\lim_{t \rightarrow 0} (\phi_{(r_1,s_1)}^{-1}(T_1)_{\phi^t(p)} \otimes \phi_{(r_2,s_2)}^{-1}(T_2)_{\phi^t(p)} - \phi_{(r_1,s_1)}^{-1}(T_1)_{\phi^t(p)} \otimes (T_2)_p \\ &\quad + \phi_{(r_1,s_1)}^{-1}(T_1)_{\phi^t(p)} \otimes (T_2)_p - (T_1)_p \otimes (T_2)_p) / t \\ &\lim_{t \rightarrow 0} \frac{\phi_{(r_1,s_1)}^{-1}(T_1)_{\phi^t(p)} \otimes (\phi_{(r_2,s_2)}^{-1}(T_2)_{\phi^t(p)} - (T_2)_p)}{t} + \lim_{t \rightarrow 0} \frac{(\phi_{(r_1,s_1)}^{-1}(T_1)_{\phi^t(p)} - (T_1)_p) \otimes (T_2)_p \otimes (T_2)_p}{t} \\ &= T_1 \otimes L_X T_2 + L_X T_1 \otimes T_2, \end{aligned}$$

что и требовалось доказать. Здесь $T_1 \in \mathcal{T}^{(r_1,s_1)}(\mathcal{M})$, $T_2 \in \mathcal{T}^{(r_2,s_2)}(\mathcal{M})$, $r = r_1 + r_2$, $s = s_1 + s_2$.

Докажем, наконец, коммутативность производной Ли с однократной сверткой (коммутативность кратной свертки доказывается вполне аналогично). Используя определение свертки (15) имеем:

$$\begin{aligned} L_X(CT) &= \lim_{t \rightarrow 0} \frac{\phi_{(r-1,s-1)}^{-1}(CT)_{\phi^t(p)} - (CT)_p}{t} \\ &\lim_{t \rightarrow 0} \frac{\phi_{(r,s)}^{-1}(T_{\phi^t(p)})(\dots, (\partial_\alpha)_p, \dots, dx_p^\alpha, \dots) - T_p(\dots, (\partial_\alpha)_p, \dots, dx_p^\alpha, \dots)}{t} \\ &\left[\lim_{t \rightarrow 0} \frac{\phi_{(r,s)}^{-1}(T)_{\phi^t(p)} - T_p}{t} \right] (\dots, (\partial_\alpha)_p, \dots, dx_p^\alpha, \dots) = C \lim_{t \rightarrow 0} \frac{\phi_{(r,s)}^{-1}(T)_{\phi^t(p)} - T_p}{t} = CL_X T. \end{aligned}$$

Сделаем несколько замечаний.

1. С точки зрения абстрактной алгебры производная Ли является элементом алгебры дифференцирований тензорной алгебры $\mathcal{T}(\mathcal{M})$, сохраняющим ее $\mathbb{Z} \times \mathbb{Z}$ -градуировку. Примерами дифференцирований тензорной алгебры, не сохраняющих градуировку, являются ковариантная производная и внешний дифференциал (повышают левую компоненту градуировки на единицу) и сопряженный к d дифференциал δ (понижает левую компоненту градуировки на единицу). Разумеется d и δ определены лишь на внешней подалгебре $\Lambda(\mathcal{M}) \subset \mathcal{T}(\mathcal{M})$.
2. Производная Ли R -линейна, но не $\mathfrak{F}$ -линейна. Это означает, что, например, отображение $\mathfrak{V}(\mathcal{M}) \times \mathfrak{V}(\mathcal{M}) \rightarrow \mathfrak{V}(\mathcal{M})$, которое определяет производная Ли: $L_X Y = Z$ не является тензором типа $(2, 1)$ (аналогично и для производных Ли тензоров высших валентностей).

9 Координатные формулы для производной Ли

Выведем теперь координатные формулы для производных Ли тензоров различных валентностей. В процессе вывода нам потребуется полезное разложение координатного представления потока ϕ^t в окрестности $t = 0$ с начальной точкой p :

$$x(t) = x(0) + \dot{x}(0)t + \ddot{x}(0)\frac{t^2}{2} + o(t^2) = x_p + X_p t + X(X)_p \frac{t^2}{2} + o(t^2) \quad (36)$$

где в последнем равенстве учтены начальные условия, уравнение потока (31) и его дифференциальное следствие:

$$\ddot{x}(0) = \frac{d}{dt} X_p|_{t=0} = \frac{\partial X}{\partial x^\alpha} \dot{x}^\alpha|_{t=0} = \frac{\partial X}{\partial x^\alpha} X^\alpha \Big|_p X(X)_p.$$

Для компонент дифференциала отображения ϕ_*^t и для обратного $(\phi_*^t)^{-1}$ имеем с точностью до $o(t)$:

$$(\phi_*^t)^\alpha_\beta = \delta^\alpha_\beta + \frac{\partial X^\alpha}{\partial x^\beta} \Big|_p t + o(t); \quad (\phi_*^t)^{-1\alpha}_\beta \approx (\phi_*^{-t})^\alpha_\beta = \delta^\alpha_\beta - \frac{\partial X^\alpha}{\partial x^\beta} \Big|_p t + o(t) \quad (37)$$

или в безындексной форме:

$$(\phi_X^t)_p = \text{id}_p + (\partial X)_p t + o(t); \quad (\phi_X^t)^{-1}_p = \text{id}_p - (\partial X)_p t + o(t).$$

Взаимная обратность приведенных матриц с точностью до $o(t)$ проверяется непосредственным вычислением.

Для скалярной функции $f \in \mathfrak{F}(\mathcal{M})$ определение (35) дает:

$$\begin{aligned} L_X f(p) &= \lim_{t \rightarrow 0} \frac{f(\phi^t(p)) - f(p)}{t} = \lim_{t \rightarrow 0} \frac{f(x_p + tX_p + o(t)) - f(p)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(p) + tX_p(f) + o(t) - f(p)}{t} = X_p(f). \end{aligned} \quad (38)$$

Таким образом, *производная Ли вдоль X от скалярной функции совпадает с дифференцированием этой функции вдоль X .*

Пример. Условие $L_X f = X(f) = 0$ для некоторого заданного векторного поля определяет такую функцию f , у которой поверхности уровня сотканы из интегральных кривых векторного поля X . Действительно, записанное равенство на инфинитезимальном языке выражает в точности тот факт, что при смещениях вдоль потока векторного поля X функция f сохраняет свое значение. Но это и означает, что при движении вдоль потока мы всегда остаемся на некоторой поверхности уровня этой функции. Наоборот, при заданной функции f это уравнение определяет векторное поле, касательное к поверхностям уровня данной функции. Например, для функции $f = f(x^2 + y^2)$ на R^2 всякое векторное поле, удовлетворяющее уравнению $L_X f = 0$, имеет вид $X = Ax\partial_y - Ay\partial_x$, где A — произвольная функция. Это поле является касательным к семейству концентрических окружностей $x^2 + y^2 = R^2$, представляющих собой семейство линий уровня f . $\square$

Для векторного поля $X \in \mathfrak{V}(\mathcal{M})$ определение (35) с учетом (36) и (37) приводит к цепочке равенств:

$$L_X Y_p = \lim_{t \rightarrow 0} \frac{\phi_*^{-1}(Y_{\phi^t(p)}) - Y_p}{t} \lim_{t \rightarrow 0} \frac{(\text{id} - (\partial X)_p t + o(t)) \cdot (Y_p + X(Y)_p t + o(t)) - Y_p}{t} \quad (39)$$

$$\lim_{t \rightarrow 0} \frac{X(Y)_p t - Y(X)_p t + o(t)}{t} = [X, Y]_p,$$

где $\cdot$ означает обычное матричное умножение. Таким образом, *производная Ли векторного поля Y вдоль векторного поля X совпадает с формально введенной ранее скобкой Ли $[X, Y]$* . В координатах производная Ли $L_X Y$ выражается формулой (4). В силу установленных ранее свойств скобки Ли имеют место формулы:

$$L_X Y = -L_Y X; \quad [L_X, L_Y]Z = L_{[X, Y]}Z. \quad (40)$$

Последнее равенство является прямым следствием тождества Якоби и справедливо не только на векторных полях но и на тензорах T произвольной валентности: $[L_X, L_Y]T = L_{[X, Y]}T$.

Пример. Пусть на многообразии $\mathcal{M}$ заданы два векторных поля X и Y и пусть их потоки задаются отображениями ϕ_X и ϕ_Y соответственно. Говорят, что *потоки ϕ_X и ϕ_Y коммутируют*, если

$$(\phi_X^t \circ \phi_Y^s)(p) = (\phi_Y^s \circ \phi_X^t)(p) \quad (41)$$

для всех $t \in R$ и $s \in R$ и для всякой $p \in \mathcal{M}$. Имеет место замечательная теорема.

Теорема 1. Следующие три утверждения эквивалентны:

- 1) Потоки ϕ_X и ϕ_Y коммутируют;
- 2) Семейства интегральных кривых векторных полей X и Y инвариантны относительно действия потоков ϕ_Y и ϕ_X соответственно;
- 3) Скобка Ли $[X, Y] = 0$.

Доказательство. Доказательство проведем по схеме $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$.

$1 \rightarrow 2$. Докажем, что интегральные кривые векторного поля X переводятся потоком ϕ_Y друг в друга. Отображение $\phi_X^t(p)$ при различных t можно рассматривать как интегральную кривую (или ее часть) векторного поля X , проходящую через точку p . Тогда отображение $\phi_Y^s \circ \phi_X^t$ при фиксированном s — это отображение интегральной кривой ϕ_X^t в некоторую новую кривую (увлеченную потоком ϕ_Y). Условие коммутативности потоков (41) утверждает, что полученная кривая совпадает с кривой $\phi_X^t \circ \phi_Y^s(p)$, которая при фиксированном s по

определению является интегральной кривой векторного поля X , проходящей через точку $\phi_Y^s(p)$. Аналогично доказывается отображение друг в друга интегральных кривых поля Y друг в друга потоком ϕ_X .

$2 \rightarrow 3$. Пусть интегральные кривые поля X переводятся потоком ϕ_Y друг в друга. При малых t и s отображения ϕ_X^t и ϕ_Y^s в окрестности некоторых точек $p \in \mathcal{M}$ и $q \in \mathcal{M}$ имеют вид:

$$x(t) = x(p) + tX_p + o(t); \quad x(t) = x(q) + sY_q + o(s).$$

Произвольная точка $x(t) = x(p) + tX_p + o(t)$ интегральной кривой $\phi_X^t(p)$ под действием потока ϕ_Y^s перейдет в точку:

$$x_s(t) = x(p) + tX_p + sY_{x(p)+tX_p+o(t)} + o(s)x(p) + tX_p + sY_p + stX(Y)_p + so(t) + o(s). \quad (42)$$

С другой стороны, из условия того, что интегральная кривая $\phi_X^t(p)$ переводятся потоком ϕ_Y в интегральную кривую $\phi_X^t \circ \phi_Y^s(p)$ имеем:

$$x_s(t) = x_s(0) + tX_{x_s(0)} + o(t) = x(p) + sY_p + tX_p + stY(X)_p + to(s) + o(t) \quad (43)$$

Приравнявая (42) и (43), получаем в порядке st требуемое равенство: $X(Y)_p = Y(X)_p \Rightarrow [X, Y]_p = 0$.

$3 \rightarrow 1$. Как уже было отмечено выше, $\phi_Y^s \circ \phi_X^t(p)$ — кривая, проходящая, через точку $\phi_X^t(p)$. В силу того, что $[X, Y] = -L_Y X = 0$, поток ϕ_Y^s по смыслу производной Ли переводит векторное X в себя. В частности, он переносит касательное к кривой $\phi_X^t(p)$ векторное поле $X|_{\phi_X^t(p)}$ в векторное поле $X|_{\phi_Y^s \circ \phi_X^t(p)}$, которое будет касательным к кривой $\phi_Y^s \circ \phi_X^t(p)$, в силу того, что

$$X|_{\phi_Y^s \circ \phi_X^t(p)} = (\phi_Y^s)_* X|_{\phi_X^t(p)} = (\phi_Y^s)_* (\phi_X^t)_*(d/dt) = (\phi_Y^s \circ \phi_X^t)_*(d/dt).$$

Таким образом, отображение $\phi_Y^s \circ \phi_X^t(p)$ переводит интегральную кривую в интегральную кривую и, в частности, точку p смещает в конечную точку интегральной кривой $\phi_X^t(p)$, а затем переводит ее в конечную точку интегральной кривой $\phi_Y^s \circ \phi_X^t(p)$, проходящей через точку $\phi_Y^s(p)$. Отображение $\phi_X^t \circ \phi_Y^s(p)$ переводит точку p в точку $\phi_Y^s(p)$, а затем смещает ее вдоль интегральной кривой $\phi_X^t \circ \phi_Y^s(p)$ поля X . Но по локальной теореме единственности интегральная кривая векторного поля, проходящая через некоторую точку, единственна. Следовательно, $\phi_Y^s \circ \phi_X^t(p) = \phi_X^t \circ \phi_Y^s(p)$. $\square$

Производную Ли 1-формы $\omega \in \mathfrak{V}^*(\mathcal{M})$ можно вычислить теперь, опираясь на производную Ли векторного поля, правило Лейбница и коммутруемость со сверткой. Учитывая, что значение $\omega(Y)$ на произвольном векторном поле Y есть скаляр, имеем цепочку равенств:

$$L_X \omega(Y) = X\omega(Y) = (L_X \omega)(Y) + \omega(L_X Y) = (L_X \omega)(Y) + \omega([X, Y]),$$

откуда

$$(L_X \omega)(Y) = X\omega(Y) - \omega([X, Y]).$$

Легко проверить, что левая часть зависит от Y $\mathfrak{F}$ -линейно и следовательно, как и должно быть, сама производная Ли $L_X(\omega)$ является 1-формой. Явное вычисление в координатах приводит к формуле:

$$(L_X \omega)_\alpha = X^\beta \partial_\beta \omega_\alpha + \omega_\beta \partial_\alpha X^\beta. \quad (44)$$

Пример. Вычислим для будущих целей производную Ли одной формы $\omega \in \mathfrak{V}^*(R^2)$ следующего вида:

$$\omega = dy - f(x, y) dx.$$

Результат вычислений по формуле (44) имеет вид:

$$(L_X \omega)_1 = -X^\alpha \partial_\alpha f - f \partial_1 X^1 + \partial_1 X^2; \quad (L_X \omega)_2 = -f \partial_2 X^1 + \partial_2 X^2. \quad (45)$$

Аналогичным образом поступим и для определения координатных формул производной Ли произвольного тензора $T \in \mathcal{T}^{r,s}(\mathcal{M})$. В силу правила Лейбница и коммутативности со сверткой имеем:

$$\begin{aligned} L_X(T(X_1, \dots, X_r, \omega_1, \dots, \omega_s)) &= X(T(X_1, \dots, X_r, \omega_1, \dots, \omega_s)) \\ (L_X T)(X_1, \dots, X_r, \omega_1, \dots, \omega_s) &+ T([X, X_1], \dots, \omega_s) + \dots + T(X_1, \dots, L_X \omega_s), \end{aligned}$$

откуда

$$\begin{aligned} (L_X T)(X_1, \dots, X_r, \omega_1, \dots, \omega_s) &= \\ X(T(X_1, \dots, X_r, \omega_1, \dots, \omega_s)) &- T([X, X_1], \dots, \omega_s) + \dots + T(X_1, \dots, L_X \omega_s), \end{aligned}$$

где все производные Ли в правой части уже определены и легко проверяется, что правая часть $\mathfrak{F}$ -линейна по всем своим аргументам, за исключением X . Явное вычисление в координатах приводит к выражению:

$$\begin{aligned} (L_X T)_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta_s} &= X(T_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta_s}) - \partial_{\alpha_1} X^\alpha T_{\alpha \dots \alpha_r}^{\beta_1 \dots \beta_s} - \dots - \partial_{\alpha_r} X^\alpha T_{\alpha_1 \dots \alpha}^{\beta_1 \dots \beta_s} + \\ &\partial_\beta X^{\beta_1} T_{\alpha_1 \dots \alpha_r}^{\beta \dots \beta_s} + \dots + \partial_\beta X^{\beta_s} T_{\alpha_1 \dots \alpha_r}^{\beta_1 \dots \beta}. \end{aligned} \quad (46)$$

Пример. Вычислим для будущих целей производную Ли $L_X g$ ковариантного симметричного тензора $g \in \mathcal{T}^{(2,0)}(\mathcal{M})$. Результат вычислений по формуле (46) принимает вид:

$$(L_X g)_{\alpha\beta} = X(g_{\alpha\beta}) + \partial_\alpha X^\gamma g_{\gamma\beta} + \partial_\beta X^\gamma g_{\alpha\gamma}. \quad (47)$$

Можно было бы стартовать с разложения $g = g_{\alpha\beta}(dx^\alpha \otimes dx^\beta)$. С учетом правила Лейбница и формулы $L_X(dx^\alpha) = \partial_\beta X^\alpha dx^\beta$, вытекающей из (44), приходим к тому же выражению (47). $\square$

10 Применения производной Ли (1): интегрирование дифференциальных уравнений.

В настоящих лекциях мы не ставим своей целью дать исчерпывающий обзор методов и результатов, связанных с дифференциально-геометрическим подходом к интегрированию дифференциальных уравнений. Основная цель настоящего параграфа проиллюстрировать основную идею этого подхода на некоторых простых примерах. Систематическое изложение этих вопросов читатель может найти в обзорной статье [18] или в известных классических книгах [19–21].

Рассмотрим обыкновенное дифференциальное уравнение вида:

$$\frac{dy}{dx} = f(x, y). \quad (48)$$

Его решение — это зависимость $y(x)$, которая превращает это уравнение в тождество. Геометрически на плоскости (x, y) это уравнение задает *поле направлений* (рис.3): каждой точке (x, y) оно приписывает тангенс угла наклона искомой зависимости $y(x)$, равный $f(x, y)$. Таким образом, *решить дифференциальное уравнение (48), это значит по заданному по-*

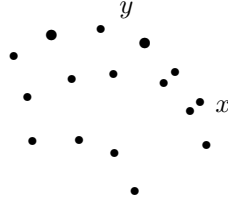


Рис. 3: К геометрической интерпретации дифференциальных уравнений

лю направлений отыскать функцию $y(x)$, поле касательных прямых к которой совпадает с заданным полем направлений. Эту же ситуацию мы можем описать на эквивалентном языке дифференциальной геометрии. Рассмотрим на плоскости R^2 векторное поле $X = \partial_1 + f(x, y)\partial_2$. Нетрудно видеть, что его интегральные кривые, являющиеся решением системы дифференциальных уравнений:

$$\frac{dx}{dt} = X^1 = 1; \quad \frac{dy}{dt} = X^2 = f(x, y),$$

в точности совпадают с решениями системы (48). Этот факт можно выразить условием:

$$\omega(X) = 0, \quad (49)$$

где $\omega = dy - f(x, y)dx$ — 1-форма, ассоциированная с дифференциальным уравнением (48). По сути уравнение (49) выражает факт принадлежности векторного поля X ядру 1-формы ω : $X \in \ker \omega$, где $\ker \omega$ и есть поле направлений дифференциального уравнения (48). В каждой точке $\ker \omega_p$ — это одномерное подпространство касательного пространства $T_p R^2$. Формально ω — это уравнение (48), переписанное в дифференциалах. Наша геометрическая интерпретация как раз и проясняет, что стоит за формальными манипуляциями с дифференциалами.

Пока мы только переформулировали проблему с аналитического языка на геометрический. Теорема локального существования и единственности, которая обычно доказывается на аналитическом языке (с помощью ломаной Эйлера или метода сжимающих отображений) гарантирует нам при некоторых достаточно общих ограничениях на правую часть, что в окрестности некоторой начальной точки решение (интегральная кривая потока векторного поля X) существует и единственно. При этом доказательство теоремы неконструктивно: оно в общем случае не дает ключа к отысканию этого единственного решения. Оказывается, геометрическая формулировка задачи допускает регулярный метод отыскания решений, который в частности, объясняет "механизм работы" всех широко известных приемов интегрирования определенных классов дифференциальных уравнений.

Векторное поле Y называется симметрией дифференциального уравнения (48), если ассоциированная с ним 1-форма ω удовлетворяет соотношению:

$$L_Y \omega = \lambda \omega, \quad (50)$$

где λ — некоторая скалярная функция на R^2 .

С геометрической точки зрения условие симметрии (50) означает, что поток ϕ_Y переводит поле направлений в себя, т.е. сохраняет ядро $\ker \omega$. При этом, если $\lambda = 0$, то ядро в точности переходит в ядро без всяких изменений. Если же $\lambda \neq 0$, то ядро переходит в ядро, но при этом векторы из ядра могут испытывать растяжение или сжатие. Очевидно, что поле направлений дифференциального уравнения при таких растяжениях или сжатиях не меняется и дифференциальное уравнение как геометрический объект остается неизменным.

Предположим, что мы каким-то образом отыскивали симметрию Y для дифференциального уравнения (48). Может иметь место два случая. В первом $\omega(Y) = 0$, т.е. симметрия осуществляет поток вдоль решений и следовательно сама является решением. Очевидно, что отыскание такой симметрии по-существу сводится к отысканию решения и потому не может упростить задачу интегрирования. Во-втором случае $\omega(Y) \neq 0$, т.е. поток симметрии трансверсален к интегральным кривым потока решений. Если при этом задача отыскания такой симметрии оказывается проще, чем задача непосредственного интегрирования исходного уравнения, то используя найденную симметрию можно легко построить интеграл уравнения (48). В общем случае он будет представлять собой неявную функцию $y(x)$, задаваемую соотношением вида $F(x, y) = \text{const}$.

Идея построения такого интеграла в рассматриваемом нами случае заключается в следующем. Интегральные кривые поля симметрии Y отмечают на плоскости R^2 линии, двигаясь вдоль которых дифференциальное уравнение переходит в себя в указанном выше смысле. Другими словами, если ввести на плоскости новую систему координат, в которой одним из двух семейств координатных линий будут линии векторного поля Y , то в дифференциальном уравнении (48) зависимость от этой координаты должна исчезнуть. При этом само уравнение превратится в уравнение с разделяющимися переменными вида

$$\frac{dy'}{dx'} = f(y') \quad \text{или} \quad \frac{dy'}{dx'} = f(x').$$

Практически, решая уравнения характеристик для векторного поля Y , мы будем иметь интеграл вида $g(x, y) = \text{const}$, описывающий семейство его интегральных кривых. В качестве новой координаты u на плоскости R^2 и следует принять эту комбинацию: $u = g(x, y)$.

Проиллюстрируем описанную выше идею на конкретном примере.

Пример. Рассмотрим уравнение вида (48) с правой частью $f(x, y) = \varphi(y/x)$. Непосредственной проверкой по формулам (45) убеждаемся, что $L_Y \omega = \omega$, где $Y = x\partial_1 + y\partial_2$, $\omega = dy - \varphi(y/x)dx$. Следовательно Y — поле симметрий, при этом $\omega(Y) \neq 0$. Интеграл уравнения характеристик

$$\frac{dx}{x} = \frac{dy}{y},$$

имеет вид $y/x = \text{const}$. После введения новой переменной $u = y/x$, исходное дифференциальное уравнение преобразуется к виду:

$$x \frac{du}{dx} + u = f(u),$$

для которого возможно разделение переменных:

$$\ln x + C = \int \frac{du}{f(u) - u}.$$

Аналогичным образом можно рассмотреть случай обобщенно-однородных уравнений вида (48), у которых

$$f(x, y) = \frac{y}{x} \varphi(y^\alpha / x^\beta),$$

где α и β — произвольные вещественные числа. Соответствующий интеграл уравнений характеристик векторного поля симметрии имеет вид:

$$u = y^\alpha / x^\beta = \text{const.} \square$$

Сделаем в конце этого раздела несколько замечаний.

1. Процедура отыскания симметрий оказывается далеко не всегда проще процедуры отыскания решения самого уравнения. Так, например, обстоит дело с общим уравнением Риккати: $y' = \varphi_2(x)y^2 + \varphi_1(x)y + \varphi_0(x)$. В этом смысле симметричный подход не является универсальным. Однако всем случаям разделения переменных в дифференциальном уравнении соответствует существование поля (или полей) симметрии Y .
2. В случае когда $\omega(Y) \neq 0$ имеет место следующая теорема:

1-форма ω , ассоциированная с уравнением (48), допускает интегрирующий множитель вида $1/\omega(Y)$. Другими словами, 1-форма $\Omega = \omega/\omega(Y)$ является в этом случае "полным дифференциалом" dF некоторой функции F .

Справедливость этой теоремы можно проверить непосредственно, вычислив величину $\partial_1\Omega_2 - \partial_2\Omega_1$ и убедившись, что она равна нулю в силу уравнений для поля симметрии Y . Более компактное доказательство, не связанное с переходом к координатам, заключается в применении аппарата внешних дифференциальных форм и теоремы Фробениуса (см. [21]).

3. Обсуждаемые здесь идеи допускают непосредственное обобщение на дифференциальные уравнения высших порядков и уравнения в частных производных. Обобщением формы ω выступает в этих случаях *распределение Картана в соответствующем расслоении джетов* [21].
4. Если у дифференциального уравнения есть несколько симметрий, то они образуют *алгебру симметрий* относительно скобки Ли. Действительно, пусть $L_Y\omega = \lambda_1\omega$ и $L_Z\omega = \lambda_2\omega$. В силу свойства (40) имеет место следующая цепочка равенств:

$$L_{[Y,Z]} = (L_ZL_Y - L_YL_Z)\omega = L_Z(\lambda_1\omega) - L_Y(\lambda_2\omega) = (Z(\lambda_1) - Y(\lambda_2))\omega,$$

откуда следует, что скобка Ли $[Y, Z]$ двух симметрий есть снова симметрия. В некоторых случаях получающаяся симметрия $[Y, Z]$ будет новой, т.е. независимой от Y и Z . Знаменитая теорема Ли гласит: Для полного интегрирования дифференциального уравнения порядка n достаточно существования $n-1$ полей симметрий (n - порядок уравнения), которые образуют разрешимую алгебру Ли [19, 21].

11 Применения производной Ли (2): тензор деформаций и изометрии многообразия

В этом и последующих разделах мы рассмотрим основную область применения производных Ли в геометрии. Речь пойдет об отыскании изометрий различных метрик. Напомним несколько предварительных определений.

Римановой метрикой g на многообразии $\mathcal{M}$ называется симметричное невырожденное положительно-определенное гладкое тензорное поле типа $\mathcal{T}^{(2,0)}(\mathcal{M})$.

Значение $g(X, Y)$ определяет скалярную функцию на $\mathcal{M}$, которая в каждой точке p определяет *скалярное произведение* векторов X_p и Y_p . На языке скалярного произведения свойства метрики, перечисленные в определении, формулируются следующим образом:

1. $g(X, Y) = g(Y, X)$ (симметричность);
2. $g(X, Y) = 0$ для всех $Y \Leftrightarrow X \equiv 0$ (невырожденность);
3. $g(X, X) \geq 0$ для всех X (равенство только для $X = 0$) (положительная определенность);
4. Если $X \in \mathfrak{V}(\mathcal{M})$ и $Y \in \mathfrak{V}(\mathcal{M})$, то $g(X, Y) \in \mathfrak{F}(\mathcal{M})$ (гладкость).

В теории относительности для отражения причинной структуры пространства-времени необходимо отказаться от условия положительной определенности метрики. Метрика g , удовлетворяющая условиям 1, 2, 4, называется *псевдоримановой*.

Введение метрики на многообразии позволяет изучать его локальную внутреннюю геометрию (длины, углы, параллельный перенос). Например, длина* $|X_p|$ вектора X_p и угол $\theta(X_p, Y_p)$ между ненулевыми векторами X_p и Y_p определяются по формуле:

$$|X_p| \equiv \sqrt{g(X_p, X_p)}; \quad \cos \theta(X_p, Y_p) \equiv \frac{g(X_p, Y_p)}{|X_p||Y_p|}. \quad (51)$$

Важным свойством метрики является ее поведение при различных отображениях многообразия в себя. В частности, особый интерес представляют такие преобразования многообразия, при которых метрика остается в определенном смысле неизменной. Такие преобразования (если они существуют) являются абстрактными дифференциально-геометрическими аналогами движений твердого тела в 3-мерном евклидовом пространстве, при которых расстояния между любыми парами точек этого тела остаются неизменными.

Рассмотрим диффеоморфизм $\varphi: \mathcal{M} \rightarrow \mathcal{M}$. Его можно интерпретировать как некоторую конечную деформацию многообразия так, как если бы многообразие представляло собой некоторую деформируемую сплошную среду.

Тензорное поле

$$u \equiv \frac{1}{2}((\varphi)^*g - g) \quad (52)$$

называется тензором конечных деформаций метрики g при диффеоморфизме φ .

Этот тензор, будучи определенным для любой точки $p \in \mathcal{M}$, определяет в ней "степень деформации" метрики. Числовые характеристики этой деформации мы получим, если рассмотрим значения u на элементах какого-нибудь ортонормированного репера $\{e_\alpha(p)\}$ в точке p и применим формулы (51). Так, $u(e_\alpha(p), e_\alpha(p)) \equiv u_{\alpha\alpha}(p)$ будет описывать относительное изменение ϵ_α длины в направлении $(\varphi)_*(e_\alpha)$ в точке $\varphi(p)$ по формуле:

$$\epsilon_\alpha = \sqrt{1 + 2u_{\alpha\alpha}(p)} - 1.$$

*Для избежания недоразумений, везде где не оговорено особо, мы рассматриваем риманову метрику.

Эта формула отнесена к системе координат в точке p . Аналогично, недиагональные компоненты $u_{\alpha\beta}$ ($\alpha \neq \beta$) описывают деформации углов в плоскостях (e_α, e_β) по формуле:

$$\cos \theta'_{\alpha\beta}(p) = \frac{u_{\alpha\beta}(p)}{(1 + \epsilon_\alpha)(1 + \epsilon_\beta)},$$

где θ' — угол между векторами $\varphi_*(e_\alpha(p))$ и $\varphi_*(e_\beta(p))$.

Теперь естественно ввести следующие определения.

Конечная деформация φ называется жесткой в точке p , если $u_p = 0$.

Конечная деформация φ называется изометрией метрики g на $\mathcal{M}$, если $u = 0$ на всем многообразии.

Аналогично тому, как при движении твердого тела в 3-мерном евклидовом пространстве оно в действительности занимает все промежуточные положения между начальным и конечным положениями, мы и в рассматриваемом нами абстрактном случае можем определить непрерывное семейство φ^t деформаций многообразия, параметризованное вещественным параметром t ("параметрическое время"). Очевидно, это семейство описывает некоторый поток на $\mathcal{M}$ и ему, в соответствии с материалом раздела 7, можно сопоставить векторное поле скорости X_φ . Для фиксированной точки p мы имеем семейство конечных тензоров деформаций:

$$u_p^t \equiv \frac{1}{2}((\varphi^t)^*g - g)_p. \quad (53)$$

Разделив правую часть на t , переходя к пределу при $t \rightarrow 0$ и используя определение производной Ли (35), приходим к определению *тензора скоростей деформаций*:

$$\dot{u} \equiv L_{X_\varphi}g,$$

который представляет собой инфинитизмальную версию тензора конечных деформаций.

Имеет место очевидная:

Теорема. Для того, чтобы φ^t было изометрией на $\mathcal{M}$ необходимо и достаточно, чтобы $\dot{u} \equiv 0$.

Доказательство. Необходимость очевидна из определения изометрии. Пусть p — некоторая произвольная фиксированная точка и пусть X_p и Y_p — пара произвольных фиксированных векторов в ней. Рассмотрим изменение величины $g(X_p, Y_p)$ под действием потока:

$$\frac{d}{dt}g((\varphi^t)_*X_p, (\varphi^t)_*Y_p) \frac{d}{dt}(\varphi^t)^*g(X_p, Y_p) = (L_{X_\varphi}g|_{\varphi^t(p)})(X_p, Y_p) = 0.$$

Следовательно длины и углы вдоль потока сохраняются и изометричность потока φ^t очевидна. $\square$

Таким образом, для отыскания изометрий метрики g достаточно найти множество инфинитизмальных изометрий — полей X , для которых выполняются уравнения

$$L_Xg = 0. \quad (54)$$

Эти уравнения называются *уравнениями Киллинга для метрики g* , а их решения X — *векторными полями Киллинга*. Отметим, что поля Киллинга образуют алгебру Ли изометрий

метрики g относительно скобки Ли (доказательство аналогично доказательству пункта 4 конца прошлого раздела). Конечные изометрии будут описываться 1-параметрическими семействами интегральных кривых найденных векторных полей Киллинга.

В качестве первого простейшего примера конкретных изометрий рассмотрим изометрии евклидовой метрики g в R^n . В декартовой системе координат евклидова метрика изображается единичной матрицей, а уравнения Киллинга с учетом формул (45) принимают вид:

$$g_{\alpha\gamma}\partial_\beta X^\gamma + g_{\beta\gamma}\partial_\alpha X^\gamma = 0. \quad (55)$$

Из диагональных уравнений при $\alpha = \beta$ вытекает, что X^α не зависит от x^α . Недиagonальные уравнения принимают вид:

$$\partial_\alpha X^\beta + \partial_\beta X^\alpha = 0$$

для всех пар $\alpha \neq \beta$. Дифференцируя это уравнение по x^α и учитывая, что $\partial_\alpha X^\alpha = 0$ (суммирования нет!), приходим к заключению, что $\partial_{\alpha\alpha}^2 X^\beta = 0$, т.е. все X^β являются линейными функциями координат:

$$X = C \cdot x + A,$$

где C и A — числовая матрица $n \times n$ и числовой n -столбец соответственно. Уравнения Киллинга тождественно удовлетворяются, если матрица коэффициентов C антисимметрична, т.е. $C^T = -C$, где $\bullet^T$ — стандартная операция матричного транспонирования. Полагая поочередно все параметры в C и A кроме одного равными нулю, получаем, таким образом, что *независимыми векторными полями алгебры изометрий евклидовой метрики являются:*

$$\mathfrak{T}_n = \{\partial_\alpha\}_{\alpha=1,\dots,n} \quad (\text{трансляции}) \quad \text{и} \quad \mathfrak{so}(n) = \{x^\alpha \partial_\beta - x^\beta \partial_\alpha\}_{\alpha < \beta} \quad (\text{плоские вращения}),$$

где посредством $\mathfrak{T}_n$ мы обозначили *подалгебру трансляций* алгебры изометрий, а посредством $\mathfrak{so}(3)$ — *подалгебру вращений*. Полученный результат носит общий и исчерпывающий характер: *все непрерывные изометрии евклидовой метрики любого числа измерений исчерпываются трансляциями $\mathfrak{T}_n$ и вращениями $\mathfrak{so}(n)$. При этом $\dim \mathfrak{T}_n = n$, $\dim \mathfrak{so}(3) = n(n-1)/2$, где n — размерность многообразия с евклидовой метрикой.* Отметим, что обозначение $\mathfrak{so}(n)$ происходит из общепринятой в теории групп Ли системы обозначений: $\mathfrak{so}(n)$ является алгеброй Ли группы вращений $SO(n)$, которая получается интегрированием уравнений потоков полей из алгебры Ли $\mathfrak{so}(n)$ (см. пример в разделе 7). В 3-мерном случае алгебра $\mathfrak{so}(3)$ характеризуется следующей системой коммутационных соотношений:

$$[X_{(i)}, X_{(j)}] = \epsilon_{ijk} X_{(k)}. \quad (56)$$

12 Применения производной Ли (3): группа движений пространства Минковского

Аналогично предыдущему случаю вычислим алгебру изометрий плоского псевдоевклидова пространства $\mathcal{M}_{p,q}$. В декартовой системе координат эта метрика имеет наиболее простой вид:

$$g = dx^0 \otimes dx^1 + \dots + dx^p \otimes dx^p - dy^1 \otimes dy^1 - \dots - dy^q \otimes dy^q. \quad (57)$$

Векторное поле симметрии запишем в координатах следующим образом:

$$Z = X^\alpha(x, y) \frac{\partial}{\partial x^\alpha} + Y^\beta(x, y) \frac{\partial}{\partial y^\beta}, \quad \alpha = 1, \dots, p; \quad \beta = 1, \dots, q.$$

Уравнения Киллинга с учетом формул (45) принимают вид (55) с заменой компонент евклидовой метрики на компоненты псевдоевклидовой метрики. Из диагональных уравнений при следует как и ранее, что X^α не зависит от x^α и Y^β не зависит от y^β . Недиagonальные уравнения принимают вид:

$$\frac{\partial X^\beta}{\partial x^\alpha} + \frac{\partial X^\alpha}{\partial x^\beta} = 0; \quad \frac{\partial Y^\beta}{\partial y^\alpha} + \frac{\partial Y^\alpha}{\partial y^\beta} = 0; \quad \frac{\partial X^\beta}{\partial y^\alpha} - \frac{\partial Y^\alpha}{\partial x^\beta} = 0.$$

Путем дифференцирования этих уравнений с учетом условия $\partial X^\alpha / \partial x^\alpha = 0$ и $\partial Y^\beta / \partial y^\beta = 0$ (суммирования нет!) аналогично случаю евклидовой метрики приходим к заключению, что все компоненты Z являются линейными функциями координат:

$$X = B \cdot x + C \cdot y + A, \quad Y = D \cdot x + F \cdot y + G,$$

где A и G — постоянные векторы размерности p и q соответственно, B — постоянная матрица $p \times p$, C — постоянная матрица $p \times q$, D — постоянная матрица $q \times p$, F — постоянная матрица $q \times q$. Уравнения Киллинга тождественно удовлетворяются, если матрицы коэффициентов удовлетворяют следующим условиям:

$$B^T = -B; \quad F^T = -F; \quad C^T = D.$$

Таким образом, *независимыми векторными полями алгебры изометрий псевдоевклидовой метрики являются:*

$$\mathfrak{T}_{p+q} = \{\partial_\alpha\}_{\alpha=1,\dots,p+q} \quad (\text{трансляции}),$$

$$\mathfrak{so}(p) \oplus \mathfrak{so}(q) = \{x^\alpha \partial / \partial x^\sigma - x^\sigma \partial / \partial x^\alpha, y^\beta \partial / \partial y^\gamma - y^\gamma \partial / \partial y^\beta\}_{\alpha < \beta} \quad (\text{плоские вращения}),$$

и

$$\mathfrak{b}(p, q) = \{x^\alpha \partial / \partial y^\beta + y^\beta \partial / \partial x^\alpha\}_{\alpha, \beta} \quad (\text{псевдовращения или бусты}).$$

Отметим, что бусты в общем случае не образуют подалгебры. Как и в случае евклидовой метрики полученный результат носит общий и исчерпывающий характер: *все непрерывные изометрии псевдоевклидовой метрики типа (p, q) исчерпываются трансляциями $\mathfrak{T}_{p+q}$, вращениями $\mathfrak{so}(p) \oplus \mathfrak{so}(q)$ и бустами $\mathfrak{b}(p, q)$. При этом $\dim \mathfrak{T} = n = p + q$, $\dim \mathfrak{so}(p) \oplus \mathfrak{so}(q) = p(p-1)/2 + q(q-1)/2$, $\dim \mathfrak{B} = pq$. Полученные результаты в частном случае $p = 1$, $q = 3$ и $p = 3$, $q = 0$ приводят к хорошо известным фактам псевдоевклидовой 4-мерной геометрии Минковского и 3-мерной евклидовой геометрии.*

13 Применения производной Ли (4): изометрии римановых метрик

Рассмотрим простейшие примеры применения производной Ли в римановой геометрии. Самыми простыми неплоскими римановыми многообразиями являются т.н. *многообразия постоянной кривизны*. Рассмотрим, к примеру, 2-мерную сферу, вложенную стандартным образом в 3-мерное евклидово пространство и метрику, индуцированную на ней объемлющей евклидовой метрикой. К выражению для последней проще всего прийти, рассмотрев евклидову метрику в сферической системе координат:

$$g = dr \otimes dr + r^2(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi).$$

Переходя на единичную сферу посредством соотношения $r = 1$, получаем метрику сферы:

$$g^{(S^2)} = d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi. \quad (58)$$

Вычислим алгебру симметрий этой метрики. Уравнения Киллинга $L_X g^{(S^2)} = 0$ записываются по формуле (45) следующим образом ($x^1 = \theta, x^2 = \varphi$):

$$\partial_\theta X^\theta = 0; \quad X^\theta + \tan \theta \partial_\varphi X^\varphi = 0; \quad \sin^2 \theta \partial_\theta X^\varphi + \partial_\varphi X^\theta = 0.$$

Из первого уравнения следует, что X^θ зависит только от φ : $X^\theta = F(\varphi)$. Из второго уравнения следует, что $X^\varphi = -\cot \theta \int F d\varphi + \psi(\theta)$, где ψ — пока произвольная функция. Подставляя эти представления в третье уравнение, приходим к уравнению на F и ψ :

$$F' + \int F d\varphi + \psi' \sin^2 \theta = 0.$$

Отсюда для стандартного условия разделения следует, что $\psi = D \cot \theta + C$, $\int F d\varphi = A \sin \varphi + B \cos \varphi + D$, где A, B, C, D — произвольные константы интегрирования. Таким образом, общее поле симметрии имеет вид:

$$X = (A \cos \varphi - B \sin \varphi) \partial_\theta - \cot \theta (A \sin \varphi + B \cos \varphi + C) \partial_\varphi. \quad (59)$$

Полагая поочередно пары коэффициентов из набора A, B, C равными нулю, приходим к следующим независимым полям симметрии сферы:

$$X_{(1)} = \cos \varphi \partial_\theta - \cot \theta \sin \varphi \partial_\varphi; \quad X_{(2)} = -\sin \varphi \partial_\theta - \cot \theta \cos \varphi \partial_\varphi; \quad X_{(3)} = \partial_\varphi.$$

Как и следовало ожидать из общих интуитивных соображений поля $X_{(i)}$ образуют алгебру Ли $\mathfrak{so}(3)$ 3-мерной группы вращений $SO(3)$, что проверяется непосредственной проверкой выполнимости соотношений (56).

В общей теории относительности 4-мерное пространство-время называется *сферически-симметричным*, если его метрика g имеет алгебру симметрии $\mathfrak{so}(3)$. В сферических координатах, в которых поля этой алгебры имеют вид (59), общая сферически симметричная метрика будет иметь вид:

$$g = A(t, r) dt \otimes dt - B(t, r) dr \otimes dr + f(t, r) (dt \otimes dr + dr \otimes dt) - C(t, r) g^{(S^2)}, \quad (60)$$

где A, B, C, f — произвольные функции временной и радиальной координат t и r . Пространство-время называется *сферически-симметричным статическим*, если его алгебра симметрии помимо $\mathfrak{so}(3)$ содержит еще одно времениподобное (т.е. удовлетворяющее условию $g(X, X) > 0$) векторное поле симметрии, коммутирующее с элементами алгебры $\mathfrak{so}(3)$. Используя свободу координатных преобразований:

$$t = t(\tau, \rho); \quad r = r(\tau, \rho),$$

сохраняющих общий вид метрики (60), ее всегда можно привести к виду:

$$g = \bar{A}(\tau, \rho) d\tau \otimes d\tau - \bar{B}(\tau, \rho) d\rho \otimes d\rho - \rho^2 g^{(S^2)}$$

в случае общей сферической симметрии и к виду:

$$g = \bar{A}(\rho)d\tau \otimes d\tau - \bar{B}(\rho)d\rho \otimes d\rho - \rho^2 g^{(S^2)}$$

в случае сферически-симметричного статического пространства-времени. Здесь $\bar{A}$ и $\bar{B}$ — некоторые функции новых переменных τ, ρ , связанные с исходными A и B посредством тензорного закона преобразований.

Замечательным фактом эйнштейновской общей теории относительности является *существование дополнительного поля Киллинга как следствие сферической симметрии*. Другими словами, (при некоторых дополнительных предположениях: в обсуждаемом нами случае — для пустого пространства-времени) *все сферически-симметричные решения уравнений Эйнштейна автоматически являются статическими и сводятся к знаменитой метрике Шварцшильда для черной дыры* (теорема Биркгоффа):

$$g = \left(1 - \frac{r_g}{r}\right) dt \otimes dt - \left(1 - \frac{r_g}{r}\right)^{-1} dr \otimes dr - r^2 g^{(S^2)},$$

где $r_g = 2GM/c^2$ — гравитационный радиус черной дыры, связанный с ее полной массой M . Метрика черной дыры и ее следствия являются основой для современного раздела общей теории относительности — *физики черных дыр* [22].

Факт допустимости любых гладких и обратимых замен координат в общей теории относительности усложняет выявление физических свойств, вытекающих из римановой метрики пространственно-временного многообразия. В частности, может оказаться так (и часто оказывается!), что одна и та же метрика имеет совершенно различный и неузнаваемый вид в разных системах координат. Это обстоятельство можно наблюдать даже в плоском пространстве времени, переходя в нем к различным криволинейным координатам. Однако в плоском пространстве времени у нас существует класс декартовых систем координат, в которых метрика имеет стандартный и самый простой вид — это утверждение можно принять за определение плоского пространства. Для римановых метрик в общем случае не существует никаких предпочтительных систем координат и поэтому для их классификации и в частности установления их тождества или различия необходимо использовать другие методы, не связанные с координатными конструкциями. Одним из подходов к инвариантной классификации римановых метрик является их классификация по симметриям. Поскольку уравнения Киллинга имеют общековариантный тензорный характер, то поля симметрий имеются или отсутствуют независимо от выбора системы координат и (при их наличии) их алгебра должна быть одной и той же в разных системах координат с точностью до линейных замен базиса в алгебре симметрий. Таким образом, *для того чтобы две метрики были эквивалентны, (т.е. чтобы они переводились друг в друга преобразованием координат) необходимо, чтобы их алгебры симметрий совпадали*. Иногда, когда алгебра симметрий является достаточно богатой, это условие является и достаточным, иногда требуется привлекать для сравнения другие инвариантные свойства многообразия (например алгебраические типы тензоров Вейля и Риччи).

Симметричный подход часто кладут в основу определения того или иного пространства-времени. Этот метод работает даже в тех случаях, когда физический смысл получающейся геометрии (т.е. ее возможные материальные источники и их свойства) не ясен. Идею "симметричного конструктора" мы проиллюстрируем на следующем примере. Пусть нас интересует

пространство-время, обладающее следующей алгеброй симметрий:

$$X_{(0)} = \partial_t, \quad X_{(1)} = \partial_x, \quad X_{(2)} = \partial_y, \quad X_{(3)} = \partial_x + y\partial_y + \frac{1}{2}(\sqrt{3}t - z)\partial_z - \frac{1}{2}(t + \sqrt{3}z)\partial_t.$$

Мы могли прийти к такой алгебре из каких-то совершенно посторонних к ОТО соображений или даже рассмотреть ее "просто так" из любопытства. Другими словами, нас интересует решение системы уравнений Киллинга:

$$L_{X_{(i)}}g = 0 \quad (i = 1, 2, 3, 4),$$

но теперь уже не относительно компонент полей $X_{(i)}$ — ведь они нам известны заранее, — а относительно компонент метрики. Если наложить дополнительное требование, чтобы искомая метрика удовлетворяла уравнениям Эйнштейна в пустоте, то оказывается (проверка в принципиальном плане не сложна, но громоздка), искомая метрика единственна и имеет вид:

$$g = k^2[e^x(\cos \sqrt{3}x(dt \otimes dt - dz \otimes dz) + \sin \sqrt{3}x(dt \otimes dz + dz \otimes dt)) - dx \otimes dx - e^{-2x}dy \otimes dy],$$

где $k = \text{const}$. Много интересной информации о классификации точных решений уравнений Эйнштейна по их симметриям приведено в книге [23].

14 Применения производной Ли (5): группа конформных симметрий евклидовой и псевдоевклидовой метрик ($n > 2$)

Рассмотрим теперь ситуацию, которая слегка обобщает непрерывные симметрии и, в частности, изометрии. Пусть T — тензорное поле некоторой валентности и пусть X — векторное поле, поток ϕ_X^t которого действует на тензорное поле T так, что выполняется равенство:

$$(\phi_X^t)^*T_{\phi^t(p)} = f(t, p)T_p,$$

где $f(t, p)$ — некоторая гладкая функция $R \times \mathcal{M} \rightarrow R$. В случае, если $f \equiv 1$, мы имеем стандартную лиеву симметрию тензора T : $L_X T = 0$. В случае $f(t, p) \neq 1$, производная Ли тензора T имеет вид:

$$L_X T \equiv \lim_{t \rightarrow 0} \frac{(\phi_X^t)^*T_{\phi^t(p)} - T_p}{t} T_p \lim_{t \rightarrow 0} \frac{f(t, p) - 1}{t} = \Phi T, \quad (61)$$

где $\Phi = \partial_t f|_{t=0}$. Симметрия тензорного поля T , описываемая уравнением (61), называется *конформной симметрией*, а векторное поле X называется *полем конформной симметрии тензорного поля T* . Конформная симметрия тензорных поле обобщает лиеву симметрию, которая описывается более простым уравнением $L_X T = 0$. Кроме преобразований лиевой симметрии (при $f = 1$ и $\Phi = 0$), конформная симметрия включает такие преобразования, которые оставляют неизменными все характеристические направления тензорного поля в каждой точке (например собственные направления симметричных тензоров валентности два), но может изменять длины характеристических векторов и значения характеристических скаляров (например длин собственных векторов и собственных значений симметричных тензоров валентности два).

С физической точки зрения конформные преобразования осуществляют переходы между системами отсчета в ситуациях, когда единичные масштабы по каким-либо причинам установить невозможно или они несущественны для круга решаемых задач. Примером подобной ситуации является мир световых сигналов в СТО или в ОТО: 4-мерные длины, измеренные вдоль световых лучей, равны нулю. Полной группой симметрии такого мира будет не группа Пуанкаре, а конформная группа, оставляющая инвариантными светоподобные интервалы и световые конуса, а не 4-мерные псевдоевклидовы интервалы. Разумеется такая группа будет включать в себя группу Пуанкаре. Преобразования из группы конформных симметрий, не принадлежащие группе Пуанкаре (или соответствующей группе изометрий для $\mathcal{M}_{p,q}$), будем называть *собственными конформными преобразованиями*.

В настоящем разделе мы вычислим такие собственные конформные преобразования для пространства $\mathcal{M}_{p,q}$ при $p + q > 2$. Для метрики (57) конформные уравнения Киллинга (61) для поля конформной симметрии $Z = X^\alpha(x, y)\partial_\alpha + Y^\beta(x, y)\partial_\beta$ принимают вид:

$$2\frac{\partial X^\alpha}{\partial x^\alpha} = \Phi; \quad 2\frac{\partial Y^\beta}{\partial y^\beta} = \Phi; \quad \text{суммирования нет!} \quad (62)$$

$$\frac{\partial X^\alpha}{\partial x^\sigma} + \frac{\partial X^\sigma}{\partial x^\alpha} = 0; \quad \frac{\partial Y^\beta}{\partial x^\gamma} + \frac{\partial Y^\gamma}{\partial y^\beta} = 0; \quad \frac{\partial X^\alpha}{\partial y^\beta} + \frac{\partial Y^\beta}{\partial x^\alpha} = 0. \quad (63)$$

Вычисляя смешанные вторые производные $\partial/\partial x^\alpha \partial x^\sigma$, $\partial/\partial y^\beta \partial y^\gamma$ и $\partial/\partial x^\alpha \partial y^\beta$ от левых частей соответственно первого, второго и третьего уравнений (63) с учетом уравнений (62), приходим к условиям на Φ :

$$\frac{\partial^2 \Phi}{\partial x^\alpha \partial x^\alpha} - \frac{\partial^2 \Phi}{\partial y^\beta \partial y^\beta} = 0; \quad \frac{\partial^2 \Phi}{\partial x^\alpha \partial x^\alpha} + \frac{\partial^2 \Phi}{\partial x^\sigma \partial x^\sigma} = 0; \quad \frac{\partial^2 \Phi}{\partial y^\beta \partial y^\beta} + \frac{\partial^2 \Phi}{\partial y^\gamma \partial y^\gamma} = 0.$$

Из этих условий непосредственно следует, что функция Φ линейна по переменным x и y и может быть представлена в виде:

$$\Phi = 2(A \cdot x) + 2B \cdot y + 2d, \quad (64)$$

где A и B — соответственно числовые p и q — векторы, d — вещественная константа, а точка означает евклидово скалярное произведение. Интегрируя теперь уравнения (62) с правой частью в виде (64), мы получаем следующее представление компонент поля конформной симметрии:

$$X^\alpha = (A \cdot x)'x^\alpha + \frac{A^\alpha(x^\alpha)^2}{2} + (B \cdot y)x^\alpha + Dx^\alpha + \varphi^\alpha; \quad (65)$$

$$Y^\beta = (B \cdot y)'y^\beta + \frac{B^\beta(y^\beta)^2}{2} + (A \cdot x)y^\beta + Dy^\beta + \psi^\beta, \quad (66)$$

где φ^α и ψ^β — семейства пока еще неопределенных функций, удовлетворяющих условиям: $\partial\varphi^\alpha/\partial x^\alpha = 0$, $\partial\psi^\beta/\partial y^\beta = 0$ (суммирования по повторяющимся индексам нет!). Штрих у скалярного произведения обозначает его неполный характер: в сумме произведений компонент скалярного произведения слагаемое с номером α или с номером β отсутствует. Подставляя

полученный общий вид полей в оставшиеся уравнения (63), приходим к следующим уравнениям на неизвестные функции φ^α и ψ^β :

$$\frac{\partial \varphi^\alpha}{\partial x^\sigma} + \frac{\partial \varphi^\sigma}{\partial x^\alpha} = -(A^\sigma x^\alpha + A^\alpha x^\sigma); \quad \frac{\partial \psi^\beta}{\partial y^\gamma} + \frac{\partial \psi^\gamma}{\partial y^\beta} = -(B^\gamma y^\beta + B^\beta y^\gamma); \quad \frac{\partial \varphi^\alpha}{\partial y^\beta} + \frac{\partial \psi^\beta}{\partial x^\alpha} = A^\alpha y^\beta + B^\beta x^\alpha. \quad (67)$$

Ввиду того, что нас интересуют только собственные конформные преобразования, достаточно найти только частные решения неоднородных уравнений (67), опуская общие решения однородных уравнений, отвечающих за уже найденные ранее изометрии. Искомые частные решения имеют вид:

$$\varphi^\alpha = -\frac{A^\alpha}{2}(s^2 - (x^\alpha)^2); \quad \psi^\beta = \frac{B^\beta}{2}(s^2 + (y^\beta)^2), \quad (68)$$

где $s^2 \equiv \sum_{\sigma=1}^p (x^\sigma)^2 - \sum_{\gamma=1}^q (y^\gamma)^2$. Окончательно, подставляя выражения (68) в (65) и (66) после некоторых упрощений приходим к следующему выражению компонент поля собственных конформных симметрий пространства $\mathcal{M}_{p,q}$:

$$X^\alpha = (A \cdot x)x^\alpha - \frac{A^\alpha}{2}s^2 + (B \cdot y)x^\alpha + Dx^\alpha; \quad (69)$$

$$Y^\beta = (B \cdot y)y^\beta + \frac{B^\beta}{2}s^2 + (A \cdot x)y^\beta + Dy^\beta. \quad (70)$$

Проверьте самостоятельно, что собственные конформные симметрии образуют алгебру относительно скобки Ли. Из формул (69)-(70) легко подсчитать ее размерность: она равна числу независимых произвольных постоянных параметров, входящих в выражения для поля симметрии. Таким образом размерность алгебры собственных конформных симметрий $\mathcal{M}_{p,q}$ равна $n_C = p + q + 1$. Для пространства Минковского мы имеем $n_C = 5$, что в совокупности с 10-параметрической группой Пуанкаре дает *полную 15-параметрическую конформную группу* $\mathcal{M}_{1,3}$. Отметим, что поле вида (69)-(70) с $A = B = 0$ и $D \neq 0$ отвечает за однородные растяжения пространства-времени. Поля с $A \neq 0$ и (или) $B \neq 0$ — это нелинейные преобразования пространства $\mathcal{M}_{p,q}$, включающие инверсию относительно единичных метрических сфер пространств $\mathcal{M}_{p,q}$. Отметим, что классический координатный подход к исследованию конформных симметрий выглядит существенно более громоздко [24].

15 Применения производной Ли (6): группа конформных симметрий евклидовой и псевдоевклидовой плоскости.

Как мы покажем в этом разделе, случай $n = 2$ является особым как в евклидовом, так и в псевдоевклидовом случае. Несмотря на то, что сами уравнения конформной симметрии при $n = 2$ являются частным случаем уравнений (62)-(63), их решения существенно богаче и допускают произвольные функции. Действительно, в двумерном случае для метрики вида:

$$g = dx^1 \otimes dx^1 + \epsilon dx^2 \otimes dx^2 \quad (\epsilon = \pm 1),$$

уравнения (62)-(63) принимают вид:

$$\partial_1 \xi^1 = \frac{\phi}{2}; \quad \partial_2 \xi^2 = \frac{\phi}{2}; \quad \partial_2 \xi^1 + \epsilon \partial_1 \xi^2 = 0. \quad (71)$$

Дифференцируя последнее уравнение последовательно по x^1 и x^2 и пользуясь первыми двумя уравнениями, приходим к условию на функцию ϕ :

$$(\partial_1^2 + \epsilon \partial_2^2)\phi = 0. \quad (72)$$

Воспользуемся тем обстоятельством, что в двумерном случае существует потенциал (сопряженная функция) χ , определяемый условиями:

$$\partial_1 \phi = \partial_2 \chi; \quad \partial_2 \phi = -\epsilon \partial_1 \chi.$$

С учетом этих определений, условие (72) выполняется тождественно. Непосредственной проверкой можно убедиться, что решения уравнений (71) можно записать через функции ϕ и χ следующим образом:

$$\begin{aligned} \xi^1 &= \frac{1}{2} \int_{(x_0^1, x_0^2)}^{(x^1, x^2)} \phi(x^1, x^2) dx^1 - \frac{\epsilon}{2} \int_{(x_0^1, x_0^2)}^{(x_0^1, x^2)} \chi(x_0^1, x^2) dx^2 + Cx^2; \\ \xi^2 &= \frac{1}{2} \int_{(x^1, x_0^2)}^{(x^1, x^2)} \phi(x^1, x^2) dx^2 + \frac{1}{2} \int_{(x_0^1, x_0^2)}^{(x^1, x_0^2)} \chi(x^1, x_0^2) dx^1 - \epsilon Cx^1. \end{aligned}$$

Здесь (x_0^1, x_0^2) — произвольная начальная точка, а интегрирование выполняется вдоль любой кривой, соединяющей точки на нижних и верхних пределах интегрирования. Таким образом, общее поле конформной симметрии евклидовой и псевдоевклидовой плоскостей содержат произвольные гармонические или псевдогармонические (волновые) функции.

Конечные конформные преобразования более наглядно записываются: в евклидовом случае на языке комплексной переменной z :

$$g = \frac{1}{2}(dz \otimes d\bar{z} + d\bar{z} \otimes dz) \xrightarrow{z=f(w)} g^f \equiv \frac{|f(w)|^2}{2}(dw \otimes d\bar{w} + d\bar{w} \otimes dw),$$

где $f(z)$ — произвольная аналитическая функция; в псевдоевклидовом случае в изотропных координатах $\{\xi^\alpha\}_{\alpha=1,2}$:

$$g = \frac{1}{2}(d\xi^1 \otimes d\xi^2 + d\xi^2 \otimes d\xi^1) \xrightarrow{\xi^\alpha=f^\alpha(\eta^\alpha)} g^f \equiv \frac{f^1(\eta^1)f^2(\eta^2)}{2}(d\eta^1 \otimes d\eta^2 + d\eta^2 \otimes d\eta^1).$$

В таком представлении бесконечномерность группы конформных симметрий двумерных евклидовой и псевдоевклидовой метрик очевидна.

16 Применения производной Ли (7): группа изометрий однородных кубических метрик

В качестве примера, в достаточной мере не отраженного в классической литературе, исследуем изометрии однородных кубических метрик вида:

$$G = G_{\alpha\beta\gamma} dx^\alpha \otimes dx^\beta \otimes dx^\gamma, \quad (73)$$

где $G_{\alpha\beta\gamma}$ — симметричная числовая кубическая матрица. Пространства с метриками такого рода традиционно относят к классу *финслеровых пространств*, а метрику G называют соответственно *финслеровой метрикой* [25]. Введем следующие обозначения:

$$G_{\alpha\alpha\alpha} = A_\alpha; \quad G_{122} = B_1; \quad G_{133} = B_2; \quad G_{233} = B_3; \quad (74)$$

$$G_{112} = C_1; \quad G_{113} = C_2; \quad G_{223} = C_3; \quad G_{123} = F,$$

где все A_α , B_β , C_γ и F — постоянные.

Не все компоненты (74) метрики G одновременно геометрически значимы. Представление (74) инвариантно относительно выбора координат среди класса аффинно-эквивалентных систем, в которых компоненты метрики остаются постоянными. Матрица невырожденного аффинного преобразования в R^3 имеет в общем случае 9 независимых компонент и ими можно распорядиться таким образом, чтобы обратить в нуль 8 из 10 компонент метрики G , поскольку уравнения вида

$$G'_{\alpha\beta\gamma} = G_{\rho\sigma\delta} h_\alpha^\rho h_\beta^\sigma h_\gamma^\delta = 0 \quad (75)$$

однородны по компонентам h . Ввиду сложного характера системы уравнений (75), конкретный набор восьми коэффициентов, которые можно обратить в нуль, сложным образом зависит от исходных значений компонент метрики. В частности, в некоторых вырожденных случаях исходной метрики в нуль можно обратить 9 из 10 компонент.

Отсюда следует два важных для нашего исследования вывода:

1. Для полного исследования достаточно исследовать метрики с небольшим числом отличных от нуля коэффициентов;
2. Следует перебрать всевозможные случаи сочетания небольшого числа отличных от нуля коэффициентов;

Мы показываем, что в действительности можно ограничиться метриками G с числом отличных от нуля коэффициентов, не превышающим 6.

Число отличных от нуля коэффициентов однородной метрики будет задавать *аффинный тип* $\tau(G)$ однородной метрики G . Отметим, что аффинный тип метрики G зависит от выбора аффинной системы координат. Инвариантной характеристикой, не зависящей от выбора системы координат, является *точный аффинный тип*:

$$\tau_0(G) \equiv \min_{\text{Aff}(R^3)} \tau(G),$$

где $\text{Aff}(R^3)$ — класс аффинных систем координат в R^3 , связанных невырожденными аффинными преобразованиями. Назовем две однородных финслеровых метрики G_1 и G_2 *эквивалентными*: $G_1 \sim G_2$, если существует такая аффинная система координат в R^3 , в которой компоненты G_1 и G_2 попарно совпадают. Очевидно, для эквивалентных метрик $G_1 \sim G_2$ может быть $\tau(G_1) \neq \tau(G_2)$, но их точные аффинные типы должны совпадать: $\tau_0(G_1) = \tau_0(G_2)$. Совпадение точных аффинных типов двух метрик, однако, не является достаточным условием их эквивалентности, поскольку, вообще говоря, компоненты, которые входят в минимальное множество отличных от нуля, для этих метрик могут быть различными.

После этих общих соображений перейдем к анализу симметрий. Система уравнений Киллинга в общем случае (74) принимает вид:

$$\left\{ \begin{array}{l} 3A_1\partial_1X^1 + 3C_1\partial_1X^2 + 3C_2\partial_1X^3 = 0; \\ 2C_1\partial_1X^1 + 2B_1\partial_1X^2 + 2F\partial_1X^3 + A_1\partial_2X^1 + C_1\partial_2X^2 + C_2\partial_2X^3 = 0; \\ B_1\partial_1X^1 + A_2\partial_1X^2 + C_3\partial_1X^3 + 2C_1\partial_2X^1 + 2B_1\partial_2X^2 + 2F\partial_2X^3 = 0; \\ 3B_1\partial_2X^1 + 3A_2\partial_2X^2 + 3C_3\partial_2X^3 = 0; \\ 2C_2\partial_1X^1 + 2F\partial_1X^2 + 2B_2\partial_1X^3 + A_1\partial_3X^1 + C_1\partial_3X^2 + C_2\partial_3X^3 = 0; \\ F\partial_1X^1 + C_3\partial_1X^2 + B_3\partial_1X^3 + C_2\partial_2X^1 + F\partial_2X^2 + B_2\partial_2X^3 + C_1\partial_3X^1 + \\ + B_1\partial_3X^2 + F\partial_3X^3 = 0; \\ 2F\partial_2X^1 + 2C_3\partial_2X^2 + 2B_3\partial_2X^3 + B_1\partial_3X^1 + A_2\partial_3X^2 + C_3\partial_3X^3 = 0; \\ B_2\partial_1X^1 + B_3\partial_1X^2 + A_3\partial_1X^3 + 2C_2\partial_3X^1 + 2F\partial_3X^2 + 2B_2\partial_3X^3 = 0; \\ B_2\partial_2X^1 + B_3\partial_2X^2 + A_3\partial_2X^3 + 2F\partial_3X^1 + 2C_3\partial_3X^2 + 2B_3\partial_3X^3 = 0; \\ 3B_2\partial_3X^1 + 3B_3\partial_3X^2 + 3A_3\partial_3X^3 = 0. \end{array} \right. \quad (76)$$

Рассмотрим последовательно все частные случаи общих метрик с различными $\tau(G)$. Там, где это возможно, мы, не оговаривая особо, используем свободу выбора масштабов координат для превращения соответствующих коэффициентов в ± 1 (канонический вид). Кроме того, не рассматриваем отдельно случаев, которые отличаются друг от друга лишь перестановкой координат. Наконец, мы всюду отбрасываем постоянные векторные поля, образующие подалгебру трансляций алгебры симметрий метрики G , существование которой очевидно и исследуем симметрии, отличные от трансляций. Мы будем называть их *нетривиальными симметриями* однородных финслеровых метрик.

16.1 Метрики с $\tau(G) = \tau_0(G) = 1$ (3 типа)

В описании случаев указываются только отличные от нуля компоненты метрики, а все остальные компоненты подразумеваются равными нулю. При этом здесь и далее мы указываем только те случаи, в которых имеются нетривиальные симметрии.

1. $F \neq 0$. Метрика принимает следующий канонический вид:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3), \quad (77)$$

где $\hat{\mathcal{S}}$ — оператор симметризации тензорного произведения. Она называется *метрикой Бервальда-Моора*. Нетривиальная алгебра симметрий этой метрики представлена следующими векторными полями [26, 27]:

$$X_1 = x^1\partial_1 - x^2\partial_2; \quad X_2 = x^1\partial_1 - x^3\partial_3,$$

описывающим согласованные (унимодулярные) дилатации координатных осей. Отметим, что эта алгебра абелева.

2. $B_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^1 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^1. \quad (78)$$

Алгебра ее нетривиальных симметрий бесконечномерна:

$$X = x^2 \partial_2 - 2x^1 \partial_1 + f(x^1, x^2, x^3) \partial_3,$$

где f — произвольная гладкая функция трех переменных, и соответствует унимодулярным дилатациям пары координат.

3. $A_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1. \quad (79)$$

Алгебра ее нетривиальных симметрий бесконечномерна:

$$X = f_2(x^1, x^2, x^3) \partial_2 + f_3(x^1, x^2, x^3) \partial_3,$$

где f_2, f_3 — произвольные гладкие функции трех переменных.

Этими случаями, по существу, исчерпывается класс метрик с $\tau_0(G) = 1$. Отметим, что метрики (78)-(79) геометрически вырождены как 3-мерные метрики, поскольку описываются всего двумя элементами 3-мерного базиса 1-форм $\{dx^1, dx^2, dx^3\}$.

16.2 Метрики с $\tau(G) = 2$ (9 типов)

1. $F \neq 0, A_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (80)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - x^3 \partial_3.$$

2. $F \neq 0, B_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^1 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^1 + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (81)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^1 \partial_1 - (x^3 + x^2/2) \partial_3; \quad x^2 \partial_2 - (x^3 + x^2) \partial_3.$$

3. $A_1 \neq 0, B_3 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^3 \otimes dx^3 + dx^3 \otimes dx^2 \otimes dx^3 + dx^3 \otimes dx^3 \otimes dx^2. \quad (82)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - (x^3/2) \partial_3.$$

4. $A_1 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^1 \otimes dx^1 \otimes dx^3 + dx^1 \otimes dx^3 \otimes dx^1 + dx^3 \otimes dx^1 \otimes dx^1. \quad (83)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^1 \partial_1 - (2x^2 + x^1) \partial_2.$$

5. $B_1 \neq 0, B_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^1 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^1 \pm (dx^1 \otimes dx^3 \otimes dx^3 + dx^3 \otimes dx^1 \otimes dx^3 + dx^3 \otimes dx^3 \otimes dx^1). \quad (84)$$

Алгебра нетривиальных симметрий 2-мерна:

$$X_1 = x^1 \partial_1 - (x^2/2) \partial_2 - (x^3/2) \partial_3; \quad X_2 = x^3 \partial_2 \mp x^2 \partial_3.$$

6. $B_1 \neq 0, B_3 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^1 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^1 + dx^2 \otimes dx^3 \otimes dx^3 + dx^3 \otimes dx^3 \otimes dx^2 + dx^3 \otimes dx^2 \otimes dx^3. \quad (85)$$

Алгебра нетривиальных симметрий 2-мерна:

$$X_1 = -2x^1 \partial_1 + x^2 \partial_2 - (x^3/2) \partial_3; \quad X_2 = x^3 \partial_1 - (x^2/2) \partial_3.$$

7. $B_1 \neq 0, C_3 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^1 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^1 + dx^3 \otimes dx^2 \otimes dx^2 + dx^2 \otimes dx^3 \otimes dx^2 + dx^2 \otimes dx^2 \otimes dx^3. \quad (86)$$

Алгебра нетривиальных симметрий бесконечномерна:

$$X_1 = x^2 \partial_2 - 2(x^1 + x^3) \partial_3; \quad X_2 = f(x^1, x^2, x^3)(\partial_1 - \partial_3),$$

где f — произвольная гладкая функция трех переменных.

8. $A_1 \neq 0, A_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^2 \otimes dx^2 \quad (87)$$

Алгебра нетривиальных симметрий бесконечномерна:

$$X = f(x^1, x^2, x^3) \partial_3,$$

где f — произвольная гладкая функция трех переменных.

9. $A_1 \neq 0, B_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 \pm \hat{S}(dx^1 \otimes dx^2 \otimes dx^2) \quad (88)$$

Алгебра нетривиальных симметрий бесконечномерна:

$$X = f(x^1, x^2, x^3) \partial_3,$$

где f — произвольная гладкая функция трех переменных.

16.3 Метрики с $\tau(G) = 3$ (13 типов)

В большей части случаев симметрии тривиальны. Нетривиальные симметрии имеются лишь в следующих тринадцати типах.

1. $F \neq 0, A_1 \neq 0, B_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 \pm \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (89)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - (x^3 \pm x^2) \partial_3.$$

2. $F \neq 0, A_1 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (90)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - (x^3 + x^1/2) \partial_3.$$

3. $F \neq 0, B_1 \neq 0, B_2 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) \pm \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (91)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^1 \partial_1 \pm (x^3/2) \partial_2 - (x^3 + x^2/2) \partial_3; \quad X_2 = (x^2 \pm x^3) \partial_2 - (x^2 + x^3) \partial_3.$$

4. $F \neq 0, B_1 \neq 0, B_3 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^2 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (92)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^1 + x^3) \partial_1 - (x^3 + x^2/2) \partial_3.$$

5. $F \neq 0, B_1 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (93)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^2 \partial_2 - (x^2 + x^3) \partial_3; \quad X_2 = x^1 \partial_1 - (x^3 + x^2/2 + x^1) \partial_3.$$

6. $F \neq 0, B_1 \neq 0, C_3 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^2 \otimes dx^2 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3). \quad (94)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + x^2/2) \partial_3 - (x^1 + x^2/2) \partial_1.$$

7. $A_1 \neq 0, A_2 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^2 \otimes dx^2 + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3). \quad (95)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^1 \partial_1 - (x^1 + 2x^3) \partial_3.$$

8. $A_1 \neq 0, B_1 \neq 0, B_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + \epsilon_1 \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \epsilon_2 \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3), \quad (96)$$

где $\epsilon_1 = \pm 1, \epsilon_2 = \pm 1$ — независимые знаковые множители. Алгебра нетривиальных симметрий одномерна:

$$X = x^3 \partial_2 - \epsilon_1 \epsilon_2 x^2 \partial_3.$$

9. $A_1 \neq 0, B_1 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 \pm \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3), \quad (97)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^1 \partial_1 - (x^2/2) \partial_2 - 2x^3 \partial_3; \quad X_2 = \mp (x^1/2) \partial_2 + x^2 \partial_3.$$

10. $A_1 \neq 0, B_1 \neq 0, C_3 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 \pm \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^2 \otimes dx^2 \otimes dx^3), \quad (98)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - 2(x^3 \pm x^1) \partial_3.$$

11. $B_1 \neq 0, B_2 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) \pm \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^2), \quad (99)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^3 \partial_2 \mp 2(x^1/2 + x^2) \partial_3.$$

12. $B_1 \neq 0, B_3 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^2 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^2). \quad (100)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^3 \partial_1 - (x^1 + x^2/2) \partial_3.$$

13. $B_1 \neq 0, B_3 \neq 0, C_3 \neq 0$. Канонический вид метрики:

$$G = \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) \pm \hat{\mathcal{S}}(dx^2 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^2 \otimes dx^2 \otimes dx^3). \quad (101)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^2 \partial_2 - (x^3/2) \partial_3 - (2x^1 + 3x^3/2) \partial_1; \quad X_2 = x^2 \partial_3 - (x^2 \pm 2x^3) \partial_1.$$

16.4 Метрики с $\tau(G) = 4$ (10 типов)

1. $F \neq 0, A_1 \neq 0, B_1 \neq 0, B_2 \neq 0$. Канонический вид метрики:

$$G = F\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3) + \epsilon_1 \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \epsilon_2 \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3) + dx^1 \otimes dx^1 \otimes dx^1. \quad (102)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + \epsilon_2 F x^2) \partial_2 - \epsilon_2 (\epsilon_1 x^2 + F x^3) \partial_3.$$

2. $F \neq 0, A_1 \neq 0, B_1 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = F\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3) \pm \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3) + dx^1 \otimes dx^1 \otimes dx^1. \quad (103)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 \pm x^2/F) \partial_3 - (x^2 + x^1/2F) \partial_2.$$

3. $F \neq 0, B_1 \neq 0, B_2 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = F\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3) \pm \hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3). \quad (104)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^2 \pm F x^3) \partial_3 \mp (F x^2 + x^3 + x^1/2) \partial_2.$$

4. $F \neq 0, B_2 \neq 0, B_3 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = F\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^2 \otimes dx^3 \otimes dx^3) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3). \quad (105)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^1 + x^3/2F) \partial_1 - (x^2 + x^1/F + x^3/2F) \partial_2.$$

5. $A_1 \neq 0, A_2 \neq 0, B_1 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^2 \otimes dx^2 + B\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^1 \otimes dx^1 \otimes dx^3). \quad (106)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^1 \partial_1 - Bx^1 \partial_2 + 2(B^2 x^2 - x^3 - x^1/2) \partial_3.$$

6. $A_1 \neq 0, A_2 \neq 0, B_1 \neq 0, C_3 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^2 \otimes dx^2 + B\hat{\mathcal{S}}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{\mathcal{S}}(dx^2 \otimes dx^2 \otimes dx^3). \quad (107)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^2 \partial_2 - (2x^3 + 2Bx^1 + x^2) \partial_3.$$

7. $A_1 \neq 0, B_1 \neq 0, B_2 \neq 0, C_1 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + \epsilon_1 \hat{S}(dx^1 \otimes dx^2 \otimes dx^2) + \epsilon_2 \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + C \hat{S}(dx^1 \otimes dx^1 \otimes dx^3). \quad (108)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^3 \partial_2 - (\epsilon_1 \epsilon_2 x^2 + \epsilon_2 C x^1 / 2) \partial_3.$$

8. $A_1 \neq 0, B_2 \neq 0, B_3 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 \pm \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{S}(dx^2 \otimes dx^3 \otimes dx^3) + C \hat{S}(dx^1 \otimes dx^1 \otimes dx^3). \quad (109)$$

Алгебра нетривиальных симметрий одномерна:

$$X = x^3 \partial_3 - C x^3 \partial_1 + (\pm C x^3 + 2(C^2 \mp 1)x^1 - 2x^2) \partial_2.$$

9. $A_1 \neq 0, B_1 \neq 0, C_1 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = dx^1 \otimes dx^1 \otimes dx^1 + B \hat{S}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{S}(dx^1 \otimes dx^1 \otimes dx^2) + \hat{S}(dx^1 \otimes dx^1 \otimes dx^3). \quad (110)$$

Алгебра нетривиальных симметрий двумерна:

$$X_1 = x^1 \partial_1 - (x^2 / 2) \partial_2 - (2x^3 + x^1 + 3x^2 / 2) \partial_3; \quad X_2 = x^1 \partial_2 - (x^1 + 2Bx^2) \partial_3.$$

10. $B_1 \neq 0, B_2 \neq 0, C_1 \neq 0, C_2 \neq 0$. Канонический вид метрики:

$$G = B \hat{S}(dx^1 \otimes dx^2 \otimes dx^2) + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{S}(dx^1 \otimes dx^1 \otimes dx^2) + \hat{S}(dx^1 \otimes dx^1 \otimes dx^3). \quad (111)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + x^1 / 2) \partial_2 - (Bx^2 + x^1 / 2) \partial_3.$$

16.5 Метрики с $\tau(G) = 5$ (5 типов)

С технической точки зрения — это самый сложный случай, так как в нем содержится самое большое количество комбинаций коэффициентов, подлежащих проверке. Эта сложность компенсируется относительной малостью случаев, в которых метрика обладает нетривиальной симметрией.

1. $F = 0, A_1 = 0, A_2 = 0, B_1 = 0, C_1 = 0$. Канонический вид метрики:

$$G = dx^3 \otimes dx^3 \otimes dx^3 + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{S}(dx^2 \otimes dx^3 \otimes dx^3) + C_2 \hat{S}(dx^1 \otimes dx^1 \otimes dx^3) + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3). \quad (112)$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + 2C_3 x^2) \partial_1 - (x^3 + 2C_2 x^1) \partial_2.$$

2. $F = 0, A_1 = 0, B_1 = 0, C_1 = 0, C_2 = 0$. Канонический вид метрики:

$$G = dx^2 \otimes dx^2 \otimes dx^2 + dx^3 \otimes dx^3 \otimes dx^3 + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + B_3 \hat{S}(dx^2 \otimes dx^3 \otimes dx^3) \quad (113) \\ + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3).$$

Алгебра нетривиальных симметрий одномерна:

$$X = ((B_3 - 1)x^3 + 2(C_3^2 - B_3)x^2 - 2x^1)\partial_1 - C_3x^3\partial_2 + x^3\partial_3.$$

3. $A_1 = 0, A_2 = 0, A_3 = 0, B_1 = 0, C_1 = 0$. Канонический вид метрики:

$$G = \hat{S}(dx^1 \otimes dx^2 \otimes dx^3) + C_2 \hat{S}(dx^1 \otimes dx^1 \otimes dx^3) + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3) + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) \quad (114) \\ + \hat{S}(dx^2 \otimes dx^3 \otimes dx^3).$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + 2x^1 + 2C_3x^2)\partial_1 - (2x^2 + 2C_2x^1 + x^3)\partial_2.$$

4. $A_1 = 0, A_2 = 0, B_1 = 0, B_2 = 0, C_1 = 0$. Канонический вид метрики:

$$G = \hat{S}(dx^1 \otimes dx^2 \otimes dx^3) + C_2 \hat{S}(dx^1 \otimes dx^1 \otimes dx^3) + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3) + \hat{S}(dx^2 \otimes dx^3 \otimes dx^3) + \quad (115) \\ + dx^3 \otimes dx^3 \otimes dx^3.$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^1 + x^2/C_2)\partial_2 - \frac{2x^1 + 2C_3x^2 + x^3}{2C_2}\partial_1.$$

5. $A_1 = 0, A_2 = 0, B_1 = 0, C_1 = 0, C_2 = 0$. Канонический вид метрики:

$$G = \hat{S}(dx^1 \otimes dx^2 \otimes dx^3) + dx^3 \otimes dx^3 \otimes dx^3 + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3) \quad (116) \\ + B_3 \hat{S}(dx^2 \otimes dx^3 \otimes dx^3).$$

Алгебра нетривиальных симметрий одномерна:

$$X = (x^3 + 2x^2)\partial_2 - (2x^1 + 2C_3x^2 + B_3x^3)\partial_1.$$

16.6 Метрики с $\tau(G) = 6$.

Существует только одна метрика общего вида этого типа при $A_1 = 0, A_2 = 0, B_1 = 0, C_1 = 0$:

$$G = F \hat{S}(dx^1 \otimes dx^2 \otimes dx^3) + \hat{S}(dx^1 \otimes dx^3 \otimes dx^3) + \hat{S}(dx^2 \otimes dx^3 \otimes dx^3) + dx^3 \otimes dx^3 \otimes dx^3 \quad (117) \\ + C_2 \hat{S}(dx^1 \otimes dx^1 \otimes dx^3) + C_3 \hat{S}(dx^2 \otimes dx^2 \otimes dx^3).$$

Алгебра ее нетривиальных симметрий одномерна:

$$X = (x^3 + 2Fx^1 + 2C_3x^2)\partial_1 - (2Fx^2 + 2C_2x^1 + x^3)\partial_2.$$

16.7 Метрики с $\tau(G) = 6, 7, 8, 9, 10$.

Среди метрик перечисленного аффинного типа нет представителей общего типа с нетривиальными симметриями.

Таким образом, имеется 41 общая кубическая метрика с различными аффинными типами, обладающих нетривиальными изометриями. Отметим, что наш анализ касается именно метрик общих аффинных типов. Внутри любого типа с тривиальными симметриями могут оказаться метрики со специальными значениями компонент метрики, при которых метрика данного типа будет обладать изометрией. В большинстве случаев такая изометрия будет эквивалентна одной из изометрий метрик с меньшим значением аффинного типа, у которых число компонент увеличилось за счет ее аффинного преобразования. В этом случае мы имеем дело с эквивалентными метриками. Но возможны ситуации, когда при частных значениях компонент метрика не эквивалентна ни одной из рассмотренных нами. Именно эти "очень специальные" метрики остаются за пределами нашего рассмотрения.

16.8 Инвариантная классификация метрик с нетривиальными изометриями

Некоторые из аффинных типов метрик, обладающих симметриями, являются аффинно-эквивалентными. Для выяснения вопроса об эквивалентности метрик из перечисленных выше 38 классам обратимся к (аффинно-)инвариантным свойствам их полей симметрий. Первичная классификация связана с размерностью алгебр симметрий. Группируя различные аффинные типы с одинаковыми размерностями алгебры симметрий, мы приходим к следующим заведомо не эквивалентным классам:

1. класс аффинных типов с двумерной алгеброй симметрии, включающий случаи (первая цифра — аффинный тип, вторая — порядковый номер в соответствующем разделе): 1.1, 2.2, 2.5, 2.6, 3.3, 3.5, 3.9, 3.13, 4.9;
2. класс аффинных типов с одномерной алгеброй симметрии, включающий случаи: 2.1, 2.3, 2.4, 3.1, 3.2, 3.4, 3.6, 3.7, 3.8, 3.10, 3.11, 3.12, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.10, 5.1, 5.2, 5.3, 5.4, 5.5, 6.1.
3. типы 1.2, 1.3, 2.7, 2.8, 2.9 с бесконечно-мерной алгеброй симметрии;
4. все, типы, которые не вошли в рассмотренные и которые не обладают нетривиальными симметриями;
5. "очень специальные метрики которые не входят ни в один из предыдущих пунктов.

Два последних класса целиком остаются за пределами нашего рассмотрения. Первые два класса допускают дальнейшую более детальную классификацию. Непосредственной проверкой убеждаемся, что коммутатор пары векторных полей симметрий метрик первого класса равен:

1. 0, для случаев 1.1, 2.2, 2.5, 3.3, 3.5;
2. $(3/2)X_2$ для случаев 2.6, 3.9, 3.13, 4.9.

Таким образом, приходим к заключению, что *группы метрик* $\{1.1, 2.2, 3.3, 3.5\}$, и $\{2.6, 3.9, 3.13, 4.9\}$ *аффинно-неэквивалентны*. Вопрос об аффинной эквивалентности метрик внутри этих групп, вообще говоря, остается открытым. Мы вернемся к нему в следующем разделе.

Перейдем к классу аффинных типов с одномерной алгеброй симметрий. Грубая классификация этих типов заключается в сравнении простейшего аффинного инварианта этих алгебр — дивергенции соответствующего векторного поля: $\operatorname{div} X \equiv \partial_i X^i$. Элементарное вычисление обнаруживает, что $\operatorname{div} X = 0$ для случаев: 2.1, 3.1, 3.2, 3.4, 3.6, 3.8, 3.11, 3.12, 4.1, 4.2, 4.3, 4.4, 4.7, 4.10, 5.1, 5.3, 5.4, 5.5, 6.1 и $\operatorname{div} X = \operatorname{const} \neq 0$ для случаев 2.3, 2.4, 3.7, 3.10, 4.5, 4.6, 4.8, 5.2. Таким образом, *аффинные типы, взятые из различных перечисленных здесь групп, аффинно-неэквивалентны*.

Для дальнейшей более детальной классификации типов внутри групп необходимо сравнивать другие аффинные инварианты. Для их построения учтем, что все векторные поля симметрий рассматриваемых типов имеют линейные и однородные по координатам компоненты. Каждому векторному полю такого вида можно поставить в соответствие *матрицу векторного поля*, определяемую соотношением:

$$X^\alpha = A^\alpha_\beta x^\beta,$$

где A^α_β — вещественные числа. Из этого определения вытекает, что A — аффинный тензор валентности $(1, 1)$. Его аффинными инвариантами будут, например, следующие величины:

$$I_1 \equiv \operatorname{Tr}(A), \quad \dots, I_n \equiv \operatorname{Tr}(A^n); \quad \Delta \equiv \det(A).$$

Отметим, что $\operatorname{div} X = I_1$. У эквивалентных метрик должны выполняться *условия коллинеарности*:

$$\sqrt[n]{\frac{I_n}{I'_n}} = \sqrt[3]{\frac{\Delta}{\Delta'}} = C = \operatorname{const} \quad (118)$$

для всех $n = 1, \dots$, где $\{I_n, \Delta\}$ — система инвариантов одной метрики, $\{I'_n, \Delta'\}$ — система инвариантов другой. Можно построить и другие инварианты, но для наших целей достаточно перечисленных.

Для класса метрик с $\operatorname{div} X \neq 0$ матрицы векторного поля и соответствующие инварианты имеют вид:

$$\begin{aligned} 1.2: \quad & \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, I_n = 1 + (-2)^n; & 2.3: \quad & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}, I_n = \frac{1+(-2)^n}{(-2)^n}; \\ 2.4: \quad & \begin{pmatrix} 1 & 0 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, I_n = 1 + (-2)^n; & 3.7: \quad & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & -2 \end{pmatrix}, I_n = 1 + (-2)^n; \\ 3.10: \quad & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ -2\epsilon & 0 & -2 \end{pmatrix}, I_n = 1 + (-2)^n; & 4.5: \quad & \begin{pmatrix} 1 & 0 & 0 \\ -B & 0 & 0 \\ 0 & 2B^2 & -2 \end{pmatrix}, I_n = 1 + (-2)^n; \\ 4.6: \quad & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 2B & 1 & -2 \end{pmatrix}, I_n = 1 + (-2)^n; & 2.3: \quad & \begin{pmatrix} 0 & 0 & -C \\ 2(C^2 \mp 1) & -2 & \pm C \\ 0 & 0 & 1 \end{pmatrix}, I_n = 1 + (-2)^n; \end{aligned}$$

$$5.2: \begin{pmatrix} -2 & 2(C_3^2 - B_3) & B_3 C_3 - 1 \\ 0 & 0 & -C_3 \\ 0 & 0 & 1 \end{pmatrix}, I_n = 1 + (-2)^n.$$

Очевидно, что условия (118) выполняются для всех метрик из группы с $\operatorname{div} X = I_1 \neq 0$.

Для группы метрик с $\operatorname{div} X = I_1 = 0$ матрицы соответствующих векторных полей и аффинные инварианты имеют следующий вид (инвариант Δ выписывается только в случае, когда он отличен от нуля):

$$\begin{aligned} 2.1: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, I_n = 1 + (-1)^n; & 3.1: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & -1 \end{pmatrix}, I_n = 1 + (-1)^n; \\ 3.2: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ -1/2 & 0 & -1 \end{pmatrix}, I_n = 1 + (-1)^n; & 3.4: & \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & -1/2 & -1 \end{pmatrix}, I_n = 1 + (-1)^n; \\ 3.6: & \begin{pmatrix} -1 & -1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}, I_n = 1 + (-1)^n; \\ 3.8: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\epsilon_1 \epsilon_2 & 0 \end{pmatrix}, I_n = (-\epsilon_1 \epsilon_2)^{n/2} (1 + (-1)^n); \\ 3.11: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \mp 1/2 & \mp 1 & 0 \end{pmatrix}, I_n = (\mp 1)^{n/2} (1 + (-1)^n); \\ 3.12: & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & -1/2 & 0 \end{pmatrix}, I_n = (-1)^{n/2} (1 + (-1)^n); \\ 4.1: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon_2 F & 1 \\ 0 & -\epsilon_1 \epsilon_2 & -\epsilon_2 F \end{pmatrix}, I_n = (F^2 - \epsilon_1 \epsilon_2)^{n/2} (1 + (-1)^n); \\ 4.2: & \begin{pmatrix} 0 & 0 & 0 \\ -1/2 F & -1 & 0 \\ 0 & \pm 1/F & 1 \end{pmatrix}, I_n = (1 + (-1)^n); \\ 4.3: & \begin{pmatrix} 0 & 0 & 0 \\ \mp 1/2 & \mp F & \mp 1 \\ 0 & 1 & \pm F \end{pmatrix}, I_n = (F^2 \mp 1)^{n/2} (1 + (-1)^n); \\ 4.4: & \begin{pmatrix} 1 & 0 & 1/2 F \\ -1/F & -1 & -1/2 F \\ 0 & 0 & 0 \end{pmatrix}, I_n = 1 + (-1)^n; \\ 4.7: & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -\epsilon_2 C/2 & -\epsilon_1 \epsilon_2 & 0 \end{pmatrix}, I_n = (-\epsilon_1 \epsilon_2)^{n/2} (1 + (-1)^n); \\ 4.10: & \begin{pmatrix} 0 & 0 & 0 \\ 1/2 & 0 & 1 \\ -1/2 & -B & 0 \end{pmatrix}, I_n = (-B)^{n/2} (1 + (-1)^n); \end{aligned}$$

$$\begin{aligned}
 5.1 : & \quad \begin{pmatrix} 0 & 2C_3 & 1 \\ -2C_2 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad I_n = (-4C_2C_3)^{n/2}(1 + (-1)^n); \\
 5.3 : & \quad \begin{pmatrix} 2 & 2C_3 & 1 \\ -2C_2 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad I_n = (4(1 - C_2C_3))^{n/2}((-1)^n + 1); \\
 5.4 : & \quad \begin{pmatrix} -1/C_2 & -C_3/C_2 & -1/2C_2 \\ 1 & 1/C_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad I_n = (1 - C_2C_3)^{n/2}(1 + (-1)^n); \\
 5.5 : & \quad \begin{pmatrix} -2 & -2C_3 & -B_3 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad I_n = 4^{n/2}(1 + (-1)^n); \\
 6.1 : & \quad \begin{pmatrix} 2F & 2C_3 & -B_3 \\ -2C_2 & -2F & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad I_n = 2^n(F^2 - C_2C_3)^{n/2}(1 + (-1)^n)
 \end{aligned}$$

Сравнительное исследование серий инвариантов обнаруживает следующие потенциальные классы аффинно-эквивалентных метрик:

1. {2.1, 3.1, 3.2, 3.4, 3.6, 3.8 ($\epsilon_1\epsilon_2 < 0$), 3.11 ($-$ в метрике), 4.1 ($F^2 > \epsilon_1\epsilon_2$), 4.2, 4.3 ($F^2 > \pm 1$), 4.4, 4.7 ($\epsilon_1\epsilon_2 < 0$), 4.10 ($B < 0$), 5.1 ($C_2C_3 < 0$), 5.3 $C_2C_3 < 1$, 5.4 ($C_2C_3 < 1$), 5.5, 6.1 ($F^2 > C_2C_3$)};
2. {3.8 ($\epsilon_1\epsilon_2 > 0$), 3.11 ($+$ в метрике), 3.12, 4.1 ($F^2 < \epsilon_1\epsilon_2$), 4.3 ($F^2 < \pm 1$), 4.7 ($\epsilon_1\epsilon_2 > 0$), 4.10 ($B < 0$), 5.1 ($C_2C_3 > 0$), 5.3 $C_2C_3 > 1$, 5.4 ($C_2C_3 < 1$), 6.1 ($F^2 < C_2C_3$)};
3. {4.1 ($F = \epsilon_1\epsilon_2$), 4.3 ($F^2 = +1$) (" $+$ " в метрике), 4.10 $B = 0$, 5.3 $C_2C_3 = 1$, 5.4 ($C_2C_3 = 1$), 6.1 ($F^2 = C_2C_3$)};

Более детальное дополнительное исследование группы случаев 3, соответствующих обнулению инвариантов, приводит к следующим уточнениям:

1. Метрика 4.1 при $F^2 = \epsilon_1\epsilon_2$ допускает векторное поле симметрии с одной произвольной функцией от всех координат, т.е. допускает бесконечномерную группу симметрии.
2. Метрика 4.3 при $F^2 = 1$ допускает двумерную неабелеву группу симметрии, с коммутатором вида: $[X_1, X_2] = (3/2)X_2$.
3. Метрика 4.10 при $B = 0$ допускает двумерную неабелеву группу симметрии, с коммутатором вида: $[X_1, X_2] = (3/2)X_2$.
4. Метрика 5.3 при $C_2C_3 = 1$ допускает двумерную неабелеву группу симметрии, с коммутатором вида: $[X_1, X_2] = (3/2)X_2$.
5. Метрика 5.4 при $C_2C_3 = 1$ допускает двумерную неабелеву группу симметрии, с коммутатором вида: $[X_1, X_2] = (3/2)X_2$.
6. Метрика 6.1 при $F^2 = C_2C_3$ допускает двумерную неабелеву группу симметрии, с коммутатором вида: $[X_1, X_2] = (3/2)X_2$.

Подводя итоги нашего исследования, можно заключить, что *все однородные кубические метрики общих аффинных типов делятся на 8 аффинно-неэквивалентных классов*:

1. класс метрик {1.1, 2.2, 2.5, 3.3, 3.5} (2-мерная абелева алгебра нетривиальных изометрий);

2. класс $\{2.6, 3.9, 3.13, 4.9, 4.3 (F^2 = 1), 4.10 (B = 0), 5, 3, 5.4 (C_2 C_3 = 1), 6.1 (F^2 = C_2 C_3)\}$ (2-мерная неабелева алгебра нетривиальных изометрий);
3. класс $\{1.2, 1.3, 2.7, 2.8, 2.9, 4.1 (F^2 = \epsilon_1 \epsilon_2)\}$ (бесконечномерная алгебра изометрий); при этом возникают подклассы (1): ∞^2 -мерной группы (1.3), (2): ∞ -мерной группы (2.8, 2.9) и (3): $\infty + 1$ -мерной группы (1.2, 2.7);
4. класс $\{2.3, 2.4, 3.7, 3.10, 4.5, 4.6, 4.8, 5.2\}$ (одномерная алгебра нетривиальных изометрий $I_1 \neq 0$);
5. класс $\{2.1, 3.1, 3.2, 3.4, 3.6, 3.8 (\epsilon_1 \epsilon_2 < 0), 3.11 (- \text{ в метрике}), 4.1 (F^2 > \epsilon_1 \epsilon_2), 4.2, 4.3 (F^2 > \pm 1), 4.4, 4.7 (\epsilon_1 \epsilon_2 < 0), 4.10 (B < 0), 5.1 (C_2 C_3 < 0), 5.4 (C_2 C_3 < 1), 5.5, 6.1 (F^2 > C_2 C_3)\}$ (одномерная алгебра нетривиальных изометрий, $\Delta = 0$, $I_n = C^{n/2}(1 + (-1)^n)$, $C = \text{const}$);
6. класс $\{3.8 (\epsilon_1 \epsilon_2 > 0), 3.11 (+ \text{ в метрике}), 3.12, 4.1 (F^2 < \epsilon_1 \epsilon_2), 4.3 (F^2 < \pm 1), 4.7 (\epsilon_1 \epsilon_2 > 0), 4.10 (B < 0), 5.1 (C_2 C_3 > 0), 5.4 (C_2 C_3 < 1), 6.1 (F^2 < C_2 C_3)\}$; (одномерная алгебра нетривиальных изометрий, $\Delta = 0$, $I_n = (-C)^{n/2}(1 + (-1)^n)$, $C = \text{const}$);
7. класс метрик, обладающих нетривиальными симметриями и не вошедших в предыдущий перечень;
8. класс метрик, имеющих только тривиальные симметрии.

Вопрос об аффинной эквивалентности метрик внутри этих классов остается открытым. В следующем разделе мы убедимся, что ответ на этот вопрос отрицателен.

16.9 Связь с проективной классификацией кубических форм

Выясним в этом разделе связь полученных результатов с известной проективной классификацией кубических форм [28]. Комбинация методов проективной геометрии и алгебры кубических матриц, приводит к следующей классификационной теореме.

ТЕОРЕМА (о классификации вещественных кубических форм) Кубическая форма над полем вещественных чисел принадлежит одному из классов аффинной вещественной эквивалентности (указаны только отличные от нуля компоненты в канонической форме метрики):

1. Общий класс $A_1 = A_2 = A_3 = 1$, с 10 неэквивалентными подклассами: $F < -(\sqrt{3} + 1)/2$, $F = -(\sqrt{3} + 1)/2$, $-(\sqrt{3} + 1)/2 < F < -1/2$, $-1/2 < F < 0$, $F = 0$, $0 < F < (\sqrt{3} - 1)/2$, $F = (\sqrt{3} - 1)/2$, $(\sqrt{3} - 1)/2 < F < 1$, $F = 1$, $F > 1$.
2. Вырожденный класс I: $A_1 = A_2$;
3. Вырожденный класс II: $A_1 = F = 1$;
4. Вырожденный класс III: $F = 1$;
5. Вырожденный класс IV: $A_1 = C_3 = 1$;
6. Вырожденный класс V: $C_1 = C_3 = 1$;
7. Вырожденный класс VI: $A_1 = A_2 = 1$;
8. Вырожденный класс VII: $C_1 = 1$;
9. Вырожденный класс VIII: $A_1 = 1$;
10. Вырожденный класс XIX: $A_3 = C_1 = B_3 = 1$;
11. Вырожденный класс X: $-A_2 = C_1 = B_3 = 1$;

12. Вырожденный класс XI: $A_2 = C_1 = B_3 = 1$;
13. Вырожденный класс XII: $C_1 = B_3 = 1$;
14. Вырожденный класс XI: $-A_2 = C_1 = 1$.

Сравнение приведенных в теореме канонических типов с рассмотренными выше аффинными типами приводит к следующим заключениям:

1. Общий класс при $F \neq -1/2$ имеет только тривиальные симметрии относится к симметричному типу 9; при $F = -1/2$ общий класс допускает 2-мерную абелеву группу симметрий и относится к симметричному типу 1;
2. Вырожденный класс I не имеет нетривиальных симметрий и также относится к типу 9;
3. Вырожденный класс II имеет 1-мерную группу симметрий с $I_1 = 0$ и относится к типу метрик 5 симметричного класса;
4. Вырожденный класс III имеет 2-мерную абелеву группу симметрий относится к симметричному типу 1;
5. Вырожденный класс IV имеет 1-мерную группу симметрий с $I_1 \neq 0$ и относится к симметричному типу 4;
6. Вырожденный класс V имеет 2-мерную неабелеву группу симметрий и относится к симметричному типу 2;
7. Вырожденный класс VI имеет ∞ -мерную группу симметрий относится к симметричному типу метрик 3(2);
8. Вырожденный класс VII имеет $\infty + 1$ -мерную группу симметрий и относится к симметричному типу метрик 3(3);
9. Вырожденный класс VIII имеет ∞^2 -мерную группу симметрий и относится к симметричному типу метрик 3(1);
10. Вырожденный класс IX не имеет нетривиальных симметрий и относится к симметричному типу метрик 8;
11. Вырожденный класс X имеет 1-мерную группу симметрий с $I_1 = 0$ и относится к симметричному типу метрик 5;
12. Вырожденный класс XI имеет 1-мерную группу симметрий с $I_1 = 0$ и относится к симметричному типу 5;
13. Вырожденный класс XII имеет 2-мерную абелеву группу симметрий и относится к симметричному типу метрик 1;
14. Вырожденный класс XIII имеет ∞ -мерную группу симметрий и относится к симметричному типу метрик 3(2);

Взаимосвязь симметричной и проективной классификаций резюмирована в таблице.

Симметричные классы	1	2	3	4	5	6	7	8
Проективные классы	III, XII	V	(1): VIII, (2): VI, XIII, (3): VII	IV	II, X, XI	?	—	Общ, I, IX

Анализ таблицы приводит к следующим важным заключениям:

1. Симметричная классификация является более грубой, чем проективная, поскольку некоторые симметричные классы содержат несколько проективных классов эквивалентных метрик.
2. Пустота клетки в 7 симметричном классе означает, что наш анализ охватил все неэквивалентные типы метрик.
3. Пустота клетки в классе 6 означает, что 5-й и 6-й симметричные классы следует считать одним и тем же классом. Общая константа в правой части условий коллинеарности (118), записанных для этих классов будет мнимая. Это, в свою очередь, означает что поля изометрий образуют, на самом деле, не $\mathbb{R}$ -модуль, как мы неявно полагали, а $\mathbb{C}$ -модуль.

Благодарности

В заключение этих лекций их автор хотел бы поблагодарить Д.Г.Павлова за стимулирующие дискуссии и предоставленную возможность озвучивания части этих лекций на осенней Школе-2008, а также всех участников школы за ценные вопросы и обсуждения.

Список литературы

- [1] Грин Б. Элегантная Вселенная. М., УРСС, 2004
- [2] Пенроуз Р. Путь к реальности. М.-Ижевск, РХД, 2007
- [3] Шутц Б. Геометрические методы в математической физике. М., Мир, 1984
- [4] Эллиот Дж., Добер П. Симметрия в физике (в 2-х томах). М., Мир, 1983
- [5] Дубровин Б.А., Новиков С.П., Фоменко А.Т. Современная геометрия. М., Наука, 1979
- [6] Постников М.М. Лекции по геометрии (сем. III: Гладкие многообразия). М., Наука, 1987
- [7] Постников М.М. Лекции по геометрии (сем. IV: Дифференциальная геометрия). М., Наука, 1988
- [8] Постников М.М. Лекции по геометрии (сем. V: Риманова геометрия). М., Факториал, 1988
- [9] Уорнер Ф. Основы теории гладких многообразий и групп Ли. М., Мир, 1987
- [10] Ленг С. Введение в теорию дифференцируемых многообразий. М., Платон, 1996
- [11] Номидзу К. Группы Ли и дифференциальная геометрия. М., Платон, 1996
- [12] Кобаяси Ш., Номидзу К. Основы дифференциальной геометрии (в 2-х томах). Новокузнецк, ИО НФМИ, 1999
- [13] Ильин В.А., Поздняк Э.Г. Математический анализ (в 2-х томах). М., Наука, 1982
- [14] Постников М.М. Лекции по геометрии (сем. I: Аналитическая геометрия). М., Физматлит, 1979
- [15] Постников М.М. Лекции по геометрии (сем. II: Линейная алгебра). М., Физматлит, 1979
- [16] Борисович Ю.Г., Близняков Н.М., Израилевич Я.А., Фоменко Т.Н. Введение в топологию. М., Наука, 1995
- [17] Brance C.H., Randall D. // gr-qc/9212003
- [18] Гараев К.Г. // Соросовский образовательный журнал, №12, 1998, с.113-118.

- [19] Арнольд В.И. Геометрические методы в теории обыкновенных дифференциальных уравнений. Ижевск, РХД, 2000
- [20] Олвер П. Приложение групп Ли к дифференциальным уравнениям. М., Мир, 1989
- [21] Симметрии и законы сохранения уравнений математической физики (под ред. А.М. Виноградова и И.С. Красильщика). М., Факториал, 1997
- [22] Новиков И.Д., Фролов В.П. Физика черных дыр. М., Наука, 1986
- [23] Точные решения уравнений Эйнштейна (под ред. Э. Шмутцера). М., Энергоиздат, 1982
- [24] Фок В.А. Теория пространства, времени и тяготения. М., Физматлит, 1961 (приложение А, с. 510)
- [25] Рунд Х. Дифференциальная геометрия финслеровых пространств. М., Наука, 1981
- [26] Павлов Д.Г., Гарасько Г.И. // Гиперкомплексные числа в геометрии и физике, 3(3), 2005
- [27] Павлов Д.Г., Кокарев С.С. // Гиперкомплексные числа в геометрии и физике, 2(10), 2008, с.3
- [28] Соколов Н.П. Пространственные матрицы и их приложения. Москва, ГИФМЛ, 1960

Основы дифференциальной геометрии и групп Ли

В.Н. Тришин

Институт гиперкомплексных систем в геометрии и физике, Фрязино, Россия.

E-mail: vtrishin@mtu-net.ru

Предлагаемые лекционные заметки посвящены изложению основных понятий и фактов из дифференциальной геометрии и теории групп Ли. Помимо стандартных объектов, таких как многообразия, расслоения, связности и метрика, обсуждается геометрический подход к дифференциальным уравнениям, основанный на G -структурах в пространстве струй.

1 Введение

Методы дифференциальной геометрии занимают важное положение в современном математическом аппарате теоретической физики. Общая теория относительности, теория калибровочных полей и их различные обобщения основываются на дифференциально - геометрических конструкциях.

В предлагаемых лекционных записях излагаются основные сведения по современной геометрии и теории групп Ли. Изложение по необходимости является весьма кратким и конспективным. Целью лекций было познакомить слушателей Школы, не являющихся математиками-геометрами, с основными понятиями, которыми оперирует современная геометрия. Для более детального изучения материала студентам было предложено ознакомиться у известными учебниками, примерный список которых приводится в конце лекций. В силу того, что лекции носили учебноознакомительный характер, большинство фактов в них приводится без доказательств. Заинтересованный слушатель может без труда найти все отсутствующие доказательства утверждений в имеющейся литературе. По той же причине отсутствуют прямые ссылки и цитирования результатов.

План лекций следующий. Основная часть предваряется вводным разделом о базовых понятиях топологии. Затем излагаются элементарные сведения из теории групп Ли. Основная часть содержит материал о главных и ассоциированных расслоениях и линейных связностях в расслоениях. Рассматривается их применение в теории калибровочных полей. Далее обсуждается геометрическая теория дифференциальных уравнений, основанная на понятии G -структуры в пространстве струй. Дано описание распределения Картана и связности Картана. В заключение кратко рассматриваются римановы и финслеровы метрики, симплектические структуры и комплексные многообразия.

2 Топологические пространства

Простейшей структурой, которую можно задать на множестве, является топологическая структура. Она определяет отношение близости между элементами множества. Существенно, что эта структура не позволяет дифференцировать функции, заданные на множестве, а также, вообще говоря, не определяет такие геометрические понятия как расстояние или угол. Для вычисления производных, векторов и т. п. необходимо определить структуру многообразия, которое можно представлять себе как пространство, склеенное из областей $\mathbb{R}^n$.

В топологическом пространстве (то есть множестве с топологической структурой) по сути можно только определить принадлежность точки к окрестности любой другой точки.

ОПРЕДЕЛЕНИЕ. **Топологическое пространство** – это пара $\{X, \tau\}$, где X – произвольное множество, а τ – набор подмножеств $\tau_i \subset X$, причем τ удовлетворяет следующим условиям:

1) $\emptyset \in \tau, X \in \tau$.

2) Объединение произвольного семейства множеств, принадлежащих τ , принадлежит τ . То есть, если $U_a \in \tau \forall a \in A$, где A множество произвольного индекса (принимаящего конечное или бесконечное число значений), то $\bigcup U_a \in \tau$.

3) Пересечение конечного семейства множеств, принадлежащих τ , принадлежит τ . То есть, если $U_i \in \tau \ i = 1, \dots, n$, то $\bigcap U_i \in \tau$. Подмножества, принадлежащие набору τ , называются *открытыми множествами*, а набор τ – *топологией* в X .

Отображение $f : X \mapsto Y$ топологических пространств называется *непрерывным*, если прообраз $f^{-1}(U)$ каждого открытого подмножества $U \subset Y$ является открытым множеством в X . Если отображение $f : X \mapsto Y$ непрерывно, взаимно однозначно и обратное к нему отображение $f^{-1} : Y \mapsto X$ тоже непрерывно, то отображение f называется *гомеоморфизмом*, а пространства X и Y – *гомеоморфными*. Заметим, что непрерывное взаимно-однозначное отображение может не иметь непрерывного обратного, например отображение $[0, 1) \ni x \rightarrow y = \exp(2\pi i x)$, не являющееся гомеоморфизмом. В этом примере отображение является *инъекцией* (injection) отрезка $[0, 1)$ в плоскость $\mathbb{R}^2$. Инъекция, являющаяся гомеоморфизмом, называется *вложением* (embedding).

Гомеоморфные пространства эквивалентны с топологической точки зрения. Можно сказать, что в топологии исследуются свойства пространств, сохраняющиеся при непрерывных отображениях. Существенно, что эти отображения могут быть даже не дифференцируемы.

Некоторые классы топологических пространств:

1. Пространство называется *хаусдорфовым*, если любую пару его точек можно окружить непересекающимися друг с другом открытыми множествами.
2. Пространство X называется *компактным*, если из любого счетного числа открытых областей, покрывающих X , можно выбрать *конечное* число покрывающих X . Например, дискретное пространство компактно тогда, и только тогда, когда оно конечно (в качестве покрывающих областей возьмем области, состоящие из одной точки). В компактном пространстве любая бесконечная последовательность имеет предельную точку.
3. Топологическое пространство *локально компактно*, если каждая точка имеет компактную окрестность. Очевидно, что все дискретные пространства локально компактны. Локально компактны также пространства $\mathbb{R}^n$, а бесконечномерные гильбертовы пространства не локально компактны.
4. Пространство называется *связным*, если оно не является объединением двух открытых непересекающихся множеств. Компонента связности точки пространства – это максимальное связное множество, содержащее эту точку. Пространство называется *линейно связным*, если любые две точки в нем можно соединить непрерывной кривой.

Для описания топологии пространство обычно снабжается некоторой дополнительной структурой.

ПРИМЕР. Важный класс топологических пространств образуют *метрические простран-*

ства. Для любых двух точек x и y метрического пространства определена функция расстояния $\rho(x, y)$ такая, что

- 1) $\rho(x, y) = \rho(y, x)$;
- 2) $\rho(x, x) = 0$, $\rho(x, y) > 0$ при $x \neq y$;

3) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ (“неравенство треугольника”). Открытыми множествами являются объединения произвольных семейств открытых шаров, где открытый шар с центром в точке x_0 радиуса ε есть совокупность таких точек x , что $\rho(x_0, x) < \varepsilon$. Такая топология называется *метрической топологией*, причем метрика здесь нужна только для того, чтобы определить близость точек. Заметим, что все метрические пространства хаусдорфовы. Кроме того, метрическое пространство компактно тогда, и только тогда, когда оно замкнуто и ограничено (то есть принадлежит шару конечного радиуса).

Другой важной структурой на топологическом пространстве является *дифференциальная структура* и различные дифференциально-геометрические объекты, построенные на ее основе. Описание топологии пространств с их помощью составляет предмет дифференциальной топологии, и будет обсуждаться в следующей части статьи, а в этой главе мы рассмотрим средства, позволяющие различать негомеоморфные пространства и не требующие для своего определения операции дифференцирования. Эти инструменты являются по сути алгебраическими.

2.1 Алгебраическая топология. Гомотопические группы

Применение алгебраических методов в топологии основывается на том соображении, что, грубо говоря, с алгебраическими структурами работать проще, чем с топологическими. Подходящий выбор таких структур позволяет описывать изменения в топологии с помощью алгебры. Топологический инвариант является алгебраическим объектом (числом, группой, линейным пространством и т. д.), который ассоциирован с топологическим пространством и сохраняется при гомеоморфизмах. Два пространства с различными алгебраическими структурами будут не гомеоморфны. Обратное, вообще говоря, не верно, то есть могут существовать пространства с одинаковыми алгебраическими (топологическими) инвариантами не гомеоморфные друг другу.

Для построения алгебраических инвариантов топологического пространства можно использовать непрерывные отображения некоторого “пробного” пространства в данное. Таким образом строятся *гомотопические группы*, где в качестве “пробного” пространства выступают сферы S^n . Определим понятие гомотопии.

ОПРЕДЕЛЕНИЕ. Непрерывные отображения f и g между топологическими пространствами X и Y называются **гомотопными**, если существует такое непрерывное отображение $F : [0, 1] \times X \rightarrow Y$, что $F(0, x) = f(x)$ и $F(1, x) = g(x)$ для всех $x \in X$.

Гомотопные отображения можно представлять себе как такие, образы которые могут быть совмещены друг с другом непрерывной деформацией. Отношение гомотопии является отношением эквивалентности, поэтому множество всех непрерывных отображений пространства X в Y можно разбить на классы эквивалентности, называемые *гомотопическими классами* $\pi(X, Y)$.

ОПРЕДЕЛЕНИЕ. n -мерная **гомотопическая группа** $\pi_n(X, x_0)$ – это множество (снабженное групповой операцией) гомотопических классов отображений сферы S^n с выделенной точкой в пространство X с выделенной точкой x_0 . Первая гомотопическая группа $\pi_1(X, x_0)$

(группа гомотопически эквивалентных петель) называется *фундаментальной группой* пространства. Групповую операцию на множестве гомотопических классов отображений можно определить как композицию двух петель, то есть двум петлям ставится в соответствие петля, получающаяся при последовательном прохождении обеих петель. Единицей группы будет класс стягиваемых петель, а обратным элементом – петля, проходимая в обратном направлении.

Заметим, что гомотопические группы $\pi_n(X, x_0)$ зависят от выделенной точки x_0 , хотя в случае линейно связных пространств эти группы изоморфны. Далее зависимость группы от выбора точки в обозначениях мы будем опускать и писать просто $\pi_n(X)$.

Некоторые свойства гомотопических групп:

1. Высшие гомотопические группы $\pi_n(X)$ при $n > 1$ всегда абелевы (коммутативны). Фундаментальная группа может быть некоммутативной, например $\pi_1(M^2)$, где M^2 – сфера с двумя и более ручками.
2. Для сфер $\pi_n(S^m) = 0$ при $n < m$, а $\pi_n(S^n) = \mathbb{Z}$.
3. Гомотопические группы прямых произведений: $\pi_n(X \times Y) = \pi_n(X) \times \pi_n(Y)$.

ПРИМЕР. Фундаментальная группа плоскости тривиальна, $\pi_1(\mathbb{R}^2) = 0$, так как любая петля может быть стянута в точку. Напротив, для плоскости с выколотой точкой $\pi_1(\mathbb{R}^2/0) = \mathbb{Z}$. Здесь целые числа, образующие группу, нумеруют количество обходов петли вокруг выколотого нуля на плоскости. Очевидно, что такие петли не гомотопны. Таким образом, мы доказали, что плоскость негомеоморфна плоскости без точки, так как эти пространства имеют различные фундаментальные группы.

С помощью фундаментальной группы легко доказать “интуитивно ясное” утверждение, что сфера не гомеоморфна тору. Действительно, на сфере любую петлю можно стянуть в точку, поэтому $\pi_1(S^2) = 0$. Для тора $T = S^1 \times S^1$ получаем $\pi_1(T) = \pi_1(S^1) \times \pi_1(S^1) = \mathbb{Z} \times \mathbb{Z}$.

Понятие фундаментальной группы ведет к определению *односвязного* пространства, для которого $\pi_1(X) = 0$. Так окружность S^1 связна, но не односвязна, $\pi_1(S^1) = \mathbb{Z}$. Важным средством для изучения неодносвязных пространств является понятие накрывающего пространства. Пространство $\tilde{X}$ называется *накрывающим* для X , если $\tilde{X}$ односвязно, и существует непрерывное отображение $f: \tilde{X} \rightarrow X$, такое что для любой точки $x \in X$ существует окрестность $U(x)$, прообраз которой есть объединение открытых в $\tilde{X}$ множеств, гомеоморфных $U(x)$. Число областей в прообразе $f^{-1}(U)$ называется числом листов накрытия. Например, для S^1 накрывающим пространством будет прямая $\mathbb{R}$ с бесконечным числом листов.

2.2 Гомологии

Гомотопические группы представляют собой одно из средств описания топологии с помощью непрерывных отображений. Другой аппарат для характеристики топологических свойств дает теория гомологий. Она основана на разбиении исходного пространства на более простые составляющие. Этому разбиению ставится в соответствие некоторая алгебраическая структура, например, клеточный комплекс, который будет описан ниже. Алгебраические характеристики этого комплекса, независимые от разбиения и сохраняющиеся при гомеоморфизмах, несут информацию о топологии пространства.

Существуют несколько различных способов для введения гомологий (заметим, что их строгое определение, в отличие от теории гомотопий, связано со значительными формальными трудностями), но все они основаны на семействе $C = \{C_n, \partial_n\}$, $n = 0, \pm 1, \dots$, где C_n

$$\dots \xleftarrow{\partial_{n-1}} C_{n-1} \xleftarrow{\partial_n} C_n \xleftarrow{\partial_{n+1}} C_{n+1} \xleftarrow{\quad} \dots$$

$$H_n(C) := \frac{\text{Ker } \partial_n}{\text{Im } \partial_{n+1}}$$

$$\Delta^k := \{x = \sum_{i=0}^k x^i e_i \mid \sum_{i=0}^k x^i = 1, 0 \leq x_i \leq 1\}$$

$$\sigma : \Delta^k \rightarrow X$$

$$\partial : C_k(X, R) \rightarrow C_{k-1}(X, R)$$

Граница стандартного симплекса $\Delta^k = [p_0, \dots, p_k]$ образована $(k-1)$ -мерными гранями – симплексами $\Delta_i^{k-1} = [p_0, \dots, \hat{p}_i, \dots, p_k]$, где $\hat{p}_i$ означает, что вершина p_i удалена. Эти грани должны быть ориентированы, чтобы их замыкание образовывало петлю, являющуюся геометрической границей исходного симплекса. В результате получаем

$$\partial\Delta^k = \sum_i (-1)^i \Delta_i^{k-1}$$

Для сингулярного k -симплекса определим грань как $\sigma^{(i)} = \sigma \circ \Delta_i^k$. Тогда границей k -симплекса будет формальная сумма $\partial\sigma = \sum_{i=0}^k (-1)^i \sigma^{(i)}$. Действие оператора границы можно продолжить по линейности на любую сингулярную цепь, то есть, если $\sigma = \sum a^k \sigma_k$, то $\partial\sigma = \sum a^k \partial\sigma_k$. Можно проверить, что $\partial^2 := \partial \circ \partial = 0$ (“граница границы равна нулю”). Фактор-пространства (для переменного k)

$$H_k(X, R) = \frac{\text{Ker } \partial_k}{\text{Im } \partial_{k+1}}$$

называются группами сингулярных гомологий с коэффициентами в R . Если в качестве кольца R берут множество целых чисел, то группу гомологий $H_k(X, \mathbb{Z})$ обозначают просто $H_k(X)$.

ПРИМЕР. Вычислим группы гомологий для сферы S^2 . Клеточное разбиение сферы состоит из 0-мерной клетки (точки a) и 2-мерной клетки (диска d). Все цепи σ_i , $i = 0, 1, 2$ являются циклами: $\partial\sigma_i = 0$ ($\partial\sigma_2 = 0$ в силу того, что граница 2-мерной клетки является одномерной, а они отсутствуют в разбиении). Найдем группу $H_0(S^2) = \text{Ker } \partial_0 / \text{Im } \partial_1$. В силу того, что одномерных клеток нет, $\text{Im } \partial_1$ состоит из одного элемента, равного 0. Кроме того, каждая 0-мерная клетка является циклом ($\partial a = 0$), и 0-мерная цепь имеет вид $\sigma_0 = \alpha a$, $\alpha \in \mathbb{Z}$. Поэтому $H_0(S^2) = \text{Ker } \partial_0 = \mathbb{Z}$. Для $H_1(S^2) = \text{Ker } \partial_1 / \text{Im } \partial_2 = 0$ (нет одномерных клеток). Для $H_2(S^2) = \text{Ker } \partial_2 / \text{Im } \partial_3$, где $\text{Ker } \partial_2 = \beta d$ (любая 2-цепь – цикл), а $\text{Im } \partial_3$ состоит только из 0 (отличных от нуля 3-мерных клеток нет), таким образом $H_2(S^2) = \mathbb{Z}$. Все группы $H_n(S^2) = 0$ при $n > 2$.

Группы гомологий, которые используются в приложениях, являются конечно - порожденными группами, то есть имеют конечное число образующих. Такие группы изоморфны прямой сумме циклических и свободных групп. Если группа гомологий содержит циклическую подгруппу, то говорят, что она содержит *кручение*, например для бутылки Клейна K^2 имеем $H_1(K^2) = \mathbb{Z}_2 \oplus \mathbb{Z}$.

Если в качестве коэффициентов берут множество действительных чисел, $R = \mathbb{R}$, то группы гомологий являются векторными пространствами над $\mathbb{R}$, и можно определить *числа Бетти* по формуле $b_k(X) := \dim H_k(X, \mathbb{R})$. Эти числа, грубо говоря, подсчитывают количество дырок в различных измерениях, в частности $b_0(X)$ – это число компонент связности пространства X . С помощью чисел Бетти можно определить важный топологический инвариант – *эйлерову характеристику*

$$\chi(X) = \sum_k (-1)^k b_k$$

В заключение заметим, что группы гомологий формально зависят от выбора кольца коэффициентов R , хотя, например, имеются простые соотношения $H_k(X, \mathbb{R}) = H_k(X) \otimes \mathbb{R}$ или $H_k(X, \mathbb{C}) = H_k(X) \otimes \mathbb{C}$. Можно сказать, что группы $H_k(X)$, $H_k(X, \mathbb{R})$ и $H_k(X, \mathbb{C})$ совпадают по модулю конечных групп (кручения). В частности группы $H_k(X, \mathbb{R})$ кручения не содержат.

3 Гладкие многообразия

В физике используются поля и полевые уравнения, поэтому только топологии на множестве недостаточно – необходимо уметь производить дифференцирования функций, заданных на топологическом пространстве. Для этого необходимо задать на нем *дифференциальную (гладкую) структуру*, превратив его тем самым в *дифференцируемое (гладкое) многообразие*.

3.1 Определение многообразия. Гладкая структура

ОПРЕДЕЛЕНИЕ. Пусть M – хаусдорфово топологическое пространство, покрытое счетным семейством открытых множеств $\mathcal{U}$ с гомеоморфизмами $\phi_U : U \in \mathcal{U} \rightarrow U_R$, где U_R – открытое множество в $\mathbb{R}^n$. Если числовые функции (*функции перехода*) $f_{UV} := \phi_U \cdot \phi_V^{-1}$, отображающие $\phi_V(U \cap V)$ в $\phi_U(U \cap V)$, являются гладкими в $\mathbb{R}^n$ (принадлежат классу C^∞), то пара (U, ϕ_U) называется (гладкой) *картой*. Семейство $A = \{\mathcal{U}, \phi_U\}$ попарно гладко согласованных карт называется (гладким) *атласом*, или гладкой структурой. Два атласа многообразия называются эквивалентными, если их объединение – тоже атлас. Пространство M наделенное семейством эквивалентных атласов называется n -мерным вещественным **гладким многообразием**.

Если функции перехода f_{UV} дифференцируемы (класса C^1), то многообразие называется дифференцируемым. Если f_{UV} аналитичны* (класса C^ω), то многообразие называется аналитическим. Формально можно рассмотреть случай, когда f_{UV} только непрерывны (класса C^0), но не обязательно дифференцируемы. Тогда многообразие называется топологическим. Можно доказать, что структура топологического многообразия определяется единственным способом, поэтому топологические многообразия – это просто подкласс топологических пространств. Заметим, что на топологическом пространстве может вообще не существовать гладкой структуры, такие пространства называются негладкими.

Далее под словом многообразие, если не оговорено другое, мы всегда будем понимать именно гладкое многообразие.

Можно рассматривать гомеоморфизмы $\phi_U : U \in \mathcal{U} \rightarrow U_C$ в открытые множества в $\mathbb{C}^n$ (M очевидно должно иметь четную размерность). Тогда, если функции f_{UV} голоморфны, многообразие M называется *комплексным*. На таком многообразии можно ввести комплексные координаты.

ПРИМЕР. Сфера S^2 является гладким многообразием, атлас которого состоит из двух карт – (U, ϕ_U) и (V, ϕ_V) , где ϕ_U (ϕ_V) – стереографическая проекция сферы U без южного полюса (соответственно сферы V без северного полюса) на плоскость $\mathbb{R}^2$. Если (x_1, y_1) и (x_2, y_2) соответствующие координаты в этих картах, то функции перехода имеют вид

$$x_2 = \frac{x_1}{x_1^2 + y_1^2}, \quad y_2 = \frac{-y_1}{x_1^2 + y_1^2}$$

Если использовать комплексную координату $z = x_1 + iy_1$, то функция перехода (голоморфная на пересечении карт) имеет вид $w = 1/z$, где $w = x_2 + iy_2$. Голоморфность функции перехода означает, что сфера является комплексным многообразием.

ПРИМЕР. Проективные пространства. Для ненулевых векторов в пространстве $\mathbb{R}^{n+1}$ определим соотношение эквивалентности, считая, что векторы v и λv , $\lambda \neq 0$ задают одну

*т. е. могут быть представлены сходящимся рядом Тейлора на всей области определения

и ту же точку. Пространство классов эквивалентности называется (действительным) *проективным пространством* $\mathbb{R}P^n$. Оно обладает структурой n -мерного многообразия, атлас которого состоит из $n + 1$ карт U_k , $k = 0, \dots, n$ с координатами

$$x_{(k)}^\mu = \left(\frac{y^0}{y^k}, \dots, \frac{y^{k-1}}{y^k}, \frac{y^{k+1}}{y^k}, \dots, \frac{y^n}{y^k} \right), \quad \mu = 1, \dots, n$$

Числа $y^i \in \mathbb{R}^{n+1}$ называют *однородными координатами* и записывают в виде $[y^0 : \dots : y^n]$. Функции перехода, например, на пересечении карт U_0 и U_1 , где $y^0 \neq 0$ и $y^1 \neq 0$, имеют вид

$$x_{(1)}^\mu = \left(\frac{1}{x_{(0)}^1}, \frac{x_{(0)}^2}{x_{(0)}^1}, \frac{x_{(0)}^3}{x_{(0)}^1}, \dots, \frac{x_{(0)}^n}{x_{(0)}^1} \right)$$

и являются гладкими дробно-линейными функциями ($x_{(0)}^1 = y^1/y^0 \neq 0$). Все проективные пространства компактны, а нечетномерные $\mathbb{R}P^{2n-1}$ – ориентируемы, в частности $\mathbb{R}P^1 \approx S^1$ и $\mathbb{R}P^3 \approx SO(3)$. Геометрия проективных пространств устроена проще, чем аффинных. Например, все прямые в проективном пространстве пересекаются.

Понятие проективного пространства можно обобщить, введя *проективную структуру* на многообразии M . Проективная структура – это класс эквивалентности проективных атласов, то есть атласов с дробно-линейными функциями перехода на пересечении областей U_i и U_j вида

$$x_{(i)}^\mu = \frac{a_0^\mu + a_1^\mu x_{(j)}^1 + \dots + a_n^\mu x_{(j)}^n}{a_0^0 + a_1^0 x_{(j)}^1 + \dots + a_n^0 x_{(j)}^n}$$

где $(n + 1) \times (n + 1)$ матрица $\|a\|_\nu^\mu$ невырождена. Грубо говоря, многообразие с проективной структурой локально выглядит так же, как $\mathbb{R}P^n$.

Из определения многообразия естественно вытекает определение гладкого отображения многообразий. $f : M \rightarrow N$ является гладким, если его выражение в терминах локальных координат принадлежит классу C^∞ . Если это отображение является гомеоморфизмом, и обратное к нему тоже гладко, то f называется *диффеоморфизмом*. С физической точки зрения диффеоморфизм представляет собой глобальную замену координат (общековариантные преобразования). Множество всех диффеоморфизмов образует (бесконечномерную) группу Ли – *группу диффеоморфизмов* многообразия.

Важно заметить, что не все гомеоморфные многообразия являются диффеоморфными. Другими словами, на топологическом пространстве могут быть *различные* дифференциальные (гладкие) структуры, которые нельзя перевести друг в друга с помощью диффеоморфизма. Так на пространствах $\mathbb{R}^n$, рассматриваемых как многообразия, существует единственная, с точностью до диффеоморфизма, гладкая структура для всех n , кроме 4. На пространстве $\mathbb{R}^4$ существует бесконечно много (континуум) гладких структур! Такие $\mathbb{R}^4$ с нестандартной гладкой структурой называются экзотическими, или фальшивыми, $\mathbb{R}^4$. Они были открыты в начале 80-х годов в работах Фридмана и Дональдсона.

3.2 Векторные поля и дифференциальные формы.

Когомологии де Рама

Для дифференцирования функций на многообразии необходимо ввести понятие векторного поля. На современном языке векторное поле – это, грубо говоря, “дифференцирование по направлению”.

ОПРЕДЕЛЕНИЕ. **Касательный вектор** v_x в точке x многообразия M – это оператор, который каждой дифференцируемой функции f на M ставит в соответствие действительное число $v_x(f)$, при этом должны выполняться условия:

- 1) линейность: $v_x(af + bg) = av_x(f) + bv_x(g)$, где a, b – константы, f, g – функции;
- 2) правило Лейбница: $v_x(fg) = f(x)v_x(g) + g(x)v_x(f)$.

Касательные векторы к n -мерному многообразию M в точке x образуют n -мерное векторное пространство $T_x M$, которое называется *касательным пространством*. Если в качестве базисных векторов этого пространства взять касательные векторы к координатным линиям на M через точку x , то полученный базис называется *голономным* (координатным) и обозначается $\{\partial_\mu\}$. Любой вектор v_x можно записать в виде

$$v_x = v^\mu(x) \frac{\partial}{\partial x^\mu} := v^\mu \partial_\mu$$

откуда видно, что локально вектор соответствует контравариантному тензору 1-го ранга.

Множество TM всех касательных пространств к многообразию M также можно наделить структурой $2n$ -мерного многообразия, задав на нем голономную систему координат $(x^\mu, \dot{x}^\mu)$ с условиями перехода

$$x'^\mu = x'^\mu(x^\nu), \quad \dot{x}'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} \dot{x}^\nu$$

где x^μ – координаты на M , а $\dot{x}^\mu$ – координаты на касательных пространствах к M относительно голономных базисов.

В касательном пространстве можно определить *коммутатор* двух векторных полей $[u, v](f) := u(v(f)) - v(u(f))$. Соответствующее векторное поле имеет вид

$$[u, v] = (u^\nu \frac{\partial v^\mu}{\partial x^\nu} - v^\nu \frac{\partial u^\mu}{\partial x^\nu}) \partial_\mu$$

Антисимметричная операция коммутирования удовлетворяет тождеству Якоби

$$[[u, v], w] + [[v, w], u] + [[w, u], v] = 0$$

поэтому векторные поля образуют (бесконечномерную) алгебру Ли $\mathfrak{a}M$ относительно коммутатора. Она является алгеброй Ли группы Ли $Diff(M)$ всех диффеоморфизмов многообразия M .

Используя тензорное произведение $T_x M \times \dots \times T_x M$ можно определить поливекторное поле w , которое локально представимо контравариантным тензором ранга k : $w = w^{\mu_1 \dots \mu_k} \partial_{\mu_1} \otimes \dots \otimes \partial_{\mu_k}$.

Дифференциальные формы.

Алгебраически двойственным объектом к касательному вектору v является дифференциальная 1-форма α . Она осуществляет отображение вектора v в действительное число посредством свертки (внутреннего произведения) $v \lrcorner \alpha$. Свертка часто обозначается так же как $\iota_v \alpha$ или $\alpha(v)$. Это отображение обладает свойством линейности $(au + bv) \lrcorner \alpha = au \lrcorner \alpha + bv \lrcorner \alpha$.

В локальных координатах 1-форма в точке x имеет вид $\alpha = \alpha_\mu(x) dx^\mu$, причем координатный базис 1-форм $dx^1, \dots, dx^n$ по определению связан с базисом векторных полей соотношением $\partial_\mu \lrcorner dx^\nu = \delta_\mu^\nu$. Локально 1-форма соответствует ковариантному тензору 1-го ранга. В точке p многообразия M 1-формы образуют *кокасательное* пространство $T_x^* M$.

Используя линейную алгебру, можно определить полилинейные функции $T_x M \times \dots \times T_x M \rightarrow \mathbb{R}$. Среди них выделим класс антисимметричных линейных k -форм, определенных соотношением

$$\omega(X_1, \dots, X_i, \dots, X_j, \dots, X_k) = -\omega(X_1, \dots, X_j, \dots, X_i, \dots, X_k)$$

где $X_i \in T_x M$. Множество дифференцируемых функций, удовлетворяющих этому соотношению, обозначим $\Lambda^k(T_x M)$. Элементы $\omega_x \in \Lambda^k(T_x M)$ называются *дифференциальными (внешними) k -формами*. Множество всех k -форм на многообразии обозначим через $\Omega^k(M)$. Пространство всех форм $\Omega(M) = \bigoplus \Omega^k(M)$ обладает структурой градуированной грассмановой алгебры, где операция произведения между элементами алгебры определена с помощью ассоциативного, антикоммутативного и дистрибутивного *внешнего умножения* $\wedge$ по формуле

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$$

где $\omega \in \Omega^p(M)$, а $\eta \in \Omega^q(M)$. В локальных координатах любая k -форма $\omega \in \Omega^k(M)$ может быть записана в виде

$$\omega = \sum_{\mu_1 < \dots < \mu_k} f_{\mu_1 \dots \mu_k}(x) dx^{\mu_1} \wedge \dots \wedge dx^{\mu_k}$$

Функции по определению являются 0-формами. Если размерность многообразия равна n , то все k -формы при $k > n$ равны нулю, а n -форма $\omega = f(x) dx^{\mu_1} \wedge \dots \wedge dx^{\mu_n}$ состоит из одного слагаемого, причем $f(x)$ не является числовой функцией, так как при координатном преобразовании делится на якобиан перехода.

Алгебру дифференциальных k -форм можно превратить в дифференциальный комплекс (частный случай цепного комплекса), называемый *комплексом де Рама*. Для этого определим операцию *внешнего дифференцирования* d . Внешний дифференциал d переводит k -форму в $(k+1)$ -форму и обладает свойствами:

- 1) $dd = 0$;
- 2) $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$, $\omega \in \Omega^k(M)$. В голономных координатах имеем

$$d\omega = \sum \frac{\partial f_{\mu_1 \dots \mu_k}(x)}{\partial x^\mu} dx^\mu \wedge dx^{\mu_1} \wedge \dots \wedge dx^{\mu_k}$$

ПРИМЕР. Для многообразия $M = \mathbb{R}^3$ комплекс де Рама имеет вид

$$0 \rightarrow \Omega^0(M) \xrightarrow{\text{grad}} \Omega^1(M) \xrightarrow{\text{rot}} \Omega^2(M) \xrightarrow{\text{div}} \Omega^3(M) \rightarrow 0$$

где над стрелками приведены обычные операции векторного анализа, соответствующие внешнему дифференцированию d . Например, в голономных координатах имеем следующие представления

$$\begin{aligned} \Omega^0(M) &\xrightarrow{d} \Omega^1(M) : & d(f(x)) &= \frac{\partial f}{\partial x^\mu} dx^\mu \\ \Omega^1(M) &\xrightarrow{d} \Omega^2(M) : & d(f_\mu(x) dx^\mu) &= \frac{\partial f_\mu}{\partial x^\alpha} dx^\alpha \wedge dx^\mu \\ \Omega^2(M) &\xrightarrow{d} \Omega^3(M) : & d(f_{\mu\nu}(x) dx^\mu \wedge dx^\nu) &= \frac{\partial f_{\mu\nu}}{\partial x^\alpha} dx^\alpha \wedge dx^\mu \wedge dx^\nu \end{aligned}$$

Форма ω называется *замкнутой*, если $d\omega = 0$, то есть $\omega \in Z^k(M) := \text{Ker} \{d : \Omega^k(M) \rightarrow \Omega^{k+1}(M)\}$. Если существует $(k-1)$ -форма η такая, что $\omega = d\eta$, то есть $\omega \in B^k(M) := \text{Im} \{d : \Omega^{k-1}(M) \rightarrow \Omega^k(M)\}$, то форма ω называется *точной*.

Фактор-пространство замкнутых k -форм по точным называется k -группой когомологий де Рама

$$H_{dR}^k(M) = \frac{\text{Ker } d^k}{\text{Im } d^{k-1}}$$

Оно образовано классами эквивалентности замкнутых форм, разность которых является точной формой. Эти группы являются линейными пространствами, причем размерность группы $H_{dR}^0(M)$, состоящей из постоянных функций, равна числу компонент связности многообразия M . При $k > n$, где n – размерность многообразия M , $H_{dR}^k(M) = 0$.

ПРИМЕР. Для пространства $\mathbb{R}^n$ все группы когомологий $H_{dR}^k(\mathbb{R}^n) = 0$ при $k > 0$, $H_{dR}^0(\mathbb{R}^n) = \mathbb{R}$ (пространство постоянных функций). Можно доказать, что для любого стягиваемого многообразия $H_{dR}^k(M) = 0$ при $k > 0$. Для нестягиваемых многообразий группы когомологий могут быть нетривиальными, в частности на таких многообразиях могут существовать неградиентные векторные поля, ротор которых равен 0. Например, на плоскости без точки $M = \mathbb{R}^2/0$ можно определить замкнутую, но не точную 1-форму $\omega = \frac{y}{x^2+y^2}dx + \frac{-x}{x^2+y^2}dy$.

Группы когомологий де Рама $H_{dR}^k(M)$ изоморфны группам сингулярных когомологий $H^k(M, \mathbb{R})$ и поэтому двойственны группам сингулярных гомологий $H_k(M, \mathbb{R})$ (теорема де Рама). Эта двойственность устанавливается билинейной формой

$$\langle \omega^k | c_k \rangle = \int_{c_k} \omega^k \quad (1)$$

определяющей интегрирование замкнутых k -форм ω^k по k -мерным циклам c_k , реализуемым подмногообразиями M .

Производная Ли.

Производная Ли $\mathcal{L}_v T$ тензорного поля T по направлению векторного поля v – это главная линейная часть приращения тензорного поля T при его преобразовании, которое индуцировано локальной однопараметрической группой диффеоморфизмов Φ_t многообразия, порожденной полем v . Например, если при диффеоморфизме точка с координатами $x^\mu(t)$ переходит в точку x_0^μ (векторное поле v является касательным к линии $x(t)$), то закон преобразования тензора T_ν^μ при переносе из $x(t)$ в x_0 имеет вид

$$(\Phi_t T)_\nu^\mu = T_\beta^\alpha \frac{\partial x_0^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x_0^\nu}$$

и для производной Ли по определению получаем выражение

$$\mathcal{L}_v T_\nu^\mu = \left. \frac{d(\Phi_t T)_\nu^\mu}{dt} \right|_{t=0}$$

Можно сказать, что производная Ли измеряет скорость изменения тензора T при деформации многообразия, вызванного диффеоморфизмом Φ_t . Отметим, что производная Ли (как и

внешнее дифференцирование форм) представляет собой операцию, определенную на многообразии без введения дополнительных структур, в частности, она не зависит от связности или метрики.

Для функции $f(x)$ производная Ли совпадает с обычной производной по направлению векторного поля v , то есть в голономных координатах $\mathcal{L}_v f = v^\mu \partial_\mu f$. Производная Ли от векторного поля u есть коммутатор полей $\mathcal{L}_v u = [v, u] = (v^\alpha \partial_\alpha u^\mu - u^\alpha \partial_\alpha v^\mu) \partial_\mu$. Для тензорного поля произвольного ранга имеем

$$\mathcal{L}_v T_{\nu_1 \dots \nu_m}^{\mu_1 \dots \mu_n} = v^\alpha \partial_\alpha T_{\nu_1 \dots \nu_m}^{\mu_1 \dots \mu_n} + (\partial_{\nu_1} v^\alpha) T_{\alpha \dots \nu_m}^{\mu_1 \dots \mu_n} + (\partial_{\nu_m} v^\alpha) T_{\nu_1 \dots \alpha}^{\mu_1 \dots \mu_n} - (\partial_\alpha v^{\mu_1}) T_{\nu_1 \dots \nu_m}^{\alpha \dots \mu_n} - (\partial_\alpha v^{\mu_n}) T_{\nu_1 \dots \nu_m}^{\mu_1 \dots \alpha}$$

Производная Ли используется для описания симметрий, а именно, если мы хотим исследовать инвариантность какого-либо объекта, заданного на M , относительно действия некоторой группы диффеоморфизмов, то достаточно рассмотреть соответствующие производные Ли. Например, векторное поле v определяет *изометрию** многообразия тогда, и только тогда, когда производная Ли метрического тензора g вдоль этого поля равна нулю, то есть $\mathcal{L}_v g = 0$. Заметим, что в приложениях, например, в общей теории относительности, это уравнение обычно записывают через ковариантные производные ∇_μ относительно связности Леви-Чивита (см. главу 5) в виде $\nabla_{(\mu} v_{\nu)} = 0$ (уравнения Киллинга), а векторное поле v , являющееся решением этого уравнения, называют *полем Киллинга*.

Конформные преобразования метрики описываются векторным полем v , удовлетворяющим уравнению $\mathcal{L}_v g = \phi(x)g$, где $\phi(x)$ – произвольная функция на M .

На дифференциальных формах производная Ли и внешнее дифференцирование связаны следующим соотношением:

$$\mathcal{L}_v \omega = v \lrcorner d\omega + d(v \lrcorner \omega)$$

Оператор Ходжа. На пространстве форм $\Omega(M)$ можно определить линейный *оператор дуальности Ходжа* $\star : \Omega^k(M) \rightarrow \Omega^{n-k}(M)$. В голономных координатах он имеет вид

$$\star(dx^{\mu_1} \wedge \dots \wedge dx^{\mu_k}) = \frac{\sqrt{|\det g_{\mu\nu}|}}{(n-k)!} \varepsilon_{\nu_{k+1} \dots \nu_n}^{\mu_1 \dots \mu_k} dx^{\nu_{k+1}} \dots dx^{\nu_n}$$

где $\varepsilon_{\nu_{k+1} \dots \nu_n}^{\mu_1 \dots \mu_k}$ – полностью антисимметричный тензор, а $g_{\mu\nu}$ – метрический тензор. В частности, оператор $\star$ переводит функции в n -формы, например $\star 1 = \nu_M$, где $\nu_M = \sqrt{|\det g_{\mu\nu}|} dx^1 \wedge \dots \wedge dx^n$ – инвариантная (относительно диффеоморфизмов) форма объема. Очевидно, оператор Ходжа зависит от выбора метрики и ориентации на M .

Действие оператора Ходжа можно расширить на группы когомологий $H_{dR}^k(M)$. Для гладких компактных многообразий это приводит к изоморфизму групп $H_{dR}^k(M)$ и $H_{dR}^{(n-k)}(M)$, что с учетом двойственности групп гомологий и когомологий ведет к *двойственности Пуанкаре*: изоморфизму групп гомологий над $\mathbb{R}$ $H_k(M, \mathbb{R})$ и групп когомологий де Рама $H_{dR}^{(n-k)}(M)$. Это соответствие позволяет использовать на компактных многообразиях билинейную форму (1), чтобы ввести скалярное произведение на пространстве $\Omega(M)$ по формуле

$$\langle \omega, \eta \rangle = \int_M \omega \wedge \star \eta$$

*преобразование, сохраняющее расстояние между точками

Используя оператор $\star$, можно определить *кодифференциал* $\delta = (-1)^{k(n-k)+1} \star d \star$. Он понижает степень формы на единицу $\delta : \Omega^k(M) \rightarrow \Omega^{k-1}(M)$ и сопряжен к оператору d :

$$\langle d\omega, \eta \rangle = \langle \omega, \delta\eta \rangle, \quad \omega \in \Omega^k(M), \eta \in \Omega^{k+1}(M)$$

Для него также справедливо свойство $\delta\delta = 0$. Группы когомологий, индуцированные δ , совпадают с когомологиями де Рама, определенными с помощью d .

Из операторов d и δ можно получить инвариантный (топологический) *лапласиан* (оператор Лапласа-Бельтрами) $\Delta_k = (d + \delta)^2 = d\delta + \delta d$ – самосопряженный оператор второго порядка $\Delta_k : \Omega^k(M) \rightarrow \Omega^k(M)$. Решения уравнения $\Delta_k \omega = 0$ называются *гармоническими* формами. Они обобщают обычные гармонические функции математического анализа. Используя их, можно записать для компактных многообразий следующее “разложение Ходжа”:

$$\Omega^k = \mathcal{H}^k \oplus \text{Im } d \oplus \text{Im } \delta,$$

где $\mathcal{H}^k = \text{Ker } \Delta_k$ – пространство всех гармонических k -форм.

Из этого разложения следует важное следствие о конечномерности групп когомологий компактных многообразий, а именно, векторное пространство гармонических k -форм изоморфно k -ой группе когомологий де Рама: $\mathcal{H}^k \simeq H_{dR}^k(M)$ (теорема Ходжа). Таким образом, мы имеем интерпретацию когомологий в терминах пространства решений дифференциального уравнения.

4 Группы Ли

Важным примером многообразий, часто встречающихся в физике, являются группы Ли.

ОПРЕДЕЛЕНИЕ. **Группа Ли** – это гладкое многообразие G , являющееся одновременно группой, причем групповые операции умножения $G \times G \rightarrow G : (g, h) \rightarrow gh$ и взятия обратного элемента $G \rightarrow G : g \rightarrow g^{-1}$ являются гладкими отображениями.

Размерность группы Ли – это размерность многообразия, параметризующего операции в группе. “Жесткость” и особые свойства структур, связанных с группами Ли, связаны с требованием непрерывности, наложенным на групповую операцию и обратное отображение. В действительности справедлив глубокий результат, что функции перехода, задающие атлас группы Ли, достаточно выбрать непрерывными (C^0), чтобы в результате получить гладкое (C^∞) и даже аналитическое (C^ω) многообразие.

ПРИМЕР. Группой Ли является пространство $\mathbb{R}^n$ с обычной операцией сложения между векторами. Другой пример – группа невырожденных матриц $GL(n, \mathbb{R})$. Подгруппы Ли групп $GL(n, \mathbb{R})$ называются *матричными* группами Ли. Заметим, что не каждая группа Ли является матричной.

ПРИМЕР.

- 1) Связная компонента единицы в $GL(n, \mathbb{R})$ определяется условием $\det a_j^i > 0$.
- 2) Условие $\det a_j^i = 1$ выделяет в $GL(n, \mathbb{R})$ связную топологическую подгруппу $SL(n, \mathbb{R})$. Ее размерность $\dim SL(n, \mathbb{R}) = n^2 - 1$.
- 3) Ортогональная группа $O(n)$ является подгруппой $GL(n, \mathbb{R})$, удовлетворяющей условиям ортогональности $\sum_{k=1}^n a_i^k a_j^k = \delta_{ij}$. Она состоит из двух связных компонент и имеет размерность

$\dim O(n) = n(n-1)/2$. Связная компонента группы $O(n)$, содержащая единицу, называется специальной ортогональной группой $SO(n)$.

Особые свойства групп Ли связаны со следующими тремя наблюдениями. Во-первых, на топологическом уровне группы Ли “однородны”, то есть каждая точка многообразия не отличима от другой. С другой стороны, на алгебраическом уровне существует выделенная точка – единица группы. В-третьих, на аналитическом уровне мы имеем нелинейный закон групповой операции $z = f(x, y)$, который может быть произвольно сложным.

Рассмотрим группы Ли как группы преобразований. Группа G действует* на многообразии M , если для любого ее элемента $g \in G$ задан диффеоморфизм $T_g(x)$ такой, что $T_g T_h = T_{gh}$, $T_1 = Id$ и $T_g(x)$ гладко зависит от пары (g, x) . Группа действует на M транзитивно, если для любых точек $x, y \in M$ существует такой элемент $g \in G$, что $T_g(x) = y$. Многообразие, на котором задано транзитивное действие группы G , называется *однородным*. Подгруппа H группы G , оставляющая точку $x \in M$ на месте, называется группой *изотропии*. Однородное пространство является фактор-пространством G/H .

ПРИМЕР. Следующие пространства однородны:

- 1) сфера $S^n = O(n+1)/O(n) = SO(n+1)/SO(n)$, а также $S^{2n-1} = U(n)/U(n-1)$;
- 2) проективное пространство $\mathbb{R}P^n = O(n+1)/(O(n) \times O(1))$;
- 3) комплексное проективное пространство $\mathbb{C}P^n = U(n+1)/(U(n) \times U(1)) = S^{2n+1}/S^1$.

Каждая группа Ли является группой преобразований самой себя. Это преобразование (*регулярное представление*) Φ дается правыми (левыми) сдвигами $R_g : h' \rightarrow h'g$ ($L_g : h' \rightarrow gh'$), а сама группа называется *главным* правым (левым) однородным пространством. Отображение Φ естественным образом индуцирует отображение Φ^* действительных функций f на G по формуле $\Phi^* f(p) = f(\Phi(p))$. Индуцированное отображение векторных полей имеет вид $\Phi_* : v \in T_p G \rightarrow \Phi_* v \in T_{\Phi(p)} G$ и определяется соотношением

$$\Phi_* v(f)|_{\Phi(p)} = v(\Phi^* f)|_p$$

Алгебры Ли Алгебра векторных полей на группе Ли, будучи бесконечномерной, мало связана с групповой операцией на самой группе G . Если рассматривать только *правоинвариантные*[†] векторные поля v , удовлетворяющие условию $(R_g)_* v = v$, то значение $v(p)$ такого поля в точке p определяется его значением в точке I – единичного элемента группы:

$$v(p) = (R_g)_* v(I), \quad v(I) = (R_{g^{-1}})_* v(p)$$

Таким образом пространство всех правоинвариантных векторных полей в группе Ли образует конечномерное пространство, изоморфное касательному пространству $T_I G$.

Это векторное пространство, снабженное билинейной антисимметричной операцией – коммутатором

$[v, u] = -[u, v]$, удовлетворяющей тождеству Якоби $[[u, v], w] + [[w, u], v] + [[v, w], u] = 0$, называется **алгеброй Ли $\mathfrak{g}$** .

*Это обобщение понятия линейного представления группы на векторном пространстве

[†]Аналогично можно рассматривать левоинвариантные поля

Для задания структуры алгебры Ли на n -мерном линейном пространстве, достаточно определить попарные коммутаторы базисных векторов e_i , то есть коэффициенты c_{ij}^k в разложении $[e_i, e_j] = c_{ij}^k e_k$ (структурные уравнения алгебры Ли). Эти коэффициенты называются *структурными константами** алгебры Ли. Они должны удовлетворять условиям

$$\begin{aligned} c_{ij}^k + c_{ji}^k &= 0 & (\text{антисимметричность}) \\ c_{is}^p c_{jk}^s + c_{js}^p c_{ki}^s + c_{ks}^p c_{ij}^s &= 0 & (\text{тождество Якоби}) \end{aligned}$$

Любое правоинвариантное векторное поле $v(x) = v^i e_j(x)$ можно записать в виде линейной комбинации базисных полей $e_j(x) = l_j^k(x) \partial_k$. Структурные уравнения алгебры Ли, записанные для поля реперов, называются уравнениями Маурера:

$$l_i^r \partial_r l_j^k - l_j^r \partial_r l_i^k = c_{ij}^r l_r^k$$

Если в этих уравнениях в качестве точки x взять единицу группы e , то мы получим, учитывая, что $l_i^k(e) = \delta_i^k$, выражение для структурных констант алгебры непосредственно через групповые функции: $c_{ij}^k = \partial_i l_j^k(e) - \partial_j l_i^k(e)$.

Вместо правоинвариантных полей можно работать с правоинвариантными дифференциальными 1-формами, которые определяются условием

$$(R_g)^* \omega(v) := \omega((R_g)_* v) = \omega(v)$$

Если ω^i – базис кокасательного пространства, двойственный базису алгебры $\mathfrak{g}$, $\omega^i(e_j) = \delta_j^i$, то уравнениям Маурера соответствуют уравнения Картана

$$d\omega^k = -\frac{1}{2} c_{ij}^k \omega^i \wedge \omega^j$$

Две алгебры Ли $\mathfrak{g}$ и $\tilde{\mathfrak{g}}$ над одним и тем же полем называются *изоморфными*, если существует линейный изоморфизм $\varphi : \mathfrak{g} \rightarrow \tilde{\mathfrak{g}}$, сохраняющий коммутатор: $\varphi[u, v] = [\varphi u, \varphi v]$. Структурные константы c_{ij}^k и $\tilde{c}_{ij}^k$ изоморфных алгебр связаны соотношением

$$\varphi_k^p c_{ij}^k = \tilde{c}_{rs}^p \varphi_i^r \varphi_j^s, \quad \text{где } \varphi e_i = \varphi_i^j \tilde{e}_j$$

Для изоморфизма алгебр Ли необходимо и достаточно, чтобы существовали базисы, в которых структурные константы алгебр совпадают.

Подпространство L алгебры $\mathfrak{g}$ является *подалгеброй*, если $[L, L] \subset L$, и *идеалом*, если $[L, \mathfrak{g}] \subset L$. Существование подалгебр и идеалов в $\mathfrak{g}$ выражается в определенных ограничениях на структурные константы. Например, если $e_1, \dots, e_k$ – базис идеала, то $c_{ij}^s = 0$ при $i \leq j$, $s > k$ и произвольном j . Идеалы в алгебрах Ли играют роль, аналогичную роли нормальных подгрупп в теории групп.

ПРИМЕР.

- 1) Алгебра Ли $\mathfrak{gl}(n, \mathbb{R})$ группы $GL(n, \mathbb{R})$ – это пространство всех матриц порядка n .
- 2) Алгебра Ли $\mathfrak{o}(n)$ ортогональной группы $O(n)$ состоит из всех антисимметричных матриц

*Эти величины не являются постоянными, несмотря на название, а преобразуются как тензор третьего ранга при замене базиса в алгебре $\mathfrak{g}$.

$\{a \mid a^T = -a\}$. Она совпадает с алгеброй Ли $\mathfrak{so}(n)$ для группы $SO(n)$.

3) Алгебра Ли $\mathfrak{u}(n)$ унитарной группы $U(n)$ состоит из всех антиэрмитовых матриц $\{a \mid a^\dagger = -a\}$. Алгебра Ли $\mathfrak{su}(n)$ для группы $SU(n)$ определяется дополнительным условием бесследовости $\text{Tr } a = 0$.

Отметим, что в отличие от групп Ли, любая алгебра Ли является матричной (теорема Адо).

Алгебру Ли можно рассматривать как “линеаризацию” группы Ли в окрестности единицы. В обратном процессе по алгебре Ли с помощью экспоненциального отображения

$$g(X) = \lim_{k \rightarrow \infty} \prod (I + \frac{1}{k}X)^k = e^X$$

можно локально (в окрестности единицы) восстановить группу Ли и получить так называемую *локальную группу Ли*. В общем случае существует взаимно-однозначное соответствие между алгеброй Ли и односвязной группой Ли. Причем каждая группа Ли с одной и той же алгеброй Ли либо односвязная (*универсальная накрывающая*) группа, либо является фактор-группой универсальной накрывающей по дискретной инвариантной подгруппе (теорема Картана).

ПРИМЕР. Алгебра Ли $\mathfrak{so}(3)$ специальной ортогональной группы $SO(3)$ (группы вращений пространства $\mathbb{R}^3$) является линейной оболочкой матриц моментов импульса

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad L_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Алгебра Ли $\mathfrak{su}(2)$ специальной унитарной группы $SU(2)$ является линейной оболочкой спинных матриц

$$s_1 = \frac{i}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad s_2 = \frac{i}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad s_3 = \frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Эти алгебры изоморфны, поскольку имеют одинаковые коммутационные соотношения $[J_i, J_k] = -\varepsilon_{ikl} J_l$. Группа $SU(2)$ односвязна и гомеоморфна S^3 . Ее максимальная дискретная инвариантная подгруппа D состоит из всех матриц кратных единичной, то есть $D = \pm I \approx \mathbb{Z}_2$. Поэтому $SO(3)$ является двусвязной и $SO(3) \approx SU(2)/\mathbb{Z}_2$.

Присоединенное представление. Метрика Киллинга Композиция сдвига в группе G на элемент g слева и на g^{-1} справа,

$$h \rightarrow ghg^{-1}, \quad h, g \in G$$

называемая внутренним автоморфизмом группы $\text{Int } G : G \rightarrow G$, задает автоморфизм касательного пространства (алгебры Ли) $\text{Ad } g : \mathfrak{g} \rightarrow \mathfrak{g}$ по формуле $\text{Ad } g(v) = gvg^{-1}$. Это отображение является представлением $\text{Ad} : G \rightarrow GL(n, \mathbb{R})$ группы Ли G в n -мерном линейном пространстве алгебры Ли $\mathfrak{g}$. Оно называется *присоединенным представлением*. Для абелевых групп представление Ad тривиально, то есть $\text{Ad } g = 1$ для $\forall g \in G$.

Дифференциал отображения Ad является отображением $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(n, \mathbb{R})$, которое задает представление алгебры Ли $\mathfrak{g}$. Его можно явно записать в виде линейного оператора ad_v , действующего в $\mathbb{R}^n$:

$$\text{ad}_v w = [v, w], \quad v, w \in \mathfrak{g}$$

В базисе e_i присоединенное представление имеет вид $\text{ad}_{e_i} e_j = c_{ij}^k e_k$. Пространство всех операторов ad_v является подалгеброй и идеалом алгебры Ли $\text{Der } \mathfrak{g}$ всех дифференцирований* $\mathfrak{g}$ и называется *присоединенной алгеброй*.

На алгебре Ли можно определить скалярное произведение по формуле

$$(v, w)_K = \text{Tr}(\text{ad}_v \cdot \text{ad}_w), \quad v, w \in \mathfrak{g}$$

которая задает на $\mathfrak{g}$ инвариантную (относительно группы всех автоморфизмов алгебры $\mathfrak{g}$) метрику Киллинга (функционал Киллинга)

$$g_{ij}^K = c_{ik}^l c_{jl}^k$$

Метрика Киллинга невырождена тогда и только тогда, когда алгебра Ли полупроста. Если G – компактная группа, то форма Киллинга является отрицательно определенной и существует базис алгебры Ли, к которому она принимает вид $g_{ij}^K = -2\delta_{ij}$. В теориях калибровочных полей как правило используют компактные группы, поскольку отрицательная определенность формы Киллинга приводит к положительной определенности энергии калибровочного поля в таких моделях.

ПРИМЕР. Метрика Киллинга на алгебре Ли общей линейной группы $\mathfrak{gl}(n, \mathbb{R})$ имеет вид

$$(v, w)_K = 2n \text{Tr}(vw) - 2 \text{Tr}(v) \text{Tr}(w)$$

где vw – обычное матричное умножение, а $\text{Tr}(v)$ обозначает след матрицы v . Эта билинейная форма является вырожденной, поскольку алгебра $\mathfrak{gl}(n, \mathbb{R})$ не является полупростой.

5 Геометрия расслоений и калибровочные теории

5.1 Расслоения. Векторные и главные расслоения

ОПРЕДЕЛЕНИЕ. Гладкое расслоенное многообразие (fiber bundle), или расслоение, – это тройка объектов

(E, π, M) , состоящая из многообразия E (тотального пространства расслоения), многообразия M (базы расслоения) и проекции $E \xrightarrow{\pi} M$, причем для любой точки $x \in M$ существует открытая окрестность U такая, что имеет место *локальное расщепление*, описываемое коммутативной диаграммой†

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\varphi} & U \times F \\ \pi \downarrow & \swarrow pr_1 & \\ U & & \end{array}$$

Гомеоморфизм $\varphi : \pi^{-1}(U) \rightarrow U \times F$ локально задает на E структуру прямого произведения, а прообраз $\pi^{-1}(x)$, $\forall x \in M$ гомеоморфен многообразию F и называется (типичным) *слоем* (standard fiber) расслоения.

*Дифференцирование D алгебры $\mathfrak{g}$ – это линейное отображение $\mathfrak{g}$ в себя, удовлетворяющее правилу Лейбница: $D([v, w]) = [D(v), w] + [v, D(w)]$ для любых $v, w \in \mathfrak{g}$

†Коммутативная диаграмма – это ориентированный граф, в вершинах которого находятся объекты, а стрелками являются морфизмы (отображения), причём результат композиции стрелок не зависит от выбранного пути

Таким образом, расслоение E над M со слоем F можно представлять себе как многообразие, которое локально является прямым произведением M и F . На E можно смотреть также как на семейство многообразий F , параметризованное точками многообразия M .

Если многообразие разбито на (погруженные) подмногообразия, то оно называется *слоением* (foliation). Пространство слоев слоения, в отличие от пространства слоев расслоения, вообще говоря, не является многообразием.

Множество всех $\{U_i, \varphi_i\}$ таких, что $\varphi_i : E|_{U_i} \simeq U_i \times F$, называется *локальными тривиализациями*, или калибровками (gauge). Эти тривиализации составляют атлас расслоения $A = \{U_i, \varphi_i, f_{ij}\}$ с функциями перехода (склейки) между слоями

$$f_{ij}(x) := \varphi_i \circ \varphi_j^{-1} : (U_i \cap U_j) \times F \rightarrow (U_i \cap U_j) \times F$$

которые осуществляют гомеоморфизм слоя F и удовлетворяют условиям

$$f_{ii}(x) = \text{Id}, \quad f_{ij}(x) = (f_{ji}(x))^{-1}, \quad f_{ij}f_{jk} = f_{ik}$$

Функции перехода можно представлять себе как матричнозначные функции $T_{ij}(x)$, действующие на координаты расслоения $x \in U_j, y \in F$ по правилу

$$f_{ij} : (x, y) \rightarrow (x, T_{ij}(x)y)$$

Набор отображений f_{ij} называется *склеивающим коциклом*. Смена тривиализации расслоения называется также *калибровочным преобразованием* (пример расслоения для поля Максвелла мы рассмотрим в следующем параграфе).

Сечением (section) расслоения E называется такое непрерывное отображение $s : M \rightarrow E$, что $\pi \circ s(x) = x$ для $\forall x \in M$. Множество всех сечений E над M обозначается $\Gamma(M, E)$. В физических приложениях сечение – это просто физическое поле на многообразии M . Существенно, что хотя локальные сечения $s : U_i \rightarrow E$ существуют всегда, глобальное сечение $s : M \rightarrow E$ есть не у каждого расслоения.

Если расслоение E является прямым произведением $E = M \times F$, то оно называется *тривиальным*. Заметим, что любое расслоение по определению локально-тривиально. У тривиального расслоения всегда существует глобальное сечение. Таким образом, существование (нетривиального) глобального сечения является важной топологической характеристикой многообразия.

ПРИМЕР. *Лист Мебиуса* представляет собой простейший пример нетривиального расслоения над базой – окружностью S^1 со слоем $F = [-1, 1]$. Структура расслоения определяется атласом, состоящим из двух карт $U_1 \times F$ и $U_2 \times F$, над покрытием

$$U_1 = \{-\varepsilon < \alpha < \pi + \varepsilon\}, \quad U_2 = \{\pi - \varepsilon < \alpha < 2\pi + \varepsilon\} \\ U_1 \cap U_2 = W_1 \cup W_2, \quad W_1 = [-\varepsilon, \varepsilon], \quad W_2 = [\pi - \varepsilon, \pi + \varepsilon]$$

с функциями перехода

$$\varphi(\alpha) : v \mapsto -v, \quad \alpha \in W_1, v \in F \\ \varphi(\alpha) : v \mapsto v, \quad \alpha \in W_2, v \in F$$

Для цилиндра (тривиального расслоения над окружностью) функция перехода имеет вид

$$\varphi(\alpha) : v \mapsto v$$

Лист Мебиуса имеет глобальное сечение $(v \circ s)\alpha = 0$.

ПРИМЕР. *Касательное расслоение* (tangent bundle) TM – это $(2n)$ -мерное многообразие касательных пространств к n -мерному многообразию M . Здесь базой расслоения является многообразие M , слоем – пространство $\mathbb{R}^n$. В качестве атласа расслоения можно взять голономный атлас с координатами $(x_i^\mu, \dot{x}_i^\mu)$ и функциями перехода

$$\|f_{ij}(x)\|_\nu^\mu = \frac{\partial x_i^\mu}{\partial x_j^\nu}, \quad \dot{x}_i^\mu = \|f_{ij}(x)\|_\nu^\mu \dot{x}_j^\nu$$

Сечения касательного расслоения TM – это векторные поля на многообразии M . Заметим, что над M существуют и другие расслоения со слоем $\mathbb{R}^n$, не эквивалентные касательному TM .

Многообразие, касательное расслоение которого тривиально, называется *параллелизуемым*. Многообразие размерности n параллелизуемо тогда и только тогда, когда на нем имеется n линейно независимых в каждой точке непрерывных векторных полей. Отметим, что все группы Ли являются параллелизуемыми многообразиями.

Примером тривиального касательного расслоения служит касательное расслоение $TM = S^1 \times \mathbb{R}$ над окружностью $M = S^1$ – двумерное многообразие, диффеоморфное бесконечно-му цилиндру. В качестве примера касательного расслоения, не являющегося тривиальным, можно привести касательное расслоение TM над сферой $M = S^2$. Это четырехмерное многообразие, которое локально имеет структуру $U_i \times \mathbb{R}^2$, $U_i \in S^2$, но глобально $TM \neq S^2 \times \mathbb{R}^2$. Как следствие, на сфере нельзя построить векторного поля, которое нигде не обращалось бы в ноль (“нельзя причесать ежа”).

Подрасслоение S касательного расслоения TM на многообразии M называется *распределением* (distribution). Оно сопоставляет каждой точке $x \in M$ k -мерное подпространство S_x из $T_x M$. Распределение называется дифференцируемым, если в окрестности U любой точки x существует k дифференцируемых векторных полей $X_1, \dots, X_k$, которые образуют базис S_y для каждой точки $y \in U$. Множество $X_1, \dots, X_k$ называется локальным базисом распределения S .

Распределение S называется *инволютивным*, или *вполне интегрируемым*, если $[X, Y]$ принадлежит S , когда векторные поля X и Y принадлежат S . Каждое инволютивное распределение задает на M максимальное интегральное подмногообразие N , то есть связное подмногообразие N в M , для которого $f_*(T_x N) = S_x$ для всех $x \in N$, где f есть вложение N в M . Таким образом, каждое инволютивное распределение определяет слоение на M .

ПРИМЕР. Интегральная кривая гладкого векторного поля без особенностей, задающего одномерное распределение на многообразии M , определяет одномерное слоение.

Дифференциальные 1-формы являются сечениями *кокасательного* расслоения (cotangent bundle) T^*M с голономными координатами $(x^\mu, \dot{x}_\mu)$ и с функциями перехода

$$\|f_{ij}(x)\|_\nu^\mu = \frac{\partial x_i^\mu}{\partial x_j^\nu}, \quad (\dot{x}_i)_\nu = \|f_{ij}(x)\|_\nu^\mu (\dot{x}_j)_\mu$$

Дифференциальные формы старших порядков определяются как сечения внешнего произведения кокасательного расслоения $\bigwedge^k T^*M = T^*M \wedge \dots \wedge T^*M$ со слоем – k -кратным антисимметризованным тензорным произведением слоя $\mathbb{R}^n$ расслоения T^*M .

Векторные и главные расслоения Расслоения, как правило, наделяют некоторой алгебраической структурой. Это может быть структура линейного пространства (над полем действительных $\mathbb{R}$ или комплексных чисел $\mathbb{C}$), и тогда мы получаем *векторное расслоение*, частным случаем которого является касательное расслоение. Слой векторного расслоения изоморфен линейному пространству $\mathbb{R}^n$ ($\mathbb{C}^n$), и гомеоморфизмы тривиализации линейны на этих слоях, по сути можно считать, что на слой действуют линейные преобразования из $GL(n, \mathbb{R})$ ($GL(n, \mathbb{C})$).

Векторное расслоение с одномерным слоем (действительным или комплексным) называется *линейным*. Комплексное расслоение с голоморфными функциями перехода называется *голоморфным* расслоением.

ПРИМЕР. Построим все *линейные голоморфные расслоения* над сферой S^2 . Для этого удобно рассматривать S^2 как компактифицированную комплексную плоскость, то есть сферу Римана $S^2 = \mathbb{C} \cup \{\infty\}$. Покроем S^2 двумя картами $\{U_0, z\}$ и $\{U_\infty, w\}$, где z – это стандартная комплексная координата в окрестности $U_0 = S^2/\{\infty\} \approx \mathbb{C}$, а $w = 1/z$ – координата на $U_\infty = S^2/\{0\} \approx \mathbb{C}$. Соответствующие сечения обозначим $s_0 : U_0 \rightarrow \mathbb{C}$ и $s_\infty : U_\infty \rightarrow \mathbb{C}$. Структура расслоения будет определена с помощью функции перехода $f_{\infty 0} : U_0 \cap U_\infty \rightarrow GL(1, \mathbb{C})$

$$s_\infty = f_{\infty 0} s_0, \quad \text{где выберем } f(z) = \frac{1}{z^n}$$

Такая функция перехода определяет для каждого целого n некоторое голоморфное расслоение над S^2 , которое обозначается $\mathcal{O}(n)$. Можно доказать, что любое голоморфное линейное расслоение изоморфно $\mathcal{O}(n)$ при некотором n . Тривиальное расслоение $S^2 \times \mathbb{C}$ будет соответствовать $\mathcal{O}(0)$. Касательное расслоение TS^2 также должно быть изоморфно $\mathcal{O}(n)$ при определенном n . Поскольку $dw = -(1/z^2)dz$, то функция перехода для касательного расслоения должна иметь вид $f_{\infty 0} = dw/dz = -1/z^2$, откуда следует, что $TS^2 \cong \mathcal{O}(2)$. Аналогично для кокасательного расслоения получаем $T^*S^2 \cong \mathcal{O}(-2)$.

Отметим важный факт, что любое голоморфное векторное расслоение E над S^2 эквивалентно прямой сумме линейных расслоений $E \cong \mathcal{O}(n_1) \oplus \dots \oplus \mathcal{O}(n_k)$ (теорема Биркгофа-Гротендика).

Векторное расслоение всегда допускает глобальное сечение, например нулевое сечение $s(x) = 0$, которое сохраняется линейными преобразованиями. Так же можно выбрать ненулевое локальное сечение внутри некоторой окрестности U . Если около границы U это сечение обращается в ноль, то его можно продолжить непрерывно как нулевое сечение за границу U и таким образом получить глобальное ненулевое сечение. Заметим, что нетривиальное глобальное сечение расслоения, которое в каждой точке многообразия отлично от нуля, существует не всегда. Пример – касательное расслоение к сфере. Вообще, среди всех двумерных замкнутых ориентируемых поверхностей только касательное расслоение над тором имеет нетривиальное сечение.

Набор $s = (s_1, \dots, s_n)$ из n гладких (непрерывных, аналитических в зависимости от типа расслоения) сечений векторного расслоения над окрестностью U n -мерного многообразия M

называется *репером*, если вектора

$\{s_1(x), \dots, s_n(x)\}$ образуют базис в слое для каждой точки $x \in U$. Выбор репера эквивалентен локальным тривиализациям расслоения.

Векторное расслоение тривиально тогда и только тогда, когда оно имеет глобально определенный над каждой точкой базы репер.

Важным классом расслоений являются *главные* расслоения (principal bundle), слоем которых является группа Ли, причем определено действие группы на себя левыми сдвигами $g : G \rightarrow gG$, $g \in G$.

ОПРЕДЕЛЕНИЕ. Расслоение $P \xrightarrow{\pi} M$ называется **главным расслоением** со *структурной* (калибровочной) группой Ли G , если задано послойное свободное и транзитивное действие группы G на P справа $R_g : p \mapsto pg$, где $p \in P$, $g \in G$. База главного расслоения M диффеоморфна фактор-пространству P/G , а проекция π совпадает с каноническим отображением $P \rightarrow P/G$.

Любое главное расслоение однозначно задается выбором базы M , структурной группы G и атласом

$A = \{U_i, \varphi_i, f_{ij}\}$, где функции перехода $f_{ij} - G$ -значные гладкие функции на $U_i \cap U_j$. Функции перехода определяются расслоением неоднозначно. Если рассмотреть различные тривиализации $\{U_i, \varphi_i\}$ и $\{U_i, \psi_i\}$ и определить отображение

$$\lambda_i(x) := \varphi_i(x) \circ \psi_i^{-1}(x) : U_i \times F \rightarrow U_i \times F$$

так, чтобы оно принадлежало структурной группе G , то функции перехода $f_{ij}(x)$ и

$$f'_{ij}(x) = \lambda_i^{-1}(x) f_{ij}(x) \lambda_j(x)$$

будут определять эквивалентные расслоения, а сами коциклы $\{f_{ij}\}$ и $\{f'_{ij}\}$ называются эквивалентными.

Глобальное сечение главного расслоения P существует тогда и только тогда, когда расслоение тривиально, то есть $P = M \times G$. Для этого необходимо и достаточно, чтобы функции перехода имели вид

$$f_{ij}(x) = \lambda_i^{-1}(x) \lambda_j(x)$$

где $\lambda_i(x)$ – гомеоморфизмы слоя, принадлежащие структурной группе G .

С главным расслоением P со структурной группой G связаны так называемые *ассоциированные расслоения* (associated bundle), которые отличаются только слоем V , которой является в этом случае пространством представления группы G . Это часто обозначают как $P \times_G V \rightarrow M$.

ПРИМЕР. Ленту Мебиуса можно рассматривать как ассоциированное расслоение над окружностью со структурной группой $\mathbb{Z}_2 = \{e, g\}$. Функции перехода имеют вид

$$f_{12}(\alpha) = \begin{cases} e, & \alpha \in W_2 \\ g, & \alpha \in W_1 \end{cases}$$

ПРИМЕР. Векторное расслоение (в частности, касательное) называется ориентируемым, если все функции перехода f_{ij} могут быть выбраны так что $f_{ij} \in GL^+(n, \mathbb{R}) := \{f \in$

$GL(n, \mathbb{R}) | \det(f) > 0 \}$. Например, из функций перехода для расслоения, определяющего лист Мебиуса, видно ($f = -1$, $\alpha \in W_1$), что оно неориентируемо.

В физических приложениях сечения ассоциированных расслоений могут описывать поля материи в некотором представлении структурной (калибровочной) группы G . Обычно используется фундаментальное представление. Например, для $G = SU(n)$ сечение $\phi = s(x)$ в точке x будет иметь вид n -компонентного столбца-вектора, преобразующегося под действием матрицы $g \in SU(2)$ как $\phi \mapsto g\phi$.

Часто используется также присоединенное представление $\text{Ad} : G \rightarrow \text{Hom}(\mathfrak{g})$ группы Ли G в ее алгебре Ли $\mathfrak{g}$ – здесь роль алгебры $\mathfrak{g}$ (как линейного пространства) играет слой *присоединенного* расслоения $\text{ad}P = P \times_G \mathfrak{g}$ главного расслоения P . Для группы $SU(n)$ сечение ϕ представляет антиэрмитову бесследовую матрицу размера $n \times n$, преобразующуюся как $\phi \mapsto g\phi g^{-1}$.

ПРИМЕР. Слой присоединенного расслоения с калибровочной группой $SU(2)$ изоморфен алгебре $\mathfrak{su}(2) \cong \mathbb{R}^3$. В качестве базиса в нем можно взять матрицы Паули

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

ПРИМЕР. Касательное расслоение TM ассоциировано с главным расслоением со структурной группой $GL(n, \mathbb{R})$, которое также называют главным реперным расслоением LM .

Индукцированные расслоения Пусть задано расслоение $E \xrightarrow{\pi} M$ со слоем F и структурной группой G и отображение $f : M' \rightarrow M$. С помощью этого отображения можно построить новое расслоение $E' \xrightarrow{\pi'} M'$ с тем же слоем F и той же группой G , которое будем называть *индуцированным* расслоением (pullback bundle). Если расслоение E задано атласом $A = \{U_i, \varphi_i, f_{ij}\}$, то структура индуцированного (с помощью отображения f) расслоения с базой M' , слоем F и группой G задается так:

- 1) покрытие $M' = \cup U'_i$, где $U'_i = f^{-1}(U_i)$;
- 2) функции перехода $f'_{ij} : (U'_i \cap U'_j) \times F \rightarrow (U'_i \cap U'_j) \times F$ определяются правилом

$$f'_{ij} : (x, y) \rightarrow (x, T'_{ij}(x)y), \quad \text{где} \quad T'_{ij}(x) = T_{ij}(f(x))$$

Справедлива коммутативная диаграмма

$$\begin{array}{ccc} E' & \xrightarrow{f'} & E \\ \pi' \downarrow & & \downarrow \pi \\ M' & \xrightarrow{f} & M \end{array}$$

При индуцировании сохраняются многие важные свойства расслоений. В частности, если пространство M может быть непрерывно стянуто в точку x_0 , то расслоение $E \rightarrow M$, индуцированное отображением $f : M \rightarrow x_0$, эквивалентно прямому произведению.

Касательное и кокасательное расслоения над $E \rightarrow M$ В геометрии и физических приложениях важную роль играют касательные TE и кокасательные T^*E расслоения над многообразиями, которые сами являются расслоениями $E \xrightarrow{\pi} M$.

Если на E ввести координаты (x^μ, y^i) с законом преобразования

$$x^\mu \rightarrow x'^\mu(x^\nu), \quad y^i \rightarrow y'^i(x^\nu, y^j),$$

то голономными координатами на касательном расслоении $TE \xrightarrow{\pi_E} E$ будут величины $(x^\mu, y^i, \dot{x}^\mu, \dot{y}^i)$, где $(\dot{x}^\mu, \dot{y}^i)$ – координаты на касательных пространствах к E , заданные относительно голономного репера $(\partial_\mu, \partial_i)$. Функции перехода для расслоения TE имеют вид

$$\dot{x}^\mu = \frac{\partial x'^\mu}{\partial x^\nu} \dot{x}^\nu, \quad \dot{y}^i = \frac{\partial y'^i}{\partial y^j} \dot{y}^j + \frac{\partial y'^i}{\partial x^\mu} \dot{x}^\mu$$

Кроме того, справедлива следующая коммутативная диаграмма

$$\begin{array}{ccc} TE & \xrightarrow{T\pi} & TM \\ \pi_E \downarrow & & \downarrow \pi_M \\ E & \xrightarrow{\pi} & M \end{array}$$

где $T\pi : (x^\mu, y^i, \dot{x}^\mu, \dot{y}^i) \rightarrow (x^\mu, \dot{x}^\mu)$ определяет TE также как расслоение над TM . Ядро этого отображения $\text{Ker } T\pi$, определяет вложенное подрасслоение, называемое *вертикальным* касательным расслоением VE , состоящее только из векторов, касательных к слоям E . Координатами на VE будут $(x^\mu, y^i, \dot{y}^i)$.

Голономными координатами на кокасательном расслоении $T^*E \xrightarrow{\pi_E} E$ будут величины $(x^\mu, y^i, \dot{x}_\mu, \dot{y}_i)$, где $(\dot{x}_\mu, \dot{y}_i)$ – координаты на кокасательных пространствах к E , заданные относительно голономного репера (dx^μ, dy^i) . В отличие от TE кокасательное расслоение T^*E не образует расслоения над T^*M , что видно из закона преобразования координат

$$\dot{x}'_\mu = \frac{\partial y^i}{\partial x'^\mu} \dot{y}_i + \frac{\partial x^\nu}{\partial x'^\mu} \dot{x}_\nu$$

где первая группа слагаемых дает “лишнюю” зависимость функций перехода от координат $\dot{y}_i$. Таким образом, справедлива следующая коммутативная диаграмма

$$\begin{array}{ccc} T^*E & & \\ \pi_E \downarrow & \searrow & \\ E & \xrightarrow{\pi} & M \end{array}$$

На многообразии M , помимо обычных дифференциальных форм, можно задать формы, принимающие значения в касательном пространстве M , то есть определить *тангенциально-значную* форму ω как сечение расслоения

$$\bigwedge^k T^*M \otimes TM \rightarrow M$$

которое в голономных координатах имеет вид $\phi = \phi_{\nu_1 \dots \nu_k}^\mu(x) dx^{\nu_1} \wedge \dots \wedge dx^{\nu_k} \otimes \partial_\mu$. В этих терминах векторное поле на многообразии – это тангенциально-значная 0-форма.

Эта конструкция позволяет эффективно работать с дифференциальными формами на многообразиях E , которые сами являются расслоениями $E \rightarrow M$. Наиболее часто используются следующие.

1. Горизонтальные формы $\phi : E \rightarrow \bigwedge^k T^*M$ – все формы на E , ограниченные на подрасслоение $T^*M \subset T^*E$. В координатах они имеют вид

$$\phi = \phi_{\nu_1 \dots \nu_k}(y) dx^{\nu_1} \wedge \dots \wedge dx^{\nu_k}$$

2. Тангенциально-значные горизонтальные формы $\phi : E \rightarrow \bigwedge^k T^*M \otimes TE$ имеют вид

$$\phi = dx^{\nu_1} \wedge \dots \wedge dx^{\nu_k} \otimes (\phi_{\nu_1 \dots \nu_k}^\mu(y) \partial_\mu + \phi_{\nu_1 \dots \nu_k}^i(y) \partial_i)$$

3. Тангенциально-значные проектируемые горизонтальные формы имеют вид

$$\phi = dx^{\nu_1} \wedge \dots \wedge dx^{\nu_k} \otimes (\phi_{\nu_1 \dots \nu_k}^\mu(x) \partial_\mu + \phi_{\nu_1 \dots \nu_k}^i(y) \partial_i)$$

4. Вертикально-значные горизонтальные формы $\phi : E \rightarrow \bigwedge^k T^*M \otimes VE$ имеют вид

$$\phi = \phi_{\nu_1 \dots \nu_k}^i(y) dx^{\nu_1} \wedge \dots \wedge dx^{\nu_k} \otimes \partial_i$$

ПРИМЕР. Вертикально-значная горизонтальная 1-форма на E $\sigma : E \rightarrow T^*M \otimes VM$, или в компонентах

$$\sigma = \sigma_\mu^i(y) dx^\mu \otimes \partial_i$$

называется *припаивающей* формой. Если $E = TM$, то существует каноническая припаивающая форма $\sigma = dx^\mu \otimes \frac{\partial}{\partial \dot{x}^\mu}$.

G -структуры. Редукция структурной группы. Возможна ситуация, когда структурная группа G *редуцирована* к некоторой своей подгруппе $G_0 \subset G$. Это означает, что локальные тривиализации могут быть выбраны так, что все функции перехода принадлежат группе G_0 . В этом случае говорят о *редукции* структурной группы.

Расслоения с редуцируемой структурной группой играют важную роль в физике, где они описывают системы со *спонтанным нарушением симметрии*, например, явление Хиггса в теории элементарных частиц или эффект Мейснера в теории сверхпроводимости.

Существуют топологические препятствия для редукции структурной группы. Укажем необходимое и достаточное условие редуцируемости G .

Структурная группа G расслоения $P \rightarrow M$ редуцируема к своей замкнутой подгруппе G_0 тогда и только тогда, когда ассоциированное расслоение $P/G_0 \rightarrow M$ со слоем G/G_0 , на который группа G действует слева, допускает глобальное сечение.

В дифференциальной геометрии редукция структурной группы соответствует так называемой G -структуре. Возьмем расслоение LM линейных реперов над M , которое является главным расслоением с группой $GL(n, \mathbb{R})$. Тогда G -**структурой** на многообразии M , где G – подгруппа Ли в $GL(n, \mathbb{R})$, называется дифференцируемое главное G -подрасслоение $P(M, G) \subset LM$ расслоения LM . Реперы $r_x \in P$ называются G -реперами.

ПРИМЕР. Если $GL(n, \mathbb{R})$ может быть редуцирована к тривиальной подгруппе, состоящей из единицы, то есть $G = \{e\}$, то многообразие M является параллелизуемым. Другими словами, на M существует абсолютный параллелизм: всякому вектору в точке x можно сопоставить вектор в точке y , который в репере r_y будет иметь те же компоненты, что и в репере

r_x . Реперное расслоение LM в этом случае является тривиальным. Аналогично можно доказать, что главное расслоение будет тривиальным тогда и только тогда, когда его структурная группа может быть редуцирована к тривиальной подгруппе, состоящей из единицы.

G -структура называется *интегрируемой*, если для любой точки $x \in M$ существует координатная окрестность (U, x^μ) такая, что координатные реперы ∂_μ являются G -реперами. На пересечении двух таких окрестностей (U, x^μ) и (W, y^μ) матрица перехода $\frac{\partial y^\mu}{\partial x^\nu}$ принадлежит группе G .

Практически G -структура на многообразии M определяется сечением $s : M \rightarrow LM/G$ соответствующего ассоциированного расслоения. Часто для этого используют тензорные поля $T(x)$ на M . Действительно, множество всех линейных преобразований $\mathbb{R}^n$, оставляющих тензор $\mathbb{T}$ инвариантным, является подгруппой Ли $G \subset GL(n, \mathbb{R})$. Если структурная группа редуцируема к G , то любой репер $r_x \in P$, осуществляющий линейный изоморфизм $r_x : \mathbb{R}^n \rightarrow T_x M$, порождает изоморфизм $T_x = r_x \mathbb{T}$ тензорных алгебр в разных точках $x \in M$. Поскольку $\mathbb{T}$ инвариантен относительно действия G , то образ изоморфизма не зависит от выбора репера, и, следовательно, мы имеем корректно определенное тензорное поле $T(x)$ – сечение расслоения LM/G .

ПРИМЕР. Пусть в $\mathbb{R}^n$ задано скалярное произведение с помощью метрического тензора $g_{\mu\nu}$. Группой инвариантности этого тензора является ортогональная группа $O(n)$. При редукации структурной группы мы получаем подрасслоение $P(M, O(n)) \subset LM$, любой репер которого определяет скалярное произведение в $T_x M$ по формуле $g_x(u, v) = (r_x^{-1}u, r_x^{-1}v)$, причем это произведение не зависит от выбора репера. Таким образом $O(n)$ -структура эквивалентна существованию *римановой метрики* на многообразии. Необходимым условием для этого является паракомпактность M . Отметим, что для ориентируемых многообразий структурная группа касательного расслоения всегда редуцируема к $SO(n)$.

$O(n)$ -структура является интегрируемой, если в окрестности любой точки можно выбрать ортогональное поле координатных реперов, в котором компоненты метрического тензора имеют диагональный вид*. Для этого необходимо и достаточно, чтобы тензор кривизны был равен нулю.

С каждой G -структурой $P \subset LM$ на многообразии M связана **группа автоморфизмов** $Aut_G M$. *Автоморфизмом* G -структуры называется диффеоморфизм $\varphi : M \rightarrow M$, если индуцированный им автоморфизм расслоения реперов $\varphi' : LM \rightarrow LM$ оставляет подрасслоение P инвариантным. Заметим, что в общем случае $Aut_G M$ не является группой Ли.

ПРИМЕР. Автоморфизмы $O(n)$ -структуры должны сохранять метрику: $\varphi_* g(u, v) = g(\varphi_* u, \varphi_* v)$, образуя таким образом группу изометрий многообразия.

На практике, в виду сложности задачи, обычно ищут не группу $Aut_G M$, а строят *инфинитезимальные* автоморфизмы – векторные поля X на M , порождающие локальную 1-параметрическую группу $g_t = e^{tX}$ автоморфизмов данной G -структуры. Если G -структура задана с помощью тензорного поля $T(x)$, то векторное поле $X = v^\mu \partial_\mu$ будет инфинитезимальным автоморфизмом тогда, и только тогда, когда производная Ли тензорного поля вдоль X обращается в ноль:

$$\mathcal{L}_X T(x) = 0$$

*т. е. для собственно римановой метрики $g_{ij} = \delta_{ij}$

и задача сводится к решению системы *линейных* дифференциальных уравнений в частных производных 1-го порядка.

ПРИМЕР. Инфинитезимальные автоморфизмы $O(n)$ -структуры – это векторные поля X , удовлетворяющие уравнениям Киллинга

$$\mathcal{L}_X g_{\mu\nu} = v^\alpha \partial_\alpha g_{\mu\nu} + g_{\alpha\nu} \partial_\mu v^\alpha + g_{\mu\alpha} \partial_\nu v^\alpha = 0$$

Множество всех полей Киллинга образует конечномерную алгебру Ли $\text{iso}(M)$ размерности $m \leq \frac{1}{2}n(n+1)$, где $n = \dim M$. Например, если M – четырехмерное пространство-время Минковского, то $m = 10$, и $\text{iso}(M)$ является алгеброй Ли группы Пуанкаре.

5.2 Связности на расслоениях

Сечения расслоений можно складывать между собой и умножать на функции на M . Можно также задать дифференцирование сечений таким образом, чтобы оно не зависело от выбора локальных тривиализаций. Таким образом получается простейшее определение *линейной связности* как *ковариантного дифференцирования*.

ОПРЕДЕЛЕНИЕ. **Ковариантной производной** ∇ на векторном расслоении $E \xrightarrow{\pi} M$ называется дифференциальный оператор 1-го порядка $\nabla : \Gamma(E) \rightarrow \Gamma(T^*M \times E)$, удовлетворяющий правилу Лейбница $\nabla(fs) = df \otimes s + f\nabla s$, где s – гладкое сечение $E \rightarrow M$, а f – гладкая функция на M .

С общей точки зрения, связность на расслоении E является дополнительной геометрической структурой, которая определяет правило *параллельного переноса* – сопоставление всякому пути γ в базе M семейства отображений слоев F_γ расслоения E над точками этого пути. В инфинитезимальной форме задание связности можно наглядно представить себе как задание в каждой точке расслоения $y \in E$ направления, в котором она будет переноситься в многообразии E , если ее проекция $x = \pi(y)$ переносится в некотором направлении в базе M . Например, для главного расслоения $P \rightarrow M$ объект связности будет 1-формой на M со значениями в алгебре $\mathfrak{g}$.

С другой стороны, если говорить на языке касательных пространств, то для каждой точки y расслоения E определяется поднятие касательного пространства $T_x M$ к базе M в касательное пространство $T_y E$ к расслоению E , то есть связность определяет *горизонтальное расщепление* касательного расслоения $TE = HE \oplus VE$, что можно записать как послойное отображение

$$\dot{x}^\mu \partial_\mu + \dot{y}^i \partial_i \mapsto \dot{x}^\mu (\partial_\mu + \Gamma_\mu^i \partial_i)$$

Таким образом, получаем геометрическое определение **связности** как тангенциально-значной проектируемой горизонтальной 1-формы Γ на E такой, что свертка $\Gamma \lrcorner \phi = \phi$ для всех горизонтальных форм ϕ . В голономных координатах на E связность имеет вид

$$\Gamma = dx^\mu \otimes (\partial_\mu + \Gamma_\mu^i(y) \partial_i)$$

с законом преобразования

$$\Gamma_\mu^i = \frac{\partial x^\nu}{\partial x'^\mu} \left(\frac{\partial y'^i}{\partial y^j} \Gamma_\nu^j + \frac{\partial y'^i}{\partial x^\nu} \right)$$

Приведенное выше определение линейной связности можно получить из общего определения, если потребовать, чтобы коэффициенты $\Gamma_\mu^i(y)$ линейно зависели от координат слоя y^i . Таким образом, **линейная связность** на векторном расслоении $E \rightarrow M$ – это 1–форма

$$\Gamma = dx^\mu \otimes (\partial_\mu - \Gamma_{\mu j}^i(x) y^j \partial_i)$$

Далее, если не оговорено противное, мы будем под связностью понимать именно линейную связность.

В физических приложениях связности на векторном расслоении $E \rightarrow M$ часто рассматривают как 1–формы на M со значениями в E . Такие формы будем обозначать как $A(x)$. Заметим, что форма связности A не имеет глобального описания на M . Ковариантное дифференцирование, определяемой этой связностью, есть просто линейное отображение $\nabla : \Omega^0(M, E) \rightarrow \Omega^1(M, E)$. Если выбрать локальные тривиализации, то есть над каждой окрестностью $U \subset M$ фиксировать репер $f = (e_1, \dots, e_r)$, то мы получим локальное представление связности

$$\nabla e_i = A_i^j e_j$$

где $A_i^j = A_i^j(f, \nabla)$ – $r \times r$ -матричнозначная 1-форма над U . Если $s \in \Omega^0(M, E)$, то мы получаем локальное представление ковариантной производной

$$\nabla s = (ds + A \cdot s)^i e_i$$

При замене репера $f \mapsto f' = f \cdot g$, где $g : U \rightarrow GL(n)$, мы получим

$$A \mapsto A' = g^{-1} A g + g^{-1} dg$$

Можно расширить определение дифференцирования на формы произвольного порядка и ввести (единственный) оператор $\nabla : \Omega^p(M, E) \rightarrow \Omega^{p+1}(M, E)$, удовлетворяющий условию

$$\nabla(\omega s) = d\omega \wedge s + (-1)^q \omega \wedge \nabla s, \quad \omega \in \Omega^q(M), s \in \Omega^p(M, E)$$

В локальном представлении

$$\nabla s = ds + A \wedge s, \quad s \in \Omega^p(M, E)$$

Последовательность отображений

$$\Omega^0(M, E) \xrightarrow{\nabla} \Omega^1(M, E) \xrightarrow{\nabla} \Omega^2(M, E) \xrightarrow{\nabla} \dots$$

будет образовывать цепной комплекс (обобщенный комплекс де Рама), только в том случае, если оператор **кривизны** связности $F := \nabla \circ \nabla = \nabla^2 : \Omega^0(M, E) \rightarrow \Omega^2(M, E)$ обращается в ноль (связность в этом случае называется *плоской*). Легко проверить, что локально

$$Fs = (dA + A \wedge A) \wedge s, \quad s \in \Omega^p(M, E)$$

Производные сечений расслоения E не входят в 2-форму кривизны, в следствие этого кривизну можно рассматривать как $\text{End}(E)$ -значную 2-форму, где $\text{End}(E)$ – эндоморфизмы слоя.

Кривизна удовлетворяет важному дифференциальному тождеству Бианки

$$\nabla F = 0$$

Если на многообразии M задать векторное поле X , то по связности ∇ можно построить оператор $\nabla_X(s) := X \lrcorner \nabla(s) : \Gamma(M, E) \rightarrow \Gamma(M, E)$, определяющий ковариантное дифференцирование вдоль поля X . Этот оператор является обобщением обычной производной по направлению. Будучи дифференциальной 2-формой, F_∇ является антисимметричной билинейной формой, действующей на касательном расслоении, поэтому свертка $F(X, Y)$ определяет линейный оператор $F(X, Y) : \Gamma(M, E) \rightarrow \Gamma(M, E)$, и мы получаем классическое выражение для кривизны как коммутатора ковариантных дифференциальных операторов:

$$F(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]}$$

Связности на главных расслоениях Опишем более подробно связности на главных расслоениях $P \rightarrow M$, где слой и функции перехода принадлежат некоторой группе Ли G . Для простоты изложения будем считать, что эта группа матричная.

Пусть $\mathfrak{g}$ – алгебра Ли структурной группы G . Введем $\mathfrak{g}$ -значную 1-форму $g^{-1}dg$ ($g \in G$), которая называется *формой Маурера-Кармана*. Она инвариантна относительно левых сдвигов на постоянный элемент $g_0 \in G$ и ее можно разложить по базису $\{\Phi^a\}$ левоинвариантных 1-форм: $g^{-1}dg = \Phi^a \lambda_a$, где λ_a – постоянные элементы (матрицы) алгебры $\mathfrak{g}$. Базис $\{\Phi^a\}$ удовлетворяет *уравнениям Маурера-Кармана*

$$d\Phi^a + \frac{1}{2} c^a_{bc} \Phi^b \wedge \Phi^c = 0$$

где c^a_{bc} – структурные константы алгебры $\mathfrak{g}$. Аналогично можно работать с базисом левоинвариантных векторных полей

$$e_a = \text{Tr} \left(g \lambda_a \frac{\partial}{\partial g^T} \right) = g^k_j \lambda_a^l \frac{\partial}{\partial g^l_j}$$

который двойственен базису 1-форм: $e_a \lrcorner \Phi^b = \delta_a^b$.

Параллельный перенос будет определяться связностью $A(x) = A_\mu^a(x) \lambda_a dx^\mu$ – матричнозначной формой на базовом многообразии M . Эта связность задает инвариантное относительно правых сдвигов расщепление касательного пространства TP к расслоению P на горизонтальную компоненту HP с базисом

$$\nabla_\mu = \frac{\partial}{\partial x^\mu} - A_\mu^a \bar{e}_a$$

и вертикальную компоненту VP с базисом $\bar{e}_a$:

$$TP = HP \oplus VP$$

где $\bar{e}_a = \text{Tr} \left(\lambda_a g \frac{\partial}{\partial g^T} \right)$ – базис правоинвариантных векторных полей на G . Кривизна связности определяется как коммутатор

$$[\nabla_\mu, \nabla_\nu] = -F_{\mu\nu}^a \bar{e}_a, \quad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + c_{bc}^a A_\mu^b A_\nu^c$$

откуда видно, что кривизна имеет только вертикальные компоненты.

При калибровочном преобразовании (автоморфизме главного расслоения), когда координата слоя g над окрестностью U связана с координатой g' над окрестностью U' функцией перехода $\varphi: g' = \varphi g$, связность на M преобразуется нетензорным образом

$$A' = \varphi A \varphi^{-1} + \varphi d\varphi^{-1}$$

Кривизна, напротив, имеет тензорный закон преобразования $F' = \varphi F \varphi^{-1}$.

В случае, когда структурной группой G является группа $O(n)$, связность называют *ортогональной*, и локально A представима 1-формой со значениями в алгебре $\mathfrak{o}(n)$ антисимметричных матриц $n \times n$. Если же G является группой $U(n)$, то связность называют *унитарной*, и локально A представима 1-формой со значениями в алгебре $\mathfrak{u}(n)$ антиэрмитовых матриц $n \times n$.

Если работать со связностью как с проектируемой формой ω , заданной на кокасательном расслоении T^*P к расслоению P , то связность имеет вид

$$\omega = g^{-1}Ag + g^{-1}dg$$

где форма Маурера-Картана описывает вертикальную компоненту связности. Эта форма преобразуется тензорным образом $\omega \rightarrow g_0^{-1}\omega g_0$ при правых сдвигах $g \rightarrow gg_0$ (форма A является инвариантной), постоянна на векторах вертикального базиса $e_a \lrcorner \omega = \lambda_a$ и обращается в ноль на векторах горизонтального базиса $\nabla_\mu \lrcorner \omega = 0$.

Кривизна $\Omega = d\omega + \omega \wedge \omega$ связности ω имеет вид

$$\Omega = g^{-1}Fg, \quad F = dA + A \wedge A = \frac{1}{2}F_{\mu\nu}^a \lambda_a dx^\mu \wedge dx^\nu$$

При калибровочном преобразовании $g' = \varphi g$ 1-форма связности ω и 2-форма кривизны Ω сохраняют свой вид в новых координатах слоя:

$$\begin{aligned} \omega &= g^{-1}Ag + g^{-1}dg = g'^{-1}Ag' + g'^{-1}dg' \\ \Omega &= g^{-1}Fg = g'^{-1}F'g' \end{aligned}$$

являясь объектами, глобально определенными на T^*P .

Связности в калибровочных теориях В калибровочных теориях элементарных частиц обычно работают не с главными расслоениями $P \rightarrow M$, а с ассоциированными с ними расслоениями $E = P \times_G V \rightarrow M$. Структурная (калибровочная) группа G расслоения описывает внутренние симметрии частиц, а поля материи, принадлежащие некоторому представлению G , соответствуют сечениям $\phi(x) \in V$ этого расслоения. Линейная связность на расслоении E представляет калибровочное поле, точнее его калибровочный потенциал A .

Обычно используют присоединенное представление группы G , а связность записывают как дифференциальную форму, принимающую значения в алгебре Ли $\mathfrak{g}$, то есть $A(x) \in \Gamma(T^*M \otimes \mathfrak{g})$.

Компоненты ковариантной производной, определяемой связностью, в присоединенном представлении имеют вид

$$\nabla_\mu \phi = \frac{\partial \phi}{\partial x^\mu} + [A_\mu, \phi]$$

а для коммутаторов получаем выражение $[\nabla_\mu, \nabla_\nu]\phi = [F_{\mu\nu}, \phi]$, где

$$F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu + [A_\mu, A_\nu]$$

Здесь $F_{\mu\nu}$ – компоненты $\mathfrak{g}$ -значной 2-формы на M , которая описывает напряженность калибровочного поля.

Электромагнитное поле (поле Максвелла) описывают связностью на расслоении с калибровочной группой $U(1)$, а слабые и сильные ядерные взаимодействия (поля Янга-Миллса) соответствуют группам $SU(2)$ и $SU(3)$.

В слое присоединенного расслоения $P \times_G \mathfrak{g}$ обычно дополнительно вводят скалярное произведение с помощью билинейной симметрической формы Киллинга $(X, Y) = \text{Tr}(\text{ad}_X \cdot \text{ad}_Y)$ на алгебре $\mathfrak{g}$. Для групп $SU(n)$ метрика Киллинга имеет простой вид

$$(X, Y) = \text{Tr}(XY)$$

где $X, Y \in \mathfrak{su}(n)$ – бесследовые антиэрмитовы матрицы.

С помощью метрики $g_{\mu\nu}$ на базе(пространстве-времени) M и (невыврожденной) метрики Киллинга в слое $\mathfrak{g}$ можно ввести лагранжиан $\mathcal{L}_{YM} := \frac{1}{4}g^{\mu\alpha}g^{\nu\beta}(F_{\mu\nu}, F_{\alpha\beta})$ для функционала действия Янга-Миллса $S_{YM} = \int \mathcal{L}_{YM} \nu_M$. Функция $\mathcal{L}_{YM} := \frac{1}{4}\text{Tr}(F_{\mu\nu}F^{\mu\nu})$ корректно определена на всем M в силу ее инвариантности при калибровочных преобразованиях 2-формы кривизны $F \mapsto \varphi F \varphi^{-1}$.

Калибровочное поле F автоматически удовлетворяет тождеству Бианки

$$\nabla F = 0, \quad \text{или} \quad \nabla_{[\alpha} F_{\mu\nu]} = 0$$

в то время как уравнения Эйлера-Лагранжа для функционала S_{YM} дают уравнения Янга-Миллса

$$\nabla \star F = 0, \quad \text{или} \quad \nabla^\mu F_{\mu\nu} = 0$$

где в присоединенном представлении $\nabla_\alpha F_{\mu\nu} = \partial_\alpha F_{\mu\nu} + [A_\alpha, F_{\mu\nu}]$.

ПРИМЕР. Монополь Дирака.

Монополь Дирака представляет собой пример простейшего калибровочного поля, в данном случае поля Максвелла, на топологически нетривиальном многообразии $\mathbb{R}^3/0$. Такое многообразие стягиваемо к S^2 , поэтому для описания монополя необходимо классифицировать все $U(1)$ расслоения над сферой. Группы гомологий сферы имеют вид

$$H_0(S^2) = \mathbb{Z}, \quad H_1(S^2) = 0, \quad H_2(S^2) = \mathbb{Z}$$

В силу нетривиальности группы $H_2(S^2) = \mathbb{Z}$ можно ожидать, что мы получим различные неэквивалентные расслоения.

Покроем базу расслоения S^2 двумя картами H_+ (верхняя полусфера) и H_- (нижняя полусфера) с угловыми координатами $(\theta, \alpha): 0 \leq \theta \leq \pi, 0 \leq \alpha \leq 2\pi$. Слой расслоения является групповое пространство $U(1) \approx S^1$ с координатой $\exp(i\psi)$. Полусферы, пересекающиеся по экватору, являются областями тривиализации расслоения. Над верхней полусферой $U(1)$ расслоение является произведением $H_+ \times U(1)$ с координатами $(\theta, \alpha, \exp(i\psi_+))$, а над нижней – произведением $H_- \times U(1)$ с координатами $(\theta, \alpha, \exp(i\psi_-))$. Функция перехода (склейки слоев)

$\varphi: \exp(i\psi_-) = \varphi \exp(i\psi_+)$ должна принадлежать группе $U(1)$, зависеть только от α и быть по α периодичной. Эти требования приводят к следующим функциям:

$$e^{i\psi_-} = e^{in\alpha} e^{i\psi_+}$$

где целое число n является топологическим инвариантом и позволяет классифицировать неэквивалентные расслоения. При $n = 0$ мы имеем тривиальное расслоение $P_{n=0} = S^2 \times S^1$. При $n \neq 0$ все расслоения только локально-тривиальны, в частности, при $n = 1$ мы получаем известное расслоение Хопфа $P_{n=1} = S^3 \rightarrow S^2$.

Введем теперь на этом расслоении линейную связность – дифференциальную 1-форму A со значениями в алгебре Ли $\mathfrak{g} = i\mathbb{R}$ группы $U(1)$. Локальные формы связности над соответствующими областями тривиализации связаны соотношением $A_- = \varphi A_+ \varphi^{-1} + \varphi d\varphi^{-1}$, что после подстановки $\varphi = \exp(in\alpha)$ дает

$$A_- = A_+ - in d\alpha$$

Здесь слагаемое $in d\alpha$ описывает калибровочное преобразование связности (электромагнитного потенциала) при смене тривиализации над пересечением карт H_+ и H_- . Алгебра мнимых чисел $i\mathbb{R}$, очевидно, изоморфна алгебре действительных чисел $\mathbb{R}$, поэтому для удобства мы далее будем считать, что формы $A_{\pm}$ принимают значения в $\mathbb{R}$, то есть являются обычными вещественными дифференциальными формами.

Мы ищем связности (калибровочные потенциалы), которые удовлетворяют уравнениям Максвелла в пространстве $\mathbb{R}^3/0$ и описывают поле магнитного монополя Дирака. Такое калибровочное поле соответствует связностям, которые локально задаются формами

$$A_{\pm} = \frac{n}{2}(\pm 1 - \cos(\theta))d\alpha = \frac{n}{2r} \frac{xdy - ydx}{z \pm r}$$

Эти формы связаны полученным выше калибровочным преобразованием и регулярны над своей областью тривиализации. Заметим также, что форму A_+ нельзя регулярно продолжить в область нижней полусферы H_- , где она становится сингулярной на струне $z < 0, x = y = 0$ (“дираковская струна”).

Калибровочное поле (2-форма кривизны связности) имеет вид

$$F = dA_{\pm} = \frac{n}{2} \sin(\theta) d\theta \wedge d\alpha = \frac{n}{2r^3} (xdy \wedge dz + ydz \wedge dx + zdx \wedge dy)$$

Оно удовлетворяет уравнениям Максвелла $dF = 0, \delta F = 0$ в $\mathbb{R}^3/0$ и описывает поле точечного магнитного заряда величины n . Сделаем несколько замечаний.

- 1) В силу коммутативности калибровочной группы $U(1)$ форма кривизны остается инвариантной при калибровочных преобразованиях: $F_- = \varphi F_+ \varphi^{-1} = F_+ = F$.
- 2) 2-форма кривизны F является замкнутой ($dF = 0$), но не является точной, так как не существует глобально определенной на $\mathbb{R}^3/0$ 1-формы A такой, что $F = dA$. Это следствие нетривиальности группы гомологий $H_2(S^2) = \mathbb{Z}$.
- 3) Квантование магнитного заряда n связано с нетривиальной топологической структурой самого расслоения P_n , что в свою очередь обусловлено топологией базы S^2 .

5.3 Характеристические классы

При изучении расслоений естественным образом возникает два вопроса. Во-первых, сколько неэквивалентных расслоений E можно построить для заданных многообразий базы M и слоя F ? Во-вторых, существуют ли какие-либо критерии, которые говорили бы нам, насколько сильно данное расслоение E отличается от тривиального $M \times F$?

Для ответа на второй вопрос заметим, что топология расслоения фактически определяется топологией базового многообразия. Например, любое расслоение над стягиваемой базой всегда тривиально. Можно сказать, что для того, чтобы получить нетривиальное векторное расслоение, нужно иметь нетривиальную топологию базового многообразия. В свою очередь, расслоение можно использовать как инструмент для исследования и описания топологических свойств базы. Такими топологическими инвариантами, определенными с помощью расслоений, являются характеристические классы. Можно сказать, что характеристический класс измеряет отклонение структуры локального произведения от структуры глобального произведения слоя и базы.

Характеристические классы – это алгебраические конструкции, элементы групп когомологий базового многообразия M , заданные с помощью расслоения $E \rightarrow M$. Такие конструкции сопоставляют расслоению замкнутую форму в базе, интегралы от которой по циклам являются топологическими инвариантами.

Общий подход к построению характеристических классов для главного $P \rightarrow M$ или ассоциированного $P \times_G V$ расслоения основан на изучении так называемого универсального расслоения EG . Для фиксированной размерности n и группы G универсальное расслоение – это, грубо говоря, такое расслоение, к которое можно вложить любое другое расслоение над многообразием размерности n и с группой G .

ОПРЕДЕЛЕНИЕ. Главное расслоение $EG \rightarrow M$ с группой G называется **универсальным** для данной группы G , если EG стягиваемо (или все гомотопические группы $\pi_i(EG) = 0$).

Справедлива следующая *классификационная теорема*: множество $S(M, G)$ расслоений с данной базой M и группой G (с точностью до эквивалентности) совпадает с множеством гомотопических классов отображений базы M в базу BG универсального расслоения, которая называется *классифицирующим пространством*.

Таким образом для заданной группы G множество $S(M, G)$ зависит только от гомотопического типа пространства M и является гомотопическим инвариантом. Фактически все расслоения получаются из отображений $f : M \rightarrow BG$ в классифицирующее пространство как индуцированные расслоения.

Характеристические классы для расслоения $E \rightarrow M$ представляют собой элементы когомологий $c(E) \in H^*(M, R)$ с коэффициентами в кольце R . Можно доказать, что на самом деле они однозначно определяются классами $c(EG) \in H^*(BG)$ универсального расслоения $EG \rightarrow BG$. Построение этих классов основывается на следующих рассуждениях.

Неэквивалентные расслоения с группой G могут быть классифицированы гомотопически классами отображений $f : M \rightarrow BG$. Далее, гомотопические группы $\pi_{k+1}(BG)$ и $\pi_k(G)$ изоморфны в силу стягиваемости EG . Теорема Гуревича устанавливает связь между когомологиями $H^*(BG)$ и $H^*(G)$. Поскольку когомологии группы Ли определяются когомологиями ее алгебры Ли, то нам необходимо найти отображение “ $H^*(\mathfrak{g}) \rightarrow H^*(M)$ ”, соответствующее $f^* : H^*(BG) \rightarrow H^*(M)$.

Обозначим через $I^k(G)$ множество симметричных полилинейных отображений

$$f : \underbrace{\mathfrak{g} \times \dots \times \mathfrak{g}}_k \rightarrow \mathbb{R}$$

и определим однородный полином $p(X) = f(X, \dots, X)$ степени k от элементов $X \in \mathfrak{g}$. Функции $f(X_1, \dots, X_k)$ называются поляризациями для полинома $p(X)$. Значения этих функций могут быть восстановлены по значениям полинома p , например при $k = 2$: $f(X, Y) = \frac{1}{2}(p(X + Y) - p(X) - p(Y))$. Далее будем рассматривать только инвариантные полиномы и, соответственно, поляризации, для которых

$$p(gXg^{-1}) = p(x), \quad g \in G$$

Введем на расслоении $E \rightarrow M$ связность, и пусть 2-форма F – кривизна этой связности. Эта форма не определена глобально на M , так как при калибровочных преобразованиях она изменяется по закону $F \mapsto gFg^{-1}$. Напротив, если p – определенный выше инвариантный полином, то $p(F) \in \Omega^{2k}(M)$ – глобально определенная $(2k)$ -форма на M . Эта форма обладает двумя важными свойствами:

- 1) $p(F)$ – замкнутая форма, то есть $dp(F) = 0$;
- 2) если F_1 и F_2 – две формы кривизны, соответствующие двум различным связностям, то существует $(2k - 1)$ -форма χ на M такая, что $p(F_1) - p(F_2) = d\chi$.

Таким образом *характеристическая форма* $p(F)$ определяет класс когомологий де Рама, независимый от связности на M . Это дает нам алгебраический *гомоморфизм Вейля* $I(G) = \sum I^k(G) \rightarrow H^*(M, \mathbb{R})$.

Классы Чженя Классы Чженя являются характеристическими классами для комплексных расслоений с группой $GL(n, \mathbb{C})$. Инвариантный полином $p(X)$ можно построить следующим образом. Пусть матрица $X \in \mathfrak{g}$ имеет собственные значения $\lambda = \{\lambda_1, \dots, \lambda_n\}$, тогда полином $p(X)$ является симметричной функцией от λ_i . Если ввести симметричные однородные полиномы

$$S_j(\lambda) = \sum_{i_1 < \dots < i_j} \lambda_{i_1} \dots \lambda_{i_j}$$

то инвариантный полином можно записать в виде следующего разложения

$$p(X) = a + bS_1(\lambda) + cS_2(\lambda) + \dots$$

Примером инвариантного полинома будет

$$\det(I + X) = 1 + S_1(\lambda) + \dots + S_k(\lambda)$$

Используя $\mathfrak{gl}(n, \mathbb{C})$ -значную форму кривизны F на расслоении, определим *полную форму Чженя* (инвариантный полином) как

$$c(F) = \det(I + \frac{i}{2\pi} F) = 1 + c_1(F) + c_2(F) + \dots$$

где $(2j)$ -формы $c_j(F)$ (полиномы степени j по F) называются *формами Чженя*. Классы когомологий этих форм являются характеристическими классами Чженя расслоения E и не зависят от выбора связности.

Для форм Чженя можно получить явные формулы, если использовать разложение по однородным полиномам для $X = \text{diag}(\lambda_1, \dots, \lambda_k)$:

$$\det(I + \frac{i}{2\pi}X) = (1 + \frac{i}{2\pi}\lambda_1) \dots (1 + \frac{i}{2\pi}\lambda_k) = 1 + \frac{i}{2\pi}S_1(\lambda) + \dots$$

где, например, $S_1(\lambda) = \text{Tr } X$, $S_2(\lambda) = \sum_{i < j}^n \lambda_i \lambda_j = \frac{1}{2}((\text{Tr } X)^2 - \text{Tr } (X^2))$ и т. д. Таким образом, получаем следующие выражения

$$\begin{aligned} c_1 &= \frac{i}{2\pi} \text{Tr } F \\ c_2 &= \frac{1}{8\pi^2} (\text{Tr } (F \wedge F) - \text{Tr } F \wedge \text{Tr } F) \\ &\dots \end{aligned}$$

Поскольку $c_i(F) \in \Omega^{2i}(M)$, то все $c_j = 0$ при $(2j) > n = \dim M$.

Заметим, что для групп $GL(n, \mathbb{C})$ и $U(n)$ инвариантные полиномы могут быть отождествлены, и характеристические классы таких расслоений совпадают. Для $SU(n)$ расслоений $c_1 = 0$, поэтому если для комплексного векторного расслоения $c_1 \neq 0$, то не существует ассоциированного ему главного $SU(n)$ расслоения. Обратное утверждение неверно, так как существуют расслоения с $c_1 = 0$, которые не допускают $SU(n)$ структуру.

Классы когомологий, к которым принадлежат классы Чженя, в действительности являются элементами целочисленных групп когомологий, поэтому можно проинтегрировать $(2i)$ -форму $c_i(F)$ по $(2i)$ -циклу на M с целыми коэффициентами и получить целое число, не зависящее от выбора связности. В случае компактной базы расслоения можно определить *числа Чженя*, как результат интегрирования инвариантных полиномов по всему многообразию M . Например, для 4-мерного пространства получается два числа Чженя

$$C_2(E) = \int_M c_2(F) \quad \text{и} \quad C_1^2(E) = \int_M c_1(F) \wedge c_1(F)$$

ПРИМЕР. Построим полную форму Чженя $c(F)$ и найдем числа Чженя для $U(1)$ расслоения над сферой S^2 , описывающего магнитный монополю Дирака. База расслоения двумерна, поэтому отлична от нуля только 2-форма Чженя $c_1(F)$. Используя форму кривизны $F = i \frac{n}{2} \sin(\theta) d\theta \wedge d\alpha$ получаем

$$c(F) = 1 + c_1(F), \quad c_1(F) = \frac{i}{2\pi} \text{Tr } F = -\frac{n}{4\pi} \sin \theta d\theta \wedge d\alpha$$

Число Чженя

$$C_1 = \int_{S^2} c_1 = -\frac{n}{4\pi} \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\alpha = -n$$

В этом примере число Чженя $C_1 = -n$ параметризует неэквивалентные $U(1)$ расслоения над S^2 , причем для тривиального расслоения $E = S^2 \times U(1)$ число Чженя $C_1 = 0$. С физической точки зрения число Чженя здесь – магнитный заряд монополя.

6 Пространство струй и геометрия дифференциальных уравнений

Пусть M и N – гладкие многообразия и φ, ψ – гладкие отображения из M в N , заданные в некоторой окрестности точки $x \in M$. Они называются *эквивалентными* порядка k в точке x , если $\varphi(x) = \psi(x)$ и имеют равные производные в этой точке до порядка k включительно.

ОПРЕДЕЛЕНИЕ. **Струя** k -го порядка (k -jet) $j_x^k \varphi$ отображения φ в точке x – это класс всех эквивалентных φ отображений с началом x и концом $y = \varphi(x)$.

В картах (U, x^μ) на M и (V, y^i) на N , где $y^i = \varphi^i(x^1, \dots, x^m)$, $(\mu = 1, \dots, m; i = 1, \dots, n)$, k -струя имеет компоненты

$$j_x^k \varphi = (x^\mu, y^i, y_{\nu_1}^i, \dots, y_{\nu_1 \dots \nu_k}^i), \quad \text{где } y_{\nu_1 \dots \nu_k}^i = \frac{\partial^k \varphi^i(x)}{\partial x^{\nu_1} \dots \partial x^{\nu_k}}$$

Поэтому k -струю можно отождествить с отрезком ряда Тейлора длины k при разложении функций $\varphi^i(x)$ в окрестности точки x .

Композиция двух отображений $M \xrightarrow{\varphi} N \xrightarrow{\psi} P$ приводит к композиции соответствующих k -струй $j_x^k(\psi \circ \varphi) = j_{\varphi(x)}^k \psi \circ j_x^k \varphi$. Если $z^a = \psi^a(y^1, \dots, y^n)$ ($a = 1, \dots, p = \dim P$), а $j_x^k(y^i, z^a, \dots, z_{i_1 \dots i_k}^a)$ – компоненты соответствующей струи, то струя $j_x^k(\psi \circ \varphi)$ имеет компоненты, образованные по правилу дифференцирования сложных функций

$$j_x^k(\psi \circ \varphi)_{\mu_1 \dots \mu_s}^a = \sum_{r=1}^s z_{i_1 \dots i_r}^a y_{\mu_1 \dots \mu_s}^{i_1 \dots i_r},$$

где

$$y_{\mu_1 \dots \mu_s}^{i_1 \dots i_r} = \frac{s!}{r!} \sum_{s_1 + \dots + s_r = s} \frac{1}{s_1! \dots s_r!} y_{(\mu_1 \dots \mu_{s_1}}^{i_1} \dots y_{\mu_{s_{r-1}} \dots \mu_{s_r}}^{i_r}$$

Например композиция 1-струй имеет компоненты $j_x^1(\psi \circ \varphi)_\mu^a = z_i^a y_\mu^i$.

Множество всех k -струй отображений $M \rightarrow N$ образует гладкое многообразие (*пространство струй*) $J^k(M, N)$. Проекция $\pi : J^k(M, N) \rightarrow M \times N$, которая струе ставит в соответствие ее начало и конец, расслаивает это многообразие над базой $M \times N$. Координатами в этом расслоении являются компоненты струи $j_x^k \varphi$. Поскольку число частных производных $y_{\nu_1 \dots \nu_k}^i$ от функции m переменных $y(x^\mu)$ равно биномиальному коэффициенту $\binom{m+k-1}{k} = \frac{(m+k-1)!}{k!(m-1)!}$, то размерность многообразия струй $J^k(M, N)$ равна

$$\dim J^k(M, N) = m + n \sum_{s=0}^k \binom{m+s-1}{m-1} = m + n \binom{m+k}{k} = m + n \frac{(m+k)!}{k! m!}$$

Пусть на многообразии N задано гладкое отображение $y \rightarrow y' = f(y)$. Оно порождает отображение $f^{(k)}$ в многообразии струй $J^k(M, N)$, которое струю $j^k \varphi$ переводит в $j^k(f \circ \varphi)$. Преобразование $j_x^k(f \circ \varphi) = j_y^k f \circ j_x^k \varphi$ называется *дифференциальным продолжением* (prolongation) k -го порядка, или *k -лифтом*, преобразования $f(y)$.

Рассмотрим отображение $\varphi(t) : \mathbb{R} \rightarrow M$ прямой в n -мерное пространство M и обозначим $\pi : T^k M \rightarrow M$ соответствующее расслоение k -струй. Компоненты струи $j_x^k \varphi = (x^\mu, x_1^\mu, \dots, x_k^\mu)$

задаются значениями производных при $t = 0$ ($x = \varphi(0)$, $x_s^\mu = \frac{d^s x^\mu}{dt^s}(0)$). $T^k M$ является гладким многообразием размерности $(k+1)n$, а расслоение, определяемое проекцией $\pi : j_x^k \varphi \rightarrow x$, называется *касательным расслоением k -го порядка*. При $k = 1$ мы имеем обычное касательное расслоение TM над M .

Введенные выше конструкции непосредственно формулируются для случая расслоенных многообразий. Если $\pi : E \rightarrow M$ – гладкое расслоение и $s : M \rightarrow E$ – его гладкое сечение, то множество струй $j_x^k s$ с началом $x \in M$ и концом $y = s(x) \in \pi^{-1}(x)$ образует гладкое расслоение над M с проекцией $\pi^k : J^k(M, E) \rightarrow M$. Локальными координатами на $J^k(M, E)$ являются компоненты струи $(x^\mu, y^i, y_{\mu_1}^i, \dots, y_{\mu_1 \dots \mu_k}^i)$, где $x^\mu = (x^1, \dots, x^m)$ – координаты в карте $U \subset M$ с тривиализацией $\pi^{-1}(U) \rightarrow U \times F$, а $y^i = (y^1, \dots, y^n)$ – координаты слоя F .

Поскольку $y_{\mu_1}^i = \partial_{\mu_1} y^i, \dots, y_{\mu_1 \dots \mu_k}^i = \partial_{\mu_k} y_{\mu_1 \dots \mu_{k-1}}^i$, то 1-формы

$$\begin{aligned}\omega^i &= dy^i - y_\mu^i dx^\mu = 0 \\ \omega_\mu^i &= dy_\mu^i - y_{\mu\nu}^i dx^\nu = 0 \\ &\dots\dots\dots \\ \omega_{\mu_1 \dots \mu_{k-1}}^i &= dy_{\mu_1 \dots \mu_{k-1}}^i - y_{\mu_1 \dots \mu_k}^i dx^{\mu_k} = 0\end{aligned}$$

определяют на многообразии струй $J^k(M, E)$ распределение $\mathcal{C}$, которое называется *распределением Картана*. Число базисных 1-форм, задающих $\mathcal{C}$, равно $N = n \binom{m+k-1}{k-1}$, поэтому размерность распределения Картана

$$\dim \mathcal{C} = \dim J^k(M, E) - N = m + n \binom{m+k-1}{m-1} = m + n \frac{(m+k-1)!}{k! (m-1)!}$$

Дифференциальное уравнение k -го порядка с геометрической точки зрения – это подмногообразие $\mathcal{F} \subset J^k(M, E)$, которое локально задается с помощью гладких функций F^a в виде

$$F^a(x^\mu, y^i, y_{\mu_1}^i, \dots, y_{\mu_1 \dots \mu_k}^i) = 0$$

В этих уравнениях координаты (x^μ, y^i) расслоенного многообразия E представляют соответственно независимые и зависимые переменные. Решение дифференциального уравнения – это сечение расслоения $E \rightarrow M$, k -струя которого удовлетворяет этому уравнению.

ПРИМЕР. Обыкновенное дифференциальное уравнение 1-го порядка $F(x, y, y') = 0$ задает в трехмерном пространстве струй $J^1(\mathbb{R}, \mathbb{R})$ с координатами (x, y, p) поверхность $\mathcal{E} : F(x, y, p) = 0$. Решение дифференциального уравнения – это интегральная кривая двумерного (к-размерности 1) распределения Картана, определяемого 1-формой $\omega = dy - p dx$. Эта кривая (на ней $\omega = 0$) лежит на поверхности $\mathcal{E}$ и без особенностей проектируется на ось x . Поскольку форма $d\omega = dx \wedge dp$ не может быть представлена в виде $\alpha \wedge \omega$, где α – некоторая 1-форма, то распределение Картана не интегрируемо по теореме Фробениуса, следовательно, его максимальное интегральное многообразие одномерно. Наглядно данную конструкцию можно представлять себе как поднятие кривой $x \rightarrow (x, y(x))$ из $\mathbb{R}^2$ в кривую $x \rightarrow (x, y(x), p(x))$, лежащую в $J^1 \approx \mathbb{R}^3$.

Пусть $f = (\chi, \phi)$ – диффеоморфизм расслоения E . Тогда на сечениях мы имеем *точечные преобразования* (point transformations)

$$x \rightarrow \tilde{x} = \chi(x, y(x)), \quad y \rightarrow \tilde{y} = \phi(x, y(x))$$

Если матрица преобразования $\left(\frac{\partial \tilde{x}}{\partial x}\right)$ невырожденная, то находя из первых уравнений x , подставляя их во вторые и тем самым исключая x , мы получим, вообще говоря, новое сечение $\tilde{s} = f \circ s : \tilde{y} = \tilde{y}(\tilde{x})$.

В частном случае *проектируемых преобразований* (fibre-preserving transformations), сохраняющих структуру расслоения $E \rightarrow M$,

$$x \rightarrow \tilde{x} = \chi(x), \quad y \rightarrow \tilde{y} = \phi(x, y(x))$$

преобразованное сечение имеет вид $\tilde{y} = \phi(\chi^{-1}(\tilde{x}), y(\phi^{-1}(\tilde{x})))$.

Каждое преобразование $f = (\chi, \phi)$ порождает преобразование всех производных $y_{\mu_1, \dots, \mu_s}(x)$. Таким образом в многообразии $J^k(E, M)$ возникают дифференциальные продолжения s -го порядка f^s ($s = 1, \dots, k$), которые преобразуют производные функций $y = y(x)$ в производные функций $\tilde{y} = \tilde{y}(\tilde{x})$. Они однозначно определяются тем, что оставляют инвариантным распределение Картана.

ПРИМЕР.* Рассмотрим обыкновенное дифференциальное уравнение (ОДУ) 3-го порядка $y''' = F(x, y, y', y'')$. При точечных преобразованиях $x \rightarrow \tilde{x} = \chi(x, y)$, $y \rightarrow \tilde{y} = \phi(x, y)$ производные функции $y(x)$ преобразуются следующим образом:

$$\begin{aligned} y' &= \frac{dy}{dx} \rightarrow \frac{d\tilde{y}}{d\tilde{x}} = \frac{D\phi}{D\chi} \\ y'' &= \frac{d^2y}{dx^2} \rightarrow \frac{d^2\tilde{y}}{d\tilde{x}^2} = \frac{D}{D\chi} \left(\frac{D\phi}{D\chi} \right) \\ y''' &= \frac{d^3y}{dx^3} \rightarrow \frac{d^3\tilde{y}}{d\tilde{x}^3} = \frac{D}{D\chi} \left(\frac{D}{D\chi} \left(\frac{D\phi}{D\chi} \right) \right) \end{aligned}$$

где введена полная производная $D = \partial_x + y' \partial_y + y'' \partial_{y'} + y''' \partial_{y''}$. В пространстве 2-струй J^2 введем координаты (x, y, p, q) . Решение данного ОДУ третьего порядка зависит от трех постоянных, которые определяются выбором значений $y(x_0)$, $y'(x_0)$ и $y''(x_0)$ в некоторой точке x_0 . Таким образом, через каждую точку четырехмерного пространства J^2 проходит точно одна кривая $(x, y(x), y'(x), y''(x))$ – лифт кривой $(x, y(x))$ (решения ОДУ). В кокасательном пространстве T^*J^2 выберем базис 1-форм

$$\begin{aligned} \omega^1 &= dy - p dx \\ \omega^2 &= dp - q dx \\ \omega^3 &= dq - F(x, y, p, q) dx \\ \omega^4 &= dx \end{aligned}$$

Если $y = f(x)$ – решение ОДУ, то на кривой $(t, f(t), f'(t), f''(t))$ формы $\omega^1 = \omega^2 = \omega^3 = 0$, а форма $\omega^4 = dt$. При точечных преобразованиях базис 1-форм переходит в

$$\begin{aligned} \omega^1 &\rightarrow \tilde{\omega}^1 = u_1 \omega^1 \\ \omega^2 &\rightarrow \tilde{\omega}^2 = u_2 \omega^1 + u_3 \omega^2 \\ \omega^3 &\rightarrow \tilde{\omega}^3 = u_4 \omega^1 + u_5 \omega^2 + u_6 \omega^3 \\ \omega^4 &\rightarrow \tilde{\omega}^4 = u_8 \omega^1 + u_7 \omega^4 \end{aligned}$$

*Пример взят из статьи P. Nurowski, M. Godlinski “Geometry of third-order ODEs” (2009), arXiv: 0902.4129

где $u_1, \dots, u_8$ – функции на J^2 , зависящие от χ и ϕ , например, $u_1 = \phi_y$. Эти преобразования являются невырожденными, $u_1 u_3 u_5 u_7 \neq 0$, и задают G -структуру на пространстве струй J^2 , где

$$G = \begin{pmatrix} u_1 & 0 & 0 & 0 \\ u_2 & u_3 & 0 & 0 \\ u_4 & u_5 & u_6 & 0 \\ u_8 & 0 & 0 & u_7 \end{pmatrix}$$

7 Дополнительные структуры на многообразиях

В этом разделе будут перечислены некоторые дополнительные геометрические структуры, которые находят приложения в физике. Список их, безусловно, далеко не полон. Каждая из этих структур подробно описана в доступной учебной литературе, примерный список которой приведен в конце обзора.

7.1 Риманова метрика в теории гравитации

Пусть задано многообразие M , и в каждой его точке определена билинейная форма $g(v, w) : T_x M \times T_x M \rightarrow \mathbb{R}$, то есть определено скалярное произведение с помощью симметричного невырожденного тензора $g_{\mu\nu}(x)$. Тогда M называется *римановым многообразием*, а $g_{\mu\nu}(x)$ – римановой метрикой. Она определяет расстояние ds между двумя бесконечно-близкими точками с координатами x^μ и $x^\mu + dx^\mu$ по формуле

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu$$

Компоненты метрики $g_{\mu\nu}$ заданы относительно голономного (координатного) базиса $\{\partial_\mu\} = \frac{\partial}{\partial x^\mu}$. Часто используют общий неголономный базис касательного пространства

$$e_a(x) = h_a^\mu(x) \partial_\mu$$

где невырожденные матрицы (компоненты припадающих форм) $h_a^\mu(x)$ называются тетрадными функциями. Неголономный базис кокасательного пространства обозначают

$$\omega^a(x) = \tilde{h}_\mu^a(x) dx^\mu$$

где $h_a^\mu \tilde{h}_\nu^a = \delta_\nu^\mu$ и $h_a^\mu \tilde{h}_\mu^b = \delta_a^b$. Тетрады связаны с метрикой соотношениями

$$g_{\mu\nu} = \eta_{ab} h_\mu^a h_\nu^b, \quad \eta^{ab} = g^{\mu\nu} h_\mu^a h_\nu^b$$

где матрица $g^{\mu\nu}$ обратна матрице $g_{\mu\nu}$. Для собственно римановых пространств с положительно определенной метрикой обычно используют ортонормированный базис, в котором $\eta_{ab} = \delta_{ab}$. Для пространства-времени с псевдоевклидовой метрикой $\eta_{ab} = \text{diag}(1, -1, -1, -1)$ (метрика Минковского).

В точке x_0 метрика $g_{\mu\nu}(x_0)$ всегда может быть приведена к диагональному виду с помощью диффеоморфизма многообразия. В окрестности точки этого сделать в общем случае невозможно, препятствием является кривизна многообразия.

Формы $\omega^a(x)$ должны удовлетворять *структурным уравнениям Картана*

$$\begin{aligned} d\omega^a + \omega_b^a \wedge \omega^b &= T^a \\ d\omega_b^a + \omega_c^a \wedge \omega_b^c &= R_b^a \end{aligned}$$

где $\omega_b^a = \Gamma_{bc}^a \omega^c$ – 1-форма связности на касательном расслоении, $T^a = \frac{1}{2} T_{bc}^a \omega^b \wedge \omega^c$ – 2-форма кручения, $R_b^a = \frac{1}{2} R_{bcd}^a \omega^c \wedge \omega^d$ – 2-форма кривизны на многообразии.

Связность ω_b^a в общем случае не зависит от метрики. Если потребовать, чтобы связность была без кручения,

$$T_{\alpha\beta}^\mu = \Gamma_{[\alpha\beta]}^\mu = 0 \quad \Leftrightarrow \quad d\omega^a + \omega_b^a \wedge \omega^b = 0$$

и метрической, то есть ковариантная производная метрики обращалась бы в ноль

$$\nabla_\alpha g_{\mu\nu} = \partial_\alpha g_{\mu\nu} - \Gamma_{\mu\alpha}^\gamma g_{\gamma\nu} - \Gamma_{\nu\alpha}^\gamma g_{\mu\gamma} = 0 \quad \Leftrightarrow \quad \omega_{(ab)} = 0$$

то можно доказать, что такая связность существует и притом единственная. Она называется *связностью Леви-Чивита* и ее компоненты в голономном базисе (символы Кристоффеля) $\left\{ \begin{smallmatrix} \alpha \\ \mu\nu \end{smallmatrix} \right\}$ однозначно определяются метрикой:

$$\left\{ \begin{smallmatrix} \alpha \\ \mu\nu \end{smallmatrix} \right\} = \frac{1}{2} g^{\alpha\beta} (\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu})$$

Общая линейная связность в голономных координатах допускает следующее разложение

$$\Gamma_{\mu\nu}^\alpha = \left\{ \begin{smallmatrix} \alpha \\ \mu\nu \end{smallmatrix} \right\} + T_{\mu\nu}^\alpha + T_{\mu\nu}^\alpha + T_{\nu\mu}^\alpha + C_{\mu\nu}^\alpha + C_{\nu\mu}^\alpha - C_{\nu\mu}^\alpha$$

где тензор неметричности $C_{\mu\nu}^\alpha = -\frac{1}{2}(\partial_\alpha g_{\mu\nu} - \Gamma_{\mu\alpha}^\gamma g_{\gamma\nu} - \Gamma_{\nu\alpha}^\gamma g_{\mu\gamma})$. В этом (общем) случае геометрия многообразия определяется метрикой, ее ковариантной производной и тензором кручения.

В *общей теории относительности* используется связность Леви-Чивита, поэтому в ОТО 1-форма связности всегда удовлетворяет условиям

$$d\omega^a = -\omega_b^a \wedge \omega^b, \quad \omega_{(ab)} = 0, \quad d\omega_b^a + \omega_c^a \wedge \omega_b^c = R_b^a$$

Используя тензор кривизны R_{bcd}^a , определим тензор Риччи $R_{ab} = R_{acb}^c$ и скалярную кривизну $R = R_a^a$.

Метрика $g_{\mu\nu}(x)$, описывающая гравитационное поле, должна удовлетворять *уравнениям Эйнштейна*

$$R_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab} = T_{ab}$$

где Λ – космологическая постоянная, а T_{ab} – тензор энергии-импульса материи. Уравнения Эйнштейна инвариантны относительно диффеоморфизмов (общековариантных преобразований) многообразия M .

7.2 Финслеровы многообразия

Длина элемента дуги $ds = F(x^1, \dots, x^n; dx^1, \dots, dx^n)$ в римановой геометрии описывается функцией $F^2 = g_{\mu\nu} dx^\mu dx^\nu$. Если отказаться от требования квадратичности и рассматривать на расслоении $TM \setminus M = \{(x, y) \in TM | x \in M, y \in T_x M \setminus \{0\}\}$ произвольные гладкие функции $F(x, y)$, однородные степени 1 по y : $F(x, \lambda y) = \lambda F(x, y)$, то мы получим *финслерову метрику* на M . Часто к условию однородности добавляют еще условие положительности гессиана $\frac{1}{2} \partial^2(F^2)/(\partial y^\mu \partial y^\nu)$.

Тензор $g_{\mu\nu}(x, y) := \frac{1}{2}(F^2)_{y^\mu y^\nu} = F F_{y^\mu y^\nu} + F_{y^\mu} F_{y^\nu}$, где $F_{y^\mu} := \partial F / \partial y^\mu$, называется *фундаментальным тензором* финслерова пространства. Он однороден по y^μ и, следовательно, определен на проективном касательном расслоении PTM . Очевидно, что если $F^2(x, y) = g_{\mu\nu}(x) y^\mu y^\nu$, то мы получаем обычную риманову метрику, заданную на M :

$$g_{\mu\nu}(x, y) := \frac{1}{2} \frac{\partial^2(F^2)}{\partial y^\mu \partial y^\nu} = g_{\mu\nu}(x)$$

Фундаментальный тензор определяет скалярное произведение векторов касательного пространства $T_x M$, так как в силу условия однородности $F^2(x, y) = g_{\mu\nu}(x, y) y^\mu y^\nu$. Гиперповерхность в $T_x M$, определяемая при фиксированном x уравнением $F^2(x, y) = 1$, называется *индикатрисой*. В случае собственно римановой метрики индикатриса – это $(n-1)$ -мерная единичная сфера евклидовой геометрии.

ПРИМЕР. Рассмотрим функцию $F(x, y) = \sqrt[n]{(h_\mu^1 y^\mu) \dots (h_\mu^n y^\mu)}$, где $\omega^a = h_\mu^a(x) dx^\mu$ – базис 1-форм в кокасательном пространстве $T_x^* M$. Она определяет *метрическую функцию Бервальда-Моора* в касательном пространстве $T_x M$ по формуле $F(y) = \sqrt[n]{y^1 \dots y^n}$.

Над $TM \setminus M$ строится индуцированное расслоение $\pi^* TM$ в соответствие с диаграммой

$$\begin{array}{ccc} \pi^* TM & \longrightarrow & TM \\ \downarrow & & \downarrow \\ TM \setminus M & \xrightarrow{\pi} & M \end{array}$$

Его слоем над точкой (x, y) , $y \neq 0$ является копия n -мерного пространства $T_x M$. Заметим, что в силу однородности по y , часто удобно работать не с индуцированным расслоением над $2n$ -мерным расслоением $TM \setminus M$, а над $(2n-1)$ -мерным проективным расслоением PTM .

Индикатриса определяет выделенное сечение $\pi^* TM$ – векторное поле $v = \frac{y^\mu}{F(x, y)} \frac{\partial}{\partial x^\mu}$, на котором фундаментальный тензор $g(v, v) = 1$. Дуальным объектом к v (сечением расслоения $\pi^* T^* M$) является *гильбертова форма* $\Omega = \Omega_\mu dx^\mu$ такая, что $\Omega(v) = 1$, где $\Omega_\mu = g_{\mu\nu} v^\nu = \partial F / \partial y^\mu$. Если на M взять гладкую кривую $\sigma : [a, b] \rightarrow M$, то ее длина $L(\sigma)$ в финслеровой метрике будет определяться интегралом $L(\sigma) := \int_a^b F(\sigma, \frac{d\sigma}{dt}) dt$, который равен интегралу от гильбертовой формы:

$$L(\sigma) = \int_\sigma \Omega$$

В случае римановой метрики, когда $F = \sqrt{g_{\mu\nu}(x) y^\mu y^\nu}$, $\Omega = \frac{1}{F} g_{\mu\nu}(x) y^\mu dx^\nu$, и мы получаем обычную длину

$$L(\sigma) = \int_a^b \sqrt{g_{\mu\nu}(x(t)) \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} dt$$

На касательном расслоении к $TM \setminus M$ можно стандартным способом определить связность через расщепление

$$T(TM \setminus M) = V(TM \setminus M) \oplus H(TM \setminus M)$$

с вертикальным касательным расслоением $V(TM \setminus M) = \text{Ker } \pi^*$. Слои $V_{(x,y)}$ вертикального подрасслоения имеют базис вертикальных векторов $\partial_{y^\mu} := \partial/\partial y^\mu$, а слои горизонтального подрасслоения $H_{(x,y)}$ имеют базис $D_\mu = \partial_{x^\mu} - \Gamma_\mu^\nu \partial_{y^\nu}$, где $\Gamma_\mu^\nu(x, y)$ – коэффициенты *нелинейной* связности. Вектора $\partial/\partial y^\mu$ и D_μ образуют базис касательного пространства $T_{(x,y)}(TM \setminus M)$.

Используя финслерову функцию $F(x, y)$, можно определить каноническую нелинейную связность

$$\Gamma_\mu^\nu = \frac{1}{2} \frac{\partial G^\mu}{\partial y^\nu}$$

где коэффициенты пульверизации $G^\mu = \frac{1}{2} g^{\mu\nu} ((F^2)_{y^\nu x^\alpha} y^\alpha - (F^2)_{x^\nu})$.

Несмотря на то, что на π^*TM не существует канонической *линейной* связности, часто удобно ввести такую связность, исходя из различных дополнительных требований. Например, известная *связность Чженя* $\omega_b^a(x, y)$ определяется условиями, аналогичными условиям для связности Леви-Чивита в римановой геометрии:

- 1) отсутствие кручения: $d\omega^a + \omega_b^a \wedge \omega^b = 0$, что в координатном базисе эквивалентно отсутствию dy^α в разложении $\omega_\nu^\mu(x, y) = \Gamma_{\nu\alpha}^\mu(x, y) dx^\alpha$ и симметричности $\Gamma_{\nu\alpha}^\mu = \Gamma_{\alpha\nu}^\mu$;
- 2) $dg_{\mu\nu} - \omega_\mu^\alpha g_{\alpha\nu} - \omega_\nu^\alpha g_{\mu\alpha} = 2C_{\mu\nu\alpha} dy^\alpha$, где введен тензор Картана $C_{\mu\nu\alpha}(x, y) = \frac{1}{4} (F^2)_{y^\mu y^\nu y^\alpha}$ и $\delta y^\alpha = dy^\alpha + \Gamma_\nu^\alpha dx^\nu$.

7.3 Симплектическая и пуассонова структуры

Симплектическая структура – это замкнутая невырожденная дифференциальная 2-форма Ω , заданная на четномерном многообразии, которое в этом случае называется симплектическим многообразием.

ПРИМЕР. Кокасательное расслоение T^*M с координатами $(x^\mu, \dot{x}_\mu)$ над многообразием M является симплектическим многообразием с канонической 2-формой

$$\Omega = d\dot{x}_\mu \wedge dx^\mu$$

В механике M называется конфигурационным пространством, а T^*M – фазовым пространством. Локальные координаты x^μ на M называются каноническими координатами, а координаты $\dot{x}_\mu$ в кокасательном пространстве относительно голономного базиса dx^μ – обобщенными импульсами. Симплектические многообразия описывают геометрию автономной механики консервативных систем.

Диффеоморфизм симплектических многообразий, переводящий симплектическую структуру одного в симплектическую структуру другого, называется *симплектоморфизмом* (каноническим преобразованием).

Условие замкнутости ($d\Omega = 0$) в определении Ω связывает кососкалярные произведения $(\dot{x}, \dot{y}) = \Omega_{\mu\nu} \dot{x}^\mu \dot{y}^\nu$ в касательных пространствах к соседним точкам таким образом, что локальная геометрия симплектических многообразий оказывается универсальной, а именно стандартной геометрией евклидова пространства. Все симплектические многообразия одинаковой размерности локально симплектоморфны (теорема Дарбу). Другими словами, на

$(2n)$ -мерном многообразии всегда существуют такие координаты (q^i, p_i) , $q^i = x^i, p_i = x_{n+i}$, $i = 1, \dots, n$, что симплектическая структура в окрестности любой точки имеет вид

$$\Omega = dp_i \wedge dq^i$$

Таким образом, в симплектической геометрии не существует локальных инвариантов, как это имеет место в геометрии римановой, где локальным инвариантом является тензор кривизны, который ограничивает группу изометрий и приводит к существованию бесконечномерного множества неэквивалентных римановых метрик. В симплектической геометрии вместо этого мы имеем бесконечномерную группу симплектоморфизмов и дискретный набор неэквивалентных *глобальных* симплектических структур.

В пространстве гладких функций на многообразии можно ввести структуру алгебры Ли – скобку Пуассона. В общем случае **скобка Пуассона** определяется как $\mathbb{R}$ -билинейное отображение на $C^\infty(M)$ такое, что для любых функций $f, g, h \in C^\infty(M)$ выполняются условия

- 1) $\{f, g\} = -\{g, f\}$ (антисимметричность)
- 2) $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$ (тождество Якоби)
- 3) $\{f, gh\} = \{f, g\}h + \{f, h\}g$ (правило Лейбница).

Локально скобка Пуассона задается антисимметричным тензором 2-го ранга $w^{\mu\nu}(x)$, удовлетворяющим условию

$$w^{\mu\alpha}\partial_\mu w^{\beta\gamma} + w^{\mu\beta}\partial_\mu w^{\gamma\alpha} + w^{\mu\gamma}\partial_\mu w^{\alpha\beta} = 0$$

Бивекторное поле $w = \frac{1}{2}w^{\mu\nu}\partial_\mu \wedge \partial_\nu$ называется *пуассоновой структурой*. Скобка Пуассона имеет вид

$$\{f, g\} = w^{\mu\nu}(x)\partial_\mu f \partial_\nu g$$

На многообразии тензор Пуассона определяет скалярное произведение ковекторов, например, градиентов функций.

Если тензор Пуассона невырожден ($\det w^{\mu\nu} \neq 0$) в каждой точке многообразия, то обратная матрица $\|w_{\mu\nu}\| = \|w^{\mu\nu}\|^{-1}$ будет определять симплектическую структуру, так как тождество Якоби для w эквивалентно условию замкнутости $d\Omega = 0$, где теперь $\Omega_{\mu\nu} = w_{\mu\nu}$. Таким образом, симплектические многообразия – это частный случай пуассоновых многообразий, когда тензор Пуассона невырожден.

ПРИМЕР. На симплектическом многообразии с координатами Дарбу (q^i, p_i) , когда $\Omega = dp_i \wedge dq^i$, скобка Пуассона имеет вид

$$\{f, g\} = \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i}$$

Пусть $f(x)$ – некоторая функция на пуассоновом многообразии, тогда векторное поле

$$v = w^{\mu\nu} \frac{\partial f}{\partial x^\mu} \partial_\nu$$

называется гамильтоновым векторным полем, ассоциированным с $f(x)$.

ПРИМЕР. Пусть $H(x)$ – гамильтониан механической системы на многообразии M . Если на кокасательном расслоении T^*M ввести канонические координаты Дарбу, то траектории гамильтонова векторного поля, ассоциированного с $H(x)$,

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}$$

определяют эволюцию системы.

7.4 Комплексные многообразия

Пусть M – многообразие размерности $2n$. Определим изоморфизм касательных расслоений $J : TM \rightarrow TM$ такой, что в каждой точке $x \in M$ $J_x^2 = -I$, где I – тождественный изоморфизм, действующий на TM . Тогда J называется **почти комплексной структурой** на многообразии M . Почти комплексная структура вводит в касательном пространстве $T_x M$, являющемся $2n$ -мерным действительным линейным пространством, структуру n -мерного комплексного пространства, определяя правило умножения вектора v на комплексные числа по формуле

$$(a + ib)v = av + bJv, \quad a, b \in \mathbb{R}, i = \sqrt{-1}$$

Почти комплексная структура существует на любом четномерном ориентируемом многообразии.

ПРИМЕР. Определим в $M = \mathbb{R}^{2n}$ с координатами (x^i, y^i) , $i = 1, \dots, n$ почти комплексную структуру J по правилу

$$J(x^1, \dots, x^n, y^1, \dots, y^n) = (-y^1, \dots, -y^n, x^1, \dots, x^n)$$

Тогда отображение $J : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ будет эквивалентно умножению на i в $\mathbb{C}^n$ с координатами $z^i = x^i + iy^i$. Тензор почти комплексной структуры J_ν^μ в этом примере имеет вид

$$J_\nu^\mu = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

где $I = \delta_j^i$ – единичная матрица $n \times n$; очевидно, что $J_\mu^\alpha J_\beta^\mu = -\delta_\beta^\alpha$. Такая структура называется стандартной комплексной структурой на $\mathbb{R}^{2n}$.

Выберем в качестве базиса касательного пространства $T_x M$ координатный базис $\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}, \frac{\partial}{\partial y^1}, \dots, \frac{\partial}{\partial y^n}\}$, тогда над полем $\mathbb{C}$ мы получим комплексифицированное касательное пространство $T_x^{\mathbb{C}} M$ с базисом

$$\frac{\partial}{\partial z^i} = \frac{1}{2} \left(\frac{\partial}{\partial x^i} - i \frac{\partial}{\partial y^i} \right), \quad \frac{\partial}{\partial \bar{z}^i} = \frac{1}{2} \left(\frac{\partial}{\partial x^i} + i \frac{\partial}{\partial y^i} \right)$$

Если коммутатор любых двух комплексных касательных векторов вида $v = v^i \frac{\partial}{\partial z^i}$ в любой точке многообразия есть тоже комплексный вектор, то почти комплексная структура называется интегрируемой. В общем случае произвольного почти комплексного многообразия коммутатор может давать комплексно-сопряженные векторы вида $w = \bar{w}^i \frac{\partial}{\partial \bar{z}^i}$. Условие интегрируемости эквивалентно обращению в ноль тензора кручения почти комплексной структуры (тензора Нейенхайса)

$$N_{\beta\gamma}^\alpha = J_\beta^\mu (\partial_\mu J_\gamma^\alpha - \partial_\gamma J_\mu^\alpha) - J_\gamma^\mu (\partial_\mu J_\beta^\alpha - \partial_\beta J_\mu^\alpha)$$

Для интегрируемости почти комплексной структуры необходимо и достаточно, чтобы многообразие M было комплексным многообразием, то есть на нем существовал атлас с голоморфными функциями перехода (теорема Ньюлендера-Ниренберга).

Для любой комплекснозначной p -формы ω возможно разложение

$$\omega = \sum \omega_{i_1 \dots i_r \bar{j}_1 \dots \bar{j}_s} dz^{i_1} \wedge \dots \wedge dz^{i_r} \wedge d\bar{z}^{j_1} \wedge \dots \wedge d\bar{z}^{j_s}$$

Такие формы $\omega_{r,s} \in \Omega^{r,s}(M)$ называются формами (r, s) типа. На комплексном многообразии внешний дифференциал раскладывается в сумму $d = \partial + \bar{\partial}$, где $\partial : \Omega^{r,s} \rightarrow \Omega^{r+1,s}$, $\bar{\partial} : \Omega^{r,s} \rightarrow \Omega^{r,s+1}$. Эти операторы удовлетворяют соотношениям

$$\partial^2 = \partial\bar{\partial} + \bar{\partial}\partial = \bar{\partial}^2 = 0$$

Соответствующие им группы когомологий называются когомологиями Дольбо:

$$H_{\partial}^{r,s} = \frac{\text{Ker } \partial}{\text{Im } \partial}, \quad H_{\bar{\partial}}^{r,s} = \frac{\text{Ker } \bar{\partial}}{\text{Im } \bar{\partial}}$$

На комплексном многообразии можно ввести *эрмитову метрику* $h(v, w) \in \mathbb{C}$, инвариантную относительно комплексной структуры: $h(v, w) = h(Jv, Jw)$. В комплексных координатах $(z^1, \dots, z^n)$ эрмитова метрика $h = h_{i\bar{k}} dz^i \otimes d\bar{z}^{\bar{k}}$ задается тензором $h_{i\bar{k}}$, который определяет эрмитово скалярное произведение касательных векторов:

$$h(v, w) = h_{i\bar{k}} v^i \bar{w}^{\bar{k}}$$

Матрица тензора $h_{i\bar{k}}$ является эрмитовой: $h_{i\bar{k}} = \overline{h_{k\bar{i}}}$. Ее вещественная часть симметрична и определяет риманову метрику $g = \frac{1}{2}(h + \bar{h})$ на касательном пространстве, а мнимая часть антисимметрична и определяет $(1, 1)$ -форму

$$\omega = \frac{i}{2} h_{i\bar{k}} dz^i \wedge d\bar{z}^{\bar{k}}$$

Риманова метрика g и форма ω инвариантны относительно комплексной структуры J и связаны с ней соотношениями

$$g(v, w) = \omega(v, Jw), \quad \omega(v, w) = g(Jv, w)$$

Если форма ω замкнута — $d\omega = 0$, то комплексное многообразие называется *кэлеровым*, а форма ω — кэлеровой. Условие $d\omega = 0$ эквивалентно ковариантному постоянству комплексной структуры $\nabla J = 0$ относительно связности Леви-Чивита, построенной по римановой метрике $g = \frac{1}{2}(h + \bar{h})$, то есть параллельный перенос индуцирует комплексно-линейное отображение на касательных пространствах. На кэлеровом многообразии форма ω определяет также симплектическую структуру, поэтому каждое кэлерово многообразие является симплектическим. Обратное неверно, существуют симплектические многообразия не являющиеся кэлеровыми.

ПРИМЕР. На комплексном проективном пространстве $\mathbb{C}P^n$ с однородными координатами $[z^0 : \dots : z^n]$ определим эрмитову метрику *Фубини-Штуды* $h = h_{i\bar{k}} dw^i \otimes d\bar{w}^{\bar{k}}$ в координатах аффинной карты $w^i = z^i/z^0$ соотношением

$$h_{i\bar{k}} = \frac{(1 + |w|^2)\delta_{i\bar{k}} - w_i \bar{w}_{\bar{k}}}{(1 + |w|^2)^2}$$

Поскольку форма $\omega = \frac{i}{2} h_{i\bar{k}} dw^i \wedge d\bar{w}^{\bar{k}}$ замкнута, то $\mathbb{C}P^n$ является кэлеровым. В частности, для комплексной проективной прямой $\mathbb{C}P^1 \approx S^2$ получаем обычную сферическую метрику $g = (1 + |w|^2)^{-2} dw d\bar{w} = d\theta^2 + \sin^2 \theta d\varphi^2$ и

$$\omega = \frac{i}{2} \frac{dw \wedge d\bar{w}}{(1 + w\bar{w})^2}$$

Благодарности

Автор благодарен Д. Г. Павлову за приглашение прочитать лекции на Осенней Школе «Финслерова геометрия, гиперкомплексные системы и теория относительности» и за финансовую поддержку.

Список литературы

- [1] Атья М. Геометрия полей Янга-Миллса // Геометрия и физика узлов. М.: Мир, 1995
- [2] Дубровин Б. А., Новиков С. П., Фоменко А. Т. Современная геометрия: Методы и приложения. М.: Наука, 1986
- [3] Кобаяси Ш., Номидзу К. Основы дифференциальной геометрии. М.: Наука, 1981
- [4] Кобаяси Ш. Группы преобразований в дифференциальной геометрии. М.: Наука, 1986
- [5] Монастырский М. И. Топология калибровочных полей и конденсированных сред. М.: ПАИМС, 1995
- [6] Новиков С. П., Тайманов И. А. Современные геометрические структуры и поля. М.: МЦНМО, 2005
- [7] Сарданашвили Г. А. Современные методы теории поля. т. 1: Геометрия и классические поля. М.: УРСС, 1996
- [8] Шапиро И. С., Ольшанецкий М. А. Лекции по топологии для физиков. Ижевск: НИЦ РХД, 2001
- [9] Шапуков Б. Н. Задачи по группам Ли и их приложениям. Москва: НИЦ РХД, 2002
- [10] Bao D., Chern S.-S., Shen Z. An introduction to Riemann-Finsler geometry. New-York: Springer, 2000
- [11] Eguchi T., Gilkey P. B., Hanson A. J. Gravitation, gauge theories and differential geometry // Physics Reports, v. 66, 1980, pp. 213-393
- [12] Nash C., Sen S. Topology and geometry for physicists. London: Academic Press, 1983
- [13] Madore J. An introduction to noncommutative differential geometry and its physical applications. Cambridge: CUP, 1995
- [14] Marathe K. B., Martucci G. The mathematical foundations of gauge theories. Amsterdam: North-Holland, 1992

Pseudo-Finsler Geometry. Symmetries and transformation groups

Vladimir Balan

University Politehnica of Bucharest, Faculty of Applied Sciences, Dept. Mathematics-Informatics I, Splaiul

Independentei 313, RO-060042 Bucharest, Romania.

E-mail: vladimir.balan@upb.ro

1 Pseudo-Finsler Geometry

1.1 Finsler structures. Finsler metric tensor field

Definitions. a) We call *Finsler structure* a couple (M, F) , where M is a real n -dimensional C^∞ differentiable manifold and $F : TM \rightarrow \mathbb{R}_+$ is a mapping (called *Finsler fundamental function* or *Finsler norm*), which obeys the following properties:

1. F is C^∞ on $\widetilde{TM} = TM \setminus \{0\} \equiv \bigcup_{x \in M} \{(x, y) \mid y \in T_x M, y \neq 0_x\}$ (F is smooth on the tangent space without the image of the null section);
2. $F(x, \lambda y) = \lambda F(x, y)$, $\forall \lambda \in [0, \infty)$ (F is positive homogeneous of degree one in y , i.e., on the fibres of the tangent bundle (TM, π, M));
3. for all $(x, y) \in TM$, the functions

$$g_{ij} = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j} \quad (1)$$

provide a symmetric positive definite matrix. They are the components of the *metric Finsler tensor field*

$$g = g_{ij}(x, y) dx^i \otimes dx^j.$$

b) We call *the indicatrix* of the Finsler structure (M, F) at the point $x \in M$, the set of unit tangent vectors $I_x = \{y \in T_x M \mid F(x, y) = 1\}$.

Remarks. a) If 3) is replaced with the (weaker) condition that g should be non-degenerate and of constant signature (not necessarily positive-definite and hence F not necessarily non-negative), then the structure (M, F) is called *pseudo-Finsler structure*.

b) If 2) is supplemented with the *reversibility condition*, i.e., $F(x, y) = F(x, -y)$, $\forall (x, y) \in TM$, then F becomes a norm, which satisfies $F(x, \lambda y) = |\lambda| F(x, y)$, $\forall \lambda \in \mathbb{R}$. This is the case of (e.g.) the Riemannian metric $F = \alpha$.

c) If F is no longer non-negative, or not defined on the whole space TM (being defined only on certain distributions of the tangent bundle), then there appear examples of new (generalized) Finsler functions, e.g., the Kropina metric $F = \alpha^2/\beta$, or the Shimada pseudo-Finsler metric $F(x, y) = \sqrt[n]{a_{i_1 \dots i_m}(x) y^{i_1} \dots y^{i_m}}$, having as particular case the Berwald-Moor metric $F(y) = \sqrt[n]{y^1 \dots y^n}$.

1.2 The non-linear connection of a Finsler structure

Consider the tangent bundle (TM, π, M) . Then a basic tool which allows to introduce a structure of vector bundle on $(TTM, d\pi, TM)$ with structural group $GL_n(\mathbb{R}) \times GL_n(\mathbb{R})$ is the *non-linear connection* $N = \{N_i^a\}$, which provides a *local* adapted basis for the module of sections of the subbundles $(HTM, d\pi, TM)$ and $(VTM, d\pi, TM)$ of the Whitney splitting $(TTM = HTM \oplus VTM, d\pi, TM)$, namely

$$\left\{ \delta_i = \frac{\delta}{\delta x^i} = \frac{\partial}{\partial x^i} - N_i^a(x, y) \frac{\delta}{\delta y^i} \right\} \subset \Gamma(HTM), \quad \left\{ \dot{\partial}_i = \frac{\partial}{\partial y^i} \right\} \subset \Gamma(VTM), \quad (2)$$

where $VTM = \text{Ker } d\pi$ and where we have denoted by Γ the sections of the corresponding bundle. There exists a canonical nonlinear connection, provided by the Finsler fundamental function F , given by (Kern’s formula):

$$N_i^a = \frac{\partial G^a}{\partial y^i}, \quad (3)$$

where g^{ij} are the components of the matrix dual (inverse) to g_{ij} , i.e., $g^{is}g_{sj} = \delta_j^i$ and

$$G^a = \frac{1}{2} g^{as} \left(\frac{\partial^2 L}{\partial y^s \partial x^j} y^j + \frac{\partial L}{\partial x^s} \right). \quad (4)$$

1.3 Linear connections and covariant derivation laws

The components $\nabla = \{\tilde{F}_{jk}^i; \mathbb{C}_{jk}^i\}$ of a linear connection which preserves the h - and v -distributions (called *linear d-connection*, [31]), locally described by (2), computed with respect to the local adapted basis of fields determined by a fixed non-linear connection N

$$\{\delta_1, \dots, \delta_n; \dot{\partial}_1, \dots, \dot{\partial}_n\}$$

are described by the relations

$$\nabla_{\delta_k} \delta_j = \tilde{F}_{jk}^i \delta_i; \quad \nabla_{\dot{\partial}_c} \dot{\partial}_b = \tilde{C}_{bc}^a \dot{\partial}_a. \quad (5)$$

These components are called the h -components $\{\tilde{F}_{jk}^i\}$ and the v -components $\{\tilde{C}_{bc}^a\}$, respectively

The non-trivial components of the torsion $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$ of ∇ are Finsler tensor fields, provided by the formulas

$$\begin{aligned} T(\delta_k, \delta_j) &= T_{jk}^i \delta_k + R_{jk}^a \dot{\partial}_a \\ T(\delta_k, \dot{\partial}_b) &= \tilde{C}_{bk}^a \delta_a + P_{bk}^a \dot{\partial}_a \\ T(\dot{\partial}_c, \dot{\partial}_b) &= S_{bc}^a \dot{\partial}_a, \end{aligned} \quad (6)$$

These non-trivial torsion fields $T \equiv \{T_{jk}^i, R_{jk}^a, P_{bk}^a, \tilde{C}_{ja}^i, S_{bc}^a\}$ described in (6), have the explicit expressions in terms of the connection coefficients $\nabla = \{\tilde{F}_{jk}^i; \tilde{C}_{jk}^i\}$ given by (5), as follows

$$T_{jk}^i = \tilde{F}_{jk}^i - \tilde{F}_{kj}^i, \quad R_{jk}^a = \delta_k N_j^a - \delta_j N_k^a, \quad P_{bk}^a = \tilde{F}_{bk}^a - \dot{\partial}_b N_k^a, \quad \tilde{C}_{ja}^i = S_{bc}^a = \tilde{C}_{bc}^a - \tilde{C}_{cb}^a. \quad (7)$$

We note that *the Cartan tensor field* which is defined in (11) is totally symmetric in all its indices and that (M, F) is a Riemannian structure (i.e., g_{ij} depend on x only) if $C_{ijk} \equiv 0$.

As well, for a Finsler linear d -connection ∇ , one can define the associated h - and the v -covariant derivatives, as follows

$$\left\{ \begin{array}{l} T_{j_1 \dots j_q | k}^{i_1 \dots i_p} = \frac{\delta T_{j_1 \dots j_q}^{i_1 \dots i_p}}{\delta x^k} + \sum_{s=\overline{1,p}} T_{j_1 \dots j_q}^{i_1 \dots t_s \dots i_p} \tilde{F}_{t_s k}^{i_s} - \sum_{s=\overline{1,q}} T_{j_1 \dots t_s \dots j_q}^{i_1 \dots i_p} \tilde{F}_{j_s k}^{t_s} \\ T_{j_1 \dots j_q}^{i_1 \dots i_p} \Big|_k = \frac{\partial T_{j_1 \dots j_q}^{i_1 \dots i_p}}{\partial y^k} + \sum_{s=\overline{1,p}} T_{j_1 \dots j_q}^{i_1 \dots t_s \dots i_p} \tilde{C}_{t_s k}^{i_s} - \sum_{s=\overline{1,q}} T_{j_1 \dots t_s \dots j_q}^{i_1 \dots i_p} \tilde{C}_{j_s k}^{t_s}, \end{array} \right.$$

for any Finsler tensor field of type (p, q) . The connection ∇ is said to be *h -metrical* if $g_{ij|k} = 0$ and *v -metrical* if $g_{ij}^{\cdot k} = 0$. There exist several well-known linear d -connections in Finsler Geometry, used in applications. None of these connections has the two basic qualities of the Levi-Civita connection from Riemannian Geometry (metricity and symmetry). Generally, the Levi-Civita connection is *not* a linear d -connection: the reasons for which such a connection ceases to exist on the tangent space TM , is that the natural extra requirement - that of preserving the h - and v -distributions locally described by (2), is fulfilled. Below is a table which contains four remarkable linear connections which obey the extra property of being d -connections, but fail to be completely metric or symmetric (e.g., [31], [49–51], [25]).

Linear d -connection	Coefficients	Metricity		Symmetry				
		$h-$	$v-$	$hh-h$	$hh-v$	$hv-h$	$hv-v$	$vv-v$
		$g_{ij k}$	$g_{ij}^{\cdot k}$	T_{jk}^i	R_{jk}^a	C_{ja}^i	P_{bk}^a	S_{bc}^a
<i>Cartan</i>	$\frac{C}{\nabla} = \{F_{jk}^i, C_{jk}^i\}$	0	0	0	#	#	#	0
<i>Chern-Rund</i>	$\frac{CR}{\nabla} = \{F_{jk}^i, 0\}$	#	#	0	#	0	#	0
<i>Berwald</i>	$\frac{B}{\nabla} = \{G_{jk}^i, 0\}$	0	#	0	#	0	0	0
<i>Hashiguchi</i>	$\frac{H}{\nabla} = \{G_{jk}^i, C_{jk}^i\}$	#	0	0	#	#	0	0

The basic formulas which provide the components of the four connections from above are:

$$G_{jk}^i = \dot{\partial}_j N_k^i \quad (8)$$

$$F_{jk}^i = \frac{1}{2} g^{is} (\delta_j g_{sk} + \delta_k g_{sj} - \delta_s g_{jk}) \quad (9)$$

$$C_{jk}^i = \frac{1}{2} g^{is} (\dot{\partial}_j g_{sk} + \dot{\partial}_k g_{sj} - \dot{\partial}_s g_{jk}) = g^{is} C_{isk}, \quad (10)$$

where

$$C_{isk} = \frac{1}{2} \frac{\partial g_{is}}{\partial y^k} = \frac{1}{4} \frac{\partial^3 F^2}{\partial y^i \partial y^s \partial y^k}. \quad (11)$$

1.4 Invariants of a Finsler structure

The invariants of Riemann type of a Finsler structure (M, F) are the following ([9])

$$\begin{cases} g = g_{ij}dx_i \otimes dx_j, & g_{ij} = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j} \\ \eta = \eta_i dx^i, & \eta_i = \partial F / \partial y^i. \end{cases}$$

We note that these are 0-homogeneous Finsler tensor fields of type $(0,2)$, $(0,3)$ and $(0,1)$, respectively. The invariants of non-Riemannian type are

$$\begin{cases} A_{ijk}dx^i \otimes dx^j \otimes dx^k, & \text{where } A_{ijk} = \frac{F}{4} \frac{\partial^3 F^2}{\partial y^i \partial y^j \partial y^k} = \frac{F}{2} \frac{\partial g_{ij}}{\partial y^k} \\ A_i dx^i, & \text{where } A_i = g^{jk} A_{ijk} \\ \tau = \frac{\sqrt{\det(g_{ij})_{i,j=\overline{1,n}}}}{\sigma(x)}, & \text{where } \sigma(x) = \frac{\text{Vol}(S^n(1))}{\text{Vol}(\{y \in \mathbb{R}^n \mid F(x, y^i \frac{\partial}{\partial x^i}) \leq 1\})}. \end{cases}$$

We note that these are 0-homogeneous Finsler tensor fields of type $(0,1)$ and $(0,0)$, respectively.

1.5 Examples of Finsler and pseudo-Finsler structures

I. The Finsler metric function $F_R(x, y) = \sqrt{g_{ij}(x)y^i y^j}$, where $\{g_{ij}(x)\}$ are the coefficients of a Riemannian metric on M , practically represents the Banach norm associated to the Riemannian structure $g = g_{ij}(x)dx^i \otimes dx^j$ of (M, g) . In particular, for $g_{ij} = \delta_{ij}$, one gets for the Finsler metric $F(y) = \sqrt{(y^1)^2 + \dots + (y^n)^2}$ of the Euclidean structure (M, δ_{ij}) .

Similarly, the pseudo-Finsler metric function $F_R(x, y) = \sqrt{g_{ij}(x)y^i y^j}$, where $\{g_{ij}(x)\}$ are the coefficients of a pseudo-Riemannian metric, is the pseudo-Riemannian norm attached to the pseudo-Riemannian structure $g = g_{ij}(x)dx^i \otimes dx^j$ of (M, g) . In particular, for $g_{ij} = \varepsilon_i \delta_{ij}$, where $\varepsilon_1 = -1, \varepsilon_2 = \dots = \varepsilon_n = 1$, one gets for the Minkowski pseudo-Finslerian norm $F(y) = \sqrt{-(y^1)^2 + (y^2)^2 + \dots + (y^n)^2}$ attached to the pseudo-Euclidean Minkowski structure $(M, \varepsilon_i \delta_{ij})$.

II. The Randers metric $F = \alpha + \beta$, with

$$\alpha = \sqrt{a_{ij}(x)y^i y^j}, \quad \beta = b_k(x)y^k, \quad (12)$$

where $a = a_{ij}(x)dx^i \otimes dx^j$ is a Riemannian metric on M and $b_k(x)dx^k$ is a 1-form on M , has the property that for $\beta = 0$, i.e., for $b_k = 0, \forall k \in \overline{1, n}$ reduces to the Finsler Riemannian structure $(M, a \equiv \{a_{ij}\})$.

The Randers spaces were first introduced by Randers in 1941, and further studied by Roman Ingarden (who gave their actual name in 1957), Ralf G. Beil, Aurel Bejancu, Ioan Buc'ataru, Vasile Sabău, etc. The Randers spaces have been shown to provide an alternative model for unified gravitation and electromagnetism.

Homework. Denote $\|b\|_a = \|b_k(x)dx^k\|_a = a^{ij}(x)b_i(x)b_j(x)$, and check that for $\|b\|_a < 1$ one gets a Finsler structure with compact indicatrix $I_x, \forall x \in M$ which encloses a convex body, while for $\|b\|_a > 1$, F defines a pseudo-Finsler structure with non-bounded indicatrix.

III. Particular cases of Randers metrics:

a) The *generalized Funk metric* given by

$$F(x, y) = \frac{\sqrt{(\|x\|^2 \langle a, y \rangle - 2 \langle a, x \rangle \langle x, y \rangle)^2 + \|y\|^2 (1 - \|a\|^2 \|x\|^4)}}{1 - \|a\|^2 \|x\|^4} - \frac{\|x\|^2 \langle a, y \rangle - 2 \langle a, x \rangle \langle x, y \rangle}{1 - \|a\|^2 \|x\|^4},$$

where $x \in D = B(0, \|a\|^{-1/2}) \subset \mathbb{R}^n, y \in T_x D, a \in \mathbb{R}^n$.

Homework. Show that (D, F) is a Randers space of scalar curvature

$$K = 3\|F\|^{-1} \langle a, y \rangle + 3 \langle a, x \rangle^2 - 2\|a\|^2 \|x\|^2,$$

and that it has isotropic S -curvature, given by $S = (n+1) \underbrace{\langle a, x \rangle}_{c(x)} F$.

b) The Finsler space $(M = B^n(1/\sqrt{-\mu}) \subset \mathbb{R}^n, F)$ with $\mu < 0$ and

$$F_\mu(x, y) = \frac{(\|y\|^2 + \mu(\|x\|^2 \|y\|^2 - \langle x, y \rangle^2))^{1/2}}{1 + \mu \|x\|^2}, \quad \forall y \in T_x M.$$

Homework. Show that (D, F) is a Randers space. Compute its flag and S -curvatures.

c) The Finsler space $(M = \mathbb{R}^2, F)$ with $\lambda \in [0, \infty), k \geq 1$ and

$$F_{\lambda, k}(y) = \sqrt{a^2 + b^2 + \lambda(a^{2k} + b^{2k})^{1/k}}, \quad \forall y = (a, b) \in T_{(x_1, x_2)} \mathbb{R}^2, (x_1, x_2) \in \mathbb{R}^2.$$

Homework. Show that (D, F) is a Randers space. Compute its flag and S -curvatures. Show that for $\lambda = 0$, the Finsler fundamental function provides the Euclidean norm on $\mathbb{R}^2$.

IV. The Kropina pseudo-Finsler metric $F = \alpha^2/\beta$, has been used in irreversible thermodynamics, as

$$F(y) = y_{n+1}^{-1} \sum_{k=1}^{n+1} y_k^2, \quad y = (y_1, \dots, y_{n+1}) \in D \subset \mathbb{R}^{n+1},$$

and in Fisher statistic Finsler geometry.

V. The (α, β) -Finsler metrics are the ones with the fundamental function of the form

$$F(x, y) = \alpha \Phi \left(\frac{\beta}{\alpha} \right),$$

with α, β given by (12), where Φ is a positive smooth function. We note that for $\Phi(s) = 1$, $\Phi(s) = 1 + s$ and $\Phi(s) = s^{-1}$, we accordingly obtain the Riemannian spaces $F = \alpha$, the Randers spaces $F = \alpha + \beta$, and the Kropina spaces $F = \alpha^2/\beta$.

1.6 Geodesics in (pseudo-)Finsler spaces

A curve $x(t)$ is called *geodesic* of the Finsler space (M, F) iff it satisfies the equations of geodesics

$$\ddot{x}^i + \gamma_{jk}^i(x, \dot{x}) \dot{x}^j \dot{x}^k = 0, \quad i = \overline{1, n}, \quad (13)$$

where $\ddot{x} = \frac{d^2x}{dt^2}$, $\dot{x} = \frac{dx}{dt}$ and $\gamma_{jk}^i = |\gamma_{jk}^i|$ are the *generalized Christoffel symbols of the second kind* provided by the Christoffel symbols of the first kind $|jk; s|$,

$$|\gamma_{jk}^i| = g^{is} |jk; s|, \quad |jk; s| = \frac{1}{2} (\partial_j g_{sk} + \partial_k g_{sj} - \partial_s g_{jk}), \quad i, j, k = \overline{1, n}.$$

Remark. Like in the (sub-)Riemannian case, when the tensor g_{ij} is *degenerate*, then the geodesics still can be defined by avoiding the usage of the Christoffel symbols of the II-nd kind (which involve the existence of the inverse tensor g^{ij}), by using the lowered index version of (13) and the symbols of I-st kind, as

$$g_{is} \ddot{x}^s + |jk; i| \dot{x}^j \dot{x}^k = 0, \quad i = \overline{1, n}. \quad (14)$$

We note that γ_{jk}^i depend on (x, y) , and that this provides a canonic non-linear connection provided by F (called *the Cartan non-linear connection*), whose coefficients are given by:

$$N_i^a = \frac{\partial(\gamma_{jk}^a y^j y^k)}{\partial y^i},$$

that the coefficients $G^i = \frac{1}{2} \gamma_{jk}^i y^j y^k$ (called *spray coefficients*) are 2-homogeneous in y , and that the equations (13) rewrite as

$$\ddot{x}^i + 2G^i(x, \dot{x}) = 0, \quad i = \overline{1, n}.$$

Homework. Using the MAPLE software, compute the generalized Christoffel symbols, the equations of geodesics and plot (as concurrent sheaf and as transversal net) the geodesics of the Euclidean space $\mathbb{R}^2$, of the parametrized sphere $\Sigma = \text{Im } r$, $r(u, v) = (\cos u \sin v, \sin u \sin v, \cos v)$, $(u, v) \in [0, 2\pi] \times [0, \pi]$, and for the Finsler structures from examples III (choose $n = 2$ and convenient initial data for solving/plotting the corresponding 2-nd order SODE).

1.7 The flag curvature of a Finsler space

For a given Finsler linear d-connection $\nabla = \{\tilde{F}_{jk}^i; \mathbb{C}_{jk}^i\}$, one can easily compute the components of the curvature $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$, which, due to the property that ∇ preserves the h - and the v -distributions, provides three nontrivial sets of coefficients, via:

$$R(\delta_l, \delta_k) \delta_j = R_{jkl}^i \delta_i, \quad R(\dot{\partial}_c, \delta_k) = P_{bkc}^a \dot{\partial}_a, \quad R(\dot{\partial}_d, \dot{\partial}_c) = S_{bcd}^a \dot{\partial}_a.$$

Homework. Check that the non-linear connection $N = \{N_i^a\}$ given by Kern's relations (3)-(4) satisfies $N_i^a = \frac{\partial(\gamma_{bc}^a y^b y^c)}{\partial y^i}$ and that its $hh-v$ -torsion R_{jk}^a provided by (7) satisfies $\frac{\partial R_{jk}^a}{\partial y^l} = R_{jkl}^a$.

By means of this curvature tensor, one can easily define the h -curvature of Riemann type, $K(X, Y, Z, V) = g(R(Z, V)Y, X)$, and further, for a family of oriented 2-tangent planes generated by two independent vector fields X, Y , the associated sectional curvature

$$K_{(x,y)}(X, Y) = \frac{K(X, Y; X, Y)}{\text{Gram}_{\det}(X, Y)}, \quad \text{where } \text{Gram}_{\det}(X, Y) = \begin{vmatrix} g(X, X) & g(X, Y) \\ g(Y, X) & g(Y, Y) \end{vmatrix}.$$

We note that this sectional curvature generally depends on the support element $(x, y) \in TM$, hence there exist *three tangent vectors* $y, X_x, Y_x \in T_x M$ which take part in this definition, in the proper Finslerian case. Still, in the particular Riemannian case, for $x \in M$ fixed, this depends just on the vectors $X_x, Y_x \in T_x M$ and provide the classical sectional curvature of the subjacent integral local submanifold $\Sigma_{X,Y} \subset M$, whose tangent plane is spanned by these two vectors, and the formula provides by the Egregium theorem the *Gauss curvature* of $\Sigma_{X,Y}$.

In the proper Finsler case we reduce the number of vectors from three to two, by replacing $X \rightarrow y, Y \rightarrow X$, and we get the so-called *flag curvature* associated to the 2-plane spanned by $\{y, X\}$,

$$K_{(x,y)}(X) = \frac{K(y, X; y, X)}{\text{Gram}_{\det}(y, X)}.$$

The Finsler space is called *of scalar (flag) curvature* if $K_{(x,y)}(X)$ does not depend on the field X , and *of constant flag curvature* if $K_{(x,y)}(X)$ is independent on both y and X .

The role played by the flag curvature is significant in Finsler geometry. Among the latest related results, we note the following:

Theorem (Rademacher 2004). *Let (M, F) be a simply connected compact Finsler structure such that $\dim M \geq 3$. If the flag curvature satisfies the condition*

$$(1 - \frac{1}{1+\lambda}) < K_{(x,y)}(X) \leq 1, \quad \forall x \in M, \forall y \in T_x M, \forall X_x \in T_x M,$$

where

$$\lambda = \sup_{F(x,y)=1} \frac{F(x, -y)}{F(x, y)} = \sup_{y \in I_x} \frac{F(x, -y)}{F(x, y)},$$

then M is homeomorphic to the sphere S^n .

Theorem (Z. Shen 2005). *Let (M, F) be a compact Finsler structure such that $\dim M \geq 3$, of scalar negative curvature. Then F is a Randers metric.*

1.8 The Legendre transform of a Finsler structure

Definition. Let (M, F) be a Finsler structure. Then the mapping $\Lambda : TM \rightarrow T^*M$ locally defined by

$$T(x, y) = (x, p), \quad \text{where } p_i = g_{ij}(x, y)y^j, \quad \forall (x, y) \in TM,$$

is called *the Legendre transform* of the given Finsler structure.

It can be proved that L is a local diffeomorphism, and that it transforms the Finsler norm F into the *dual norm* $F^* : T^*M \rightarrow \mathbb{R}_+$ defined by $F^*(x, p) = \sup_{F(x,y)=1} |p(y)| = \sup_{y \in I_x} |p(y)|$.

Remarks. a) The *dual indicatrix*, i.e., the indicatrix $I_x^* = \{p \in T^*M \mid F^*(x, p) = 1\}$ of F^* can be locally described by its Cartesian implicit equation which is obtained by eliminating the parameters $y^1, \dots, y^n$ from the system of $n+1$ relations

$$\begin{cases} F(x; y^1, \dots, y^n) = 1 \\ F_{y^1}(x; y^1, \dots, y^n) = 0 \\ \dots \\ F_{y^n}(x; y^1, \dots, y^n) = 0. \end{cases}$$

b) One can easily see that the (local) inverse $\Lambda^{-1} : T^*M \rightarrow TM$ is generally hard to obtain. Namely, the pre-image $(x, y) \in TM$ of $(x, p) \in T^*M$ can be determined by finding the unique solution $y = (y^1, \dots, y^n)$ of the system of highly non-linear equations

$$\begin{cases} g(x, y)_{11} \cdot y^1 + \dots + g(x, y)_{1n} \cdot y^n = p_1 \\ \dots \\ g(x, y)_{n1} \cdot y^1 + \dots + g(x, y)_{nn} \cdot y^n = p_n. \end{cases}$$

c) In the (particular) Riemannian case, when $g(x, y)_{ij}$ depends only on $x \in M$, this system is linear and compatible, with the unique solution provided by $y^i = g^{ij} p_j, i = \overline{1, n}$. We then easily note that the Legendre transform becomes in this case a vector bundle isomorphism, acting on the fibres as

$$\Lambda_x : T_x M \rightarrow T_x^* M, \quad \Lambda_x(y) = p, \quad p_i = g_{ij} y^j,$$

thus "lowering" the index ($L_x = L_b$). As well, its inverse acts on fibres as

$$\Lambda_x^{-1} : T_x^* M \rightarrow T_x M, \quad \Lambda_x^{-1}(p) = y, \quad y^i = g^{ij} p_j,$$

thus "raising" the index ($L_x^{-1} = L_{\#}^{-1}$). Hence, often L is called (in the Riemannian case), *the musical isomorphism*.

Homework. Show that the Cartesian indicatrix I_x and the dual indicatrix I_x^* ($x \in \mathbb{R}^4$) of the Berwald-Moor pseudo-Finsler metric are Tzitzeica surfaces, provided by the same Cartesian equations: $I_x : y^1 y^2 y^3 y^4 = 1$; $I_x^* : p_1 p_2 p_3 p_4 = 1$.

1.9 Exam quizzes - part I.

I. Solve several selected by the examiner homeworks.

II. Thoroughly comment the MAPLE code related to geodesic plots.

III. Provide evidence of assimilating basic knowledge regarding one selected by the examiner section.

IV. Check that the provided examples of Finsler structures satisfy the corresponding axioms for (pseudo-)Finsler structures.

2 Symmetries and transformation groups

2.1 Lie algebras and Lie groups

Lie Algebras

Definition We call *Lie algebra*, a vector space V with $[\cdot, \cdot] : V \times V \rightarrow V$ satisfying:

a) $\mathbb{R}$ -bilinearity

b) $[a, b] = -[b, a]$ (skew symmetry)

c) $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$ (the Jacobi identity).

Remark. Using $\text{ad} : V \rightarrow \text{End}(V)$, $\text{ad } a(b) = j_{a*}(b) = [a, b]$ ($j_a(b) = aba^{-1}$), the Jacobi identity confirms that $\text{ad } a$ is a *derivation* $\forall a \in V$, i.e.

$$\text{ad } a([b, c]) = [\text{ad } a(b), c] + [b, \text{ad } a(c)].$$

Examples of Lie algebras

- $\mathbb{R}^3$ with $[u, v] = u \times v$ (the cross product);
- $End(V)$ with $[A, B] = A \circ B - B \circ A$ (using composition);
- $M_n(K)$ with $[A, B] = A \cdot B - B \cdot A$ (using matrix multiplication);
- $\mathcal{X}(M)$ with $[X, Y] = (X^i \cdot \partial_i(Y^j) - Y^i \cdot \partial_i(X^j))\partial_i, \forall X, Y \in \mathcal{X}(U)$;
- for $M = G$ Lie group, $(\mathfrak{g}, [\cdot, \cdot]), \mathfrak{g} \equiv T_e G \equiv \mathcal{X}_r(G) \equiv \mathcal{X}_l(G)$;
- induced Lie bracket on submanifolds of M ;

Lie groups

a) We call *Lie group*, a group $(G, \cdot)$, with G differentiable manifold endowed with a group operation and inversion both smooth mappings.

b) We call *Lie subgroup* a subset $H \subset G$ which is both algebraic subgroup and submanifold of G .

Remark. Any Lie group is an orientable and parallelizable manifold.

Examples of Lie groups

- $(K^n, +)$ and $(K^*, \cdot)$, $K \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$
- $(M_n(V), +); (GL(V), \circ)$
- $(GL(n, K) = GL(K^n), \cdot)$, $K \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$
- $Aff(S) = GL(V) \oplus T$, (V = the linear space associated to the affine space S)

More examples

- G finite or countable set (with the discrete topology of 0-dimensional manifolds)
- for G, G' Lie groups, it results that $G \times G'$ is a Lie group; as consequence: $K^m \times T^n$ where $T^n = \underbrace{S^1 \times \dots \times S^1}_{n \text{ times}}$ is a Lie group as well;
- any closed subgroup of a Lie group, is a Lie group;
- any normal subgroup H of G ($H \triangleleft G \Leftrightarrow \forall g \in G, gHg^{-1} = H$) is a Lie group;
- real Lie groups: subgroups of $GL(n, \mathbb{R})$: $O(n), SL(n, \mathbb{R}), SO(n) = O(n) \cap SL(n, \mathbb{R})$
- complex Lie groups: subgroups of $GL(n, \mathbb{C})$: $U(n), SL(n, \mathbb{C}), SU(n) = U(n) \cap SL(n, \mathbb{C})$
- $T = T^1 = \{z \in \mathbb{C}^* \mid |z| = 1\} = S^1 \subset \mathbb{C}^*$
- W additive subspace of V ; $(K^*)^n \subset GL(K^n)$; $K \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$;
- the Heisenberg-Weil triangular subgroups of $GL(n, K)$ (used in second quantization and coherent states):

$$\begin{aligned} H_R &= \{(a_{ij}) \mid a_{ii} = 1, a_{ij} = 0, \forall i > j\}, \\ H_L &= \{(a_{ij}) \mid a_{ii} = 1, a_{ij} = 0, \forall i < j\} \end{aligned}$$

2.2 Homomorphisms of Lie groups

Definition. We call a mapping $G \xrightarrow{\varphi} G'$ *homomorphism of Lie groups* if φ is both a smooth and an algebraic morphism.

Examples.

- $x \in \mathbb{R} \rightarrow e^x \in \mathbb{R}_+^*$
- $A \in GL(n, K) \rightarrow \det A \in K^*$
- $G \xrightarrow{\varphi} G, \varphi(g) \in \{ag, ga, aga^{-1}\}$ for some $a \in G$
- $x \in \mathbb{R} \rightarrow e^{ix} \in S^1 \in \mathbb{C}^*$
- $(A, b) \in \text{Aff}(S) \rightarrow A \in GL(V)$ (V associated to the affine space S)
- any algebraic morphism from G finite or countable to a Lie group G'

2.3 Lie algebras of Lie groups

- $sl(n, K) = \{A \in M_n(K) | \text{Tr}(A) = 0\}, K \in \{\mathbb{R}, \mathbb{C}\}$
- $o(n, K) = \{A \in M_n(K) | A^t = -A\}, (K \in \{\mathbb{R}, \mathbb{C}\}, o(n) = o(n, \mathbb{R}))$
- $so(n, K) = o(n, K) \cap sl(n, K)$
- $o(p, q) = \{A \in M_n(\mathbb{R}) | AJ + JA^t = 0\}$
 $= \{AJ | A^t = -A\} = o(n)J, \quad J = \text{diag}(I_p, -I_q).$
- $so(p, q) = o(p, q) \cap sl(n, \mathbb{R})$

Examples of Lie algebras

- $u(n) = \{A \in GL(n, \mathbb{C}) | \bar{A}^t = -A\}$ (skew-Hermitian)
- $su(n) = u(n) \cap sl(n, \mathbb{C})$
- $u(p, q) = \{A \in M_n(\mathbb{C}) | AJ + J\bar{A}^t = 0\} = \{AJ | \bar{A}^t = -A\} = su(n)J, J = \text{diag}(I_p, -I_q).$
- $su(p, q) = u(p, q) \cap sl(n, \mathbb{C})$
- Heisenberg-Weil groups have the Lie algebras $Lie(H_U) = \{(a_{ij})_{i,j=\overline{1,n}} | a_{ij} = 0, \forall i \geq j\}$, similar $Lie(H_L)$.
- $sl(2, \mathbb{R}) = \text{Span}_{\mathbb{R}}(\{(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}), (\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}), (\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix})\})$

Special cases

- $so(3, \mathbb{R}) = \text{Span}_{\mathbb{R}}(\{[R_{x,\pi/2}], [R_{y,\pi/2}], [R_{z,\pi/2}]\}, [\cdot, \cdot])$
- $\mathbb{R}^3 = \text{Span}_{\mathbb{R}}(\{e_1, e_2, e_3\}), \times$
- $su(2) = \text{Span}_{\mathbb{R}}(\{(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}), (\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}), (\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix})\}, [\cdot, \cdot])$
- The three Lie algebras from above are isomorphic.

2.4 Lie transformation groups

2.4.1 Transformation groups in the 2D plane

- $T \equiv \mathbb{R}^2$ - translations $x \rightarrow x + x_0$
- $H \equiv \mathbb{R}_+^*$ - homotheties $x \rightarrow kx$
- $H \otimes T$ semidirect product $((a, x) \circ (b, y) = (ab, x + ky))$
- $GL_2(\mathbb{R})$ - linear isomorphisms of $\mathbb{R}^2$
- $\text{Aff}(\mathbb{R}^2) = GL_2(\mathbb{R}) \otimes T$ - the affine group

A. Isometries

- $O(2) = \{A \in GL_2(\mathbb{R}) \mid A^t A = I_2\}$ - orthogonal transformations (linear isometries)
- $\text{Iso}(\mathbb{R}^2) = O(2)_{\otimes} T$ - isometries
- the special orthogonal group (the connected component of $O(2)$, positive linear isometries)

$$SO(2) = \{A \in O(2) \mid \det(A) = 1\} = \{[R_\theta] = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \mid \theta \in \mathbb{R}\},$$

- $SO(2)_{\otimes} T$ (proper movements, positive isometries)

B. Conformal transformations

- $\text{Conf}(\mathbb{R}^2) = (H \times O(2))_{\otimes} T$ - the conformal group
- $\text{Conf}_+(\mathbb{R}^2) = (H \times SO(2))_{\otimes} T$ - orientation preserving (positive) conformal transformations
- $O(2) = SO(2) \cup \{A \cdot \text{diag}(1, -1) \mid A \in SO(2)\}$
- glide reflections (negative isometries; not a group !):

$$(O(2) \setminus SO(2))_{\otimes} T = \{\text{symm}_\Delta \circ T_{v \parallel v_\Delta} \mid \Delta \subset \mathbb{R}^2\}$$

- $\text{Conf}_+(\mathbb{R}^2) \sim \text{Aff}(1, \mathbb{C}) = \{z \rightarrow wz + w_0 \mid z \in \mathbb{C}^*, w_0 \in \mathbb{C}\}$
- $\text{Conf}(\mathbb{R}^2) \sim \{z \rightarrow wz + w_0 \text{ or } z \rightarrow w\bar{z} + w_0 \mid z \in \mathbb{C}^*, w_0 \in \mathbb{C}\}$

C. Projective and concircular transformations

- $PGL(2, \mathbb{R}) = \{x \rightarrow \pi_{12}([A(x^t, 1)]) \mid A \in GL(3, \mathbb{R})\}$ (the projective transformations of the plane)
- the circular transformations:

$$\text{Circ}_2 = PGL(1, \mathbb{C}) \cup \underbrace{\{z \rightarrow \varphi(\bar{z}) \mid \varphi \in PGL(1, \mathbb{C})\}}_{PGL(1, \mathbb{C})'}$$

- $PGL(1, \mathbb{C})$ = proper fractional linear transformations = inversions $\circ$ spiral similarities $\circ$ reflections $\circ$ translations
- $PGL(1, \mathbb{C})'$ = improper fractional linear transformations = inversions $\circ$ spiral similarities $\circ$ translations
- $SO(2)_{\otimes} H$ - spiral similarities of center O
- $x \rightarrow \rho \cdot x / ((x^1)^2 + (x^2)^2)$, inversion of center O

2.4.2 The isometries of $\mathbb{R}^n$

- $n = 3$: positive isometries (proper motions) $\subset$ isometries $\subset$ affine transformations

$$SO(3)_{\otimes} T \subset O(3)_{\otimes} T \subset GL_3(\mathbb{R})_{\otimes} T$$

- isometries of $\mathbb{R}^n$:

$$SO(n)_{\otimes} T \subset O(n)_{\otimes} T \subset GL(n, \mathbb{R})_{\otimes} T$$

- Galilean transformations of classical mechanics (in $\mathbb{R}_{3,1}^4$)

$$(t, x) \rightarrow (t, Ax + x_0 - vt) \sim \begin{pmatrix} A & -v & x_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in PGL(\mathbb{R}^4)$$

- Kepler (celestial mechanics - $T_{rot}^2 P \sim d^3$, $d = d_{min}(P, Sun)$ perihelion):

$$(t, x) \rightarrow (\beta t, \alpha x) = (\beta t, \sqrt[3]{\alpha^2} x), \quad \beta^2 = \alpha^3, \beta > 0$$

2.4.3 Lorentz transformations

- Lorentz transformations which fix the origin in $(\mathbb{R}_{1,1}^2, ds^2 = c^2 dt^2 - (dx^1)^2)$

$$\begin{aligned} O(1, 1) &= \left\{ A = \frac{\pm 1}{\sqrt{1-\beta^2}} \begin{pmatrix} 1 & \pm\beta \\ \beta & \pm 1 \end{pmatrix} \mid |\beta| < 1 \right\} \\ &= \{ A = \pm \begin{pmatrix} c & \pm s \\ s & \pm c \end{pmatrix} \mid \psi \in \mathbb{R} \} \end{aligned}$$

for $\beta = \tanh \psi, c = \cosh \psi, s = \sinh \psi$; four connected pieces (the extremal two are orthocronous):

$$\begin{aligned} &\left\{ \begin{pmatrix} c & s \\ s & c \end{pmatrix}, \begin{pmatrix} c & -s \\ s & -c \end{pmatrix}, \begin{pmatrix} -c & s \\ -s & c \end{pmatrix}, \begin{pmatrix} -c & -s \\ -s & -c \end{pmatrix} \right\} \text{ containing} \\ &\{ I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, PT = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \} \end{aligned}$$

- Lorentz transformations on $\mathbb{R}_{1,n-1}^n$ which fix the origin in

$$(\mathbb{R}_{p,q}^n, ds^2 = c^2 dt^2 - (dx^1)^2 - \dots - (dx^{n-1})^2), \quad p + q = n$$

form $O(p, q)$ (four connected components !)

2.5 Complex transformation groups

Realization of complex Lie groups

- $\mathbb{C}^n$ as real vector space spanned by $\{e_1, \dots, e_n, ie_1, \dots, ie_n\}$
- $\Lambda = A + iB \in M_n(\mathbb{C}) \rightarrow r(\Lambda) = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in M_{2n}(\mathbb{R})$
- $\det(r(\Lambda)) = |\det(\Lambda)|^2$
- $r(GL(n, \mathbb{C})) = \{A \in GL_{2n}(\mathbb{R}) \mid AJ = JA\}, J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$
- $SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) \mid \det(A) = 1\}$
- Realization: $z = x + iy \equiv (x, y), w = u + iv \equiv (u, v)$; inner products:

$$\langle z, w \rangle_{\mathbb{C}} = \sum z^i \bar{w}^i, \quad \langle z, w \rangle_{\mathbb{R}} = \sum z^i u^i + \sum y^i v^i$$

and realization implies $\Re \langle z, w \rangle_{\mathbb{C}} = \langle z, w \rangle_{\mathbb{R}}$.

- $U(n) = \{\Lambda \in GL(n, \mathbb{C}) \mid \langle \Lambda z, \Lambda w \rangle_{\mathbb{C}} = \langle z, w \rangle_{\mathbb{C}}\} = \{\Lambda \in GL(n, \mathbb{C}) \mid \bar{\Lambda}^t \Lambda = I_n\}$
- $r(U(n)) = SO(2n) \cap r(GL(n, \mathbb{C}))$

- similarly in $(\mathbb{C}_{p,q}^n, \langle z, w \rangle_{\mathbb{C}} = \sum_{i=1}^p z^i \bar{w}^i - \sum_{i=p+1}^n z^i \bar{w}^i)$, the metric is preserved by $U(p, q) \supset SU(p, q)$

Particular cases

- $n = 1$: $GL(1, \mathbb{C}) \sim (\mathbb{C}^*, \cdot)$; $r(U(1)) = SO(2)$
- $n = 2$: $SL_2(\mathbb{C}) = \{A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid \det(A) = 1\}$ admits an epimorphism to linear fractional transformations $L = \{\varphi_A : \mathbb{C} \rightarrow \mathbb{C} \mid \varphi_A(z) = \frac{az+b}{cz+d}\}$ of kernel $\mathbb{Z}_2$.
It follows $L \sim SL(n, \mathbb{C})/\mathbb{Z}_2$
- $U(2) = \{A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid |a|^2 + |b|^2 = 1\}$
- $U(1, 1) = \{A = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \mid |a|^2 - |b|^2 = 1\}$
- The latter one admits an injection - not covering $\left\{ \begin{pmatrix} 0 & d \\ -\bar{d} & 0 \end{pmatrix} \mid d \in \mathbb{C} \right\}$,

$$\begin{pmatrix} c & d \\ \bar{d} & \bar{c} \end{pmatrix} \in SU(1, 1) \rightarrow \begin{pmatrix} 1/c & d/\bar{c} \\ -d/\bar{c} & 1/\bar{c} \end{pmatrix} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \in SU(2)$$

2.6 Group actions, orbits and stabilizers

Group actions

Definition. We call *action of a Lie group G on a manifold M* , a mapping $\mu : G \times M \rightarrow M$ such that:

- μ is differentiable,
- $\tau_{gg'} = \tau_g \circ \tau_{g'}$ (where $\tau_g(x) = \mu(g, x)$) and
- $\tau_e = Id_M$.

Examples of actions

- $G \rightarrow \text{Diff}(G)$ via $L_a(g) = ag$, $R_a(g) = ga$, $j_a(g) = aga^{-1}$
- $GL(n, K) \xrightarrow{\varphi} \text{Diff}(P(K^n))$, $\varphi_A(x) = [Ax]$
- linear representations (linear actions): $G \rightarrow GL(V)$
- affine actions $G \rightarrow \text{Aff}(S)$, where S affine space with $\text{Aff}(S) = G \otimes T$
- any homomorphism $H \rightarrow G$ pull-backs the actions of G on manifolds

Orbits and stabilizers

- *free* action, if $\tau_g x = x \Rightarrow g = e$
- *effective* action, if $\tau_g = Id_M \Rightarrow g = e$
- *orbit* of $x \in M$ is $\mathcal{O}_x = \{\tau_g(x) \mid g \in G\} \subset M$
- *transitive* action if $\forall x \in M, \mathcal{O}_x = M$
- *stability (isotropy) group* of $x \in M$ is $G_x = \{g \in G \mid \tau_g(x) = x\}$
- transformation group if $\tau_g \in \text{Diff}(M)$.

Example. $M = \mathbb{R}^3, G = SO(3), S_{e_3} = \text{diag}(1, SO(2)) \sim SO(2) \subset SO(3)$ and

$$\mathcal{O}_{e_3} = \{e_3\}, \mathcal{O}_{e_2} = \{(\cos t, \sin t, 0), t \in \mathbb{R}\}.$$

Remarks.

- for $y = gx \in \mathcal{O}_x$, $G_y = gG_xg^{-1}$ (the stability groups are conjugate along the orbit)
- for $\mu_x : g \rightarrow M$ ($x \in M$), the mapping $\mu_{x*,e} : \mathfrak{g} \equiv T_e G \rightarrow T_x M$ maps the right/left invariant vector fields of $\mathfrak{g}$ to $\mathcal{X}(M)$; its image is a Lie subalgebra of $\mathcal{X}(M)$.

Theorem. If μ_x has constant rank k , then:

- G_x is Lie subgroup of G of codimension k and $T_e(G_x) = \text{Ker}(\mu_{x*,e})$
- $\exists U \ni e$ neighborhood in G s.t. $\tau_U(x) \subset M$ is a submanifold of dimension k and $T_x(\tau_U(x)) = \mu_{x*,e}T_e G$
- if $\mathcal{O}_x$ is a submanifold of M then $\dim \mathcal{O}_x = k$.

2.7 Symmetric spaces

Theorem. If G acts transitively on M , then $M \sim G/H$, where $H = G_x, x \in M$.

Examples

- $S^2 \sim SO(3)/SO(2)$
- $\mathbb{R}P^2 = SO(3)/(SO(2) \times \mathbb{Z}_2)$
- $CP^n \sim S^{2n+1}/S^1$

More examples

- real Stieffel manifolds (k -orthonormal frames in $\mathbb{R}^n$)
 $V_{n,k}^{\mathbb{R}} \sim O(n)/O(n-k) = SO(n)/SO(n-k)$
- complex Stieffel manifolds (k -orthonormal frames in $\mathbb{C}^n$)
 $V_{n,k}^{\mathbb{C}} \sim U(n)/U(n-k) = SU(n)/SU(n-k)$
- real Grassmann (k -dimensional linear subspaces in $\mathbb{R}^n$)
 $G_{n,k}^{\mathbb{R}} \sim O(n)/(O(n-k) \times O(k))$
- oriented k linear subspaces in $\mathbb{R}^n$:
 $\tilde{G}_{n,k}^{\mathbb{R}} \sim SO(n)/(SO(n-k) \times SO(k))$
- complex Grassmann (k -dimensional linear subspaces in $\mathbb{C}^n$)
 $G_{n,k}^{\mathbb{C}} \sim U(n)/(U(n-k) \times U(k))$
- particular case for $k=1$, $CP^{n-1} = G_{n,1}^{\mathbb{C}} \sim U(n)/(U(n-1) \times U(1))$

2.8 Notable examples of actions

- $G = GL(V) \xrightarrow{\varphi} \text{Diff}(B_+(V))$, $\varphi_A(b) = A^t b A$ ($M = B_+(V)$ = bilinear symmetric tensors on V , $\dim V = n$). $G_b = O(V, b)$ is a Lie subgroup of G . If b is *non-degenerate*, then O_b is open in M and

$$\dim O(V, b) = \dim GL(V) - \dim \mathcal{O}_b = \frac{n(n-1)}{2}$$

- $G = GL(V) \xrightarrow{\varphi} \text{Diff}(B_-(V))$, $\varphi_A(b) = A^t b A$ ($M = B_-(V)$ = bilinear skew-symmetric tensors on V , $\dim V = n$). $G_b = Sp(V, b)$ is a Lie subgroup of G . If b is *non-degenerate*, then O_b is open in M and

$$\dim Sp(V, b) = \dim GL(V) - \dim \mathcal{O}_b = \frac{n(n+1)}{2}$$

- $G = GL(V) \xrightarrow{\varphi} \text{Diff}(T_2^1(V))$, $(\varphi_A(b))_{st}^r = b_{ij}^k A_s^i A_t^j (A^{-1})_k^r$

2.9 Infinitesimal isometries (Killing vector fields)

Definition. Let (M, g) be an n -dimensional (pseudo-)Riemannian manifold. $X \in \mathcal{X}(M)$ is a *Killing vector field* if its local 1-parameter group is metric-preserving ($\varphi_{t*}g = g, \forall t \in I$) **Remarks.**

- X Killing iff $(L_X g) = 0 \Leftrightarrow L_X(g(Y, Z)) = g(L_X Y, Z) + g(Y, L_X Z), \forall Y, Z \in \mathcal{X}(M)$
- For $\{e_i\}_{i=1, \dots, n}$ local field orthonormal basis, $L_X e_i = a_i^j(x) e_j$ whence X Killing iff $a_{ij} + a_{ji} = 0$ (with lowering via g).
- the Killing vector fields form a Lie subalgebra in $\mathcal{X}(M)$ of dimension $m \leq \frac{n(n+1)}{2}$.

Example. Let $M = \mathbb{R}^2, g_{ij} = \delta_{ij}$. The isometry group $\text{Iso}(\mathbb{R}^2) = O(2) \ltimes T$ has its Lie algebra of (maximal) dimension 3, generated by the Killing fields

$$\left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right\}$$

2.10 Groups of automorphisms of geometric structures

Definition. A subset G of $\text{Diff}(M)$ is a *Lie transformation group* of M iff G is a Lie subgroup of $\text{Diff}(M)$ and any flow of G is the flow of an infinitesimal symmetry of $\text{Lie}(G)$.

Theorem (Bochner-Montgomery-Zippin, 1955). *Let G be a subgroup of $\text{Diff}(M)$, where M is a differentiable manifold. If G is locally compact, then G is a Lie transformation group of M .*

2.11 Invariant 4-tensors in $\mathbb{R}^4$

Theorem. ([54, 17.1.4]) $T \in \mathcal{T}_4^0(\mathbb{R}^4)$ invariant under $SO(4)$ has the form

$$T_{ijkl} = \lambda \delta_{ik} \delta_{jl} + \mu \delta_{ij} \delta_{kl} + \nu \delta_{il} \delta_{jk}, \quad (\lambda, \mu, \nu \in \mathbb{R}).$$

Theorem. $T \in \mathcal{T}_2^2(\mathbb{R}^4)$ invariant under $GL_4(\mathbb{R})$ has the form

$$T_{kl}^{ij} = \lambda \delta_k^i \delta_l^j + \mu \delta_l^i \delta_k^j, \quad (\lambda, \mu \in \mathbb{R}).$$

2.12 G-structures associated to tensors

Theorem [58]. *Let K be a tensor over $\mathbb{R}^n$ and G the subgroup of $GL(n, \mathbb{R})$ which invariants K . Let P be a G -structure on $M = \mathbb{R}^n$ (a subbundle of the bundle of frames $L(M)$ with fibre G) and let T be the tensor field induced on M by K and P . Then P is integrable (locally the canonic basis of TM provides frames in P) iff there exist coordinates in M on which T has constant components.*

Theorem. *If a G -structure is integrable, then it admits a torsionless connection.*

Remarks.

a) $G = GL(n, \mathbb{R})$. Any smooth manifold admits the G -structure $P = L(M)$, integrable. Any diffeomorphism of M provides an automorphism of P and any vector field on M is an infinitesimal automorphism of P ;

b) $G = GL^+(n, \mathbb{R})$. Any *orientable* smooth manifold admits a G -structure; this is integrable. Any orientation-preserving diffeomorphism of M provides an automorphism of P and any vector field on M with orientation-preserving groups with 1 parameter is an infinitesimal automorphism of P .

2.13 Riemannian isometries

Definition. A map $\varphi : (M, g) \rightarrow (N, h)$ between Riemannian manifolds is a *Riemannian isometry* iff $\varphi^*h = g$ ($h(\varphi_*X, \varphi_*Y) = g(X, Y), \forall X, Y \in \mathcal{X}(M)$).

Theorem (Hopf). Let (M, g) complete connected Riemannian manifold of constant sectional curvature k . Let $(\tilde{M}, \tilde{g})$ be the standard model of simply connected complete Riemannian space with constant curvature k (hyperbolic for $k < 0$, Euclidean for $k = 0$ and standard sphere of radius $1/\sqrt{k}$, for $k > 0$). Then:

- a) There exists a covering projection ("onto"map), smooth local isometry $\varphi : (\tilde{M}, \tilde{g}) \rightarrow (M, g)$.
- b) If M is simply connected as well, then φ is injective, hence diffeomorphism, and the two manifolds are isometric.

2.14 Finslerian isometries

Definition. A map $\varphi : (\tilde{M}, \tilde{F}) \rightarrow (M, F)$ between Finslerian manifolds is a *Finslerian isometry* iff $\varphi^*F = \tilde{F}$ ($F(\varphi_*X) = \tilde{F}(X), \forall X \in \mathcal{X}(\tilde{M})$).

Theorem (Bao-Chern-Shen). If two Finsler structures $(\tilde{M}, \tilde{F})$ and (M, F) are positively homogeneous of degree one, and both are connected, $\tilde{F}$ is geodesically complete and there exists a smooth local isometry $\varphi : (\tilde{M}, \tilde{F}) \rightarrow (M, F)$, then

- a) (M, F) is geodesically complete as well;
- b) φ is a surjective covering projection.

Definition. An isometry in (M, F) is a diffeomorphism $\varphi : M \rightarrow M$ s.t. $F \circ \varphi_* = F$.

Remark. F induces the distance $(d(\varphi(x), \varphi(y))) = d(x, y), \forall x, y \in M$, where $d(x, y) = \inf\{l(c) \mid c(o) = x, c(1) = y\}$ and $l(c) = \int_0^1 F(\frac{dc}{dt})dt$.

Theorem (Deng-Hou 2002, generalization of the Myers-Steenrod theorem). Any smooth map $\varphi \in (\text{Diff}(M))$ is an isometry on (M, F) iff it is distance-preserving.

Theorem (Deng-Hou 2002). The group of isometries $G = \text{Iso}(M)$ of a Finsler space (M, F) is a Lie transformation group and the isotropy subgroup G_x is compact for all $x \in M$.

Theorem (H.C.Wang 1947). If (M, F) is a Finsler space of dimension $n \neq 4$, and if $\dim \text{Iso}(M) > \frac{n(n-1)}{2} + 1$, then (M, F) is Riemannian. **Theorem (Chao-Hao 1957; Egorov 1974).** There exist Finsler non-Riemannian spaces with $\dim \text{Iso}(M) = \frac{n(n-1)}{2} + 1$.

Theorem (L.Kozma 2008). a) If $\|\cdot\|_1, \|\cdot\|_2$ are two Minkowski norms on $\mathbb{R}^n$ and $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with $\|\varphi(a) - \varphi(b)\|_2 = \|a - b\|_1, \forall a, b \in \mathbb{R}^n$, then $\varphi \in \text{Diff}(M)$.

b) If φ is a distance-preserving mapping on (M, F) , then $\varphi \in \text{Diff}(M)$.

2.15 Seminar addenda I. The LU decomposition algorithm.

1. Consider as input the matrix $A \in GL_n(\mathbb{R})$. Check that the Jacobi minors

$$\Delta_k \equiv \det A^{[k]} = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix},$$

of A are all nonzero. In affirmative case, the algorithm continues.

2. Denote $A = A_1$ and subsequently build the matrices $L_k = [e_1, \dots, e_{k-1}, m_k, e_{k+1}, \dots, e_n]$ and $A_{k+1} = L_k A_k$ ($k = \overline{1, n-1}$), where $A_k = (\tilde{a}_{ij})_{i,j=\overline{1,n}}$ and

$$\begin{cases} e_j = (0, \dots, 0, \underbrace{1}_{j\text{-th}}, 0, \dots, 0)^T, & j = \overline{1, n} \\ m_k = \left(0, \dots, 0, \underbrace{1}_{k\text{-th}}, -\frac{\tilde{a}_{k+1,k}}{\tilde{a}_{k,k}}, \dots, -\frac{\tilde{a}_{n,k}}{\tilde{a}_{k,k}} \right)^T, & k = \overline{1, n-1}. \end{cases}$$

3. Denote $L_* = L_1 \cdot \dots \cdot L_{n-1}$. Check that $L_* \in H_L$ and that $U = A_n$ is upper triangular.
4. Compute $L = L_*^{-1}$.
5. Check that $L \in H_L$ and that $A = L \cdot U$.

Remarks. a) For solving a linear system $AX = b$, use the algorithm to build $L_* \cdot A = U$ and re-write the system equivalently as the triangular system $UX = L_* X$, which can be easily solved by backward substitution, by successively solving and replacing the solutions in the backward sequence of equations $n \rightarrow n-1 \rightarrow \dots \rightarrow 1$.

b) The algorithm from above provides the Doolittle decomposition, in which $L \in H_L$ and U is upper triangular.

Further, the LDU decomposition is $A = \hat{L} \hat{D} \hat{U}$, with $\hat{L} \in H_L$, $\hat{D} \in \mathbb{D} = \{\text{diag}(d_1, \dots, d_n) \mid d_k \in \mathbb{R}^*, k = \overline{1, n}\}$ and $\hat{U} \in H_R$, where $\hat{L} = L$, $\hat{D} = \text{diag}(U_{11}, \dots, U_{nn})$ and $\hat{U} = [U_{11}^{-1} L_1^T, \dots, U_{nn}^{-1} L_n^T]^T$, where $L_1, \dots, L_n$ are the lines of the matrix U .

As well, the Crout decomposition provides $A = \bar{L} \bar{U}$, where $\bar{L} = \hat{L} \hat{D}$ is lower triangular and $\bar{R} = \hat{R} \in H_R$.

2.16 Seminar addenda II. The QR decomposition algorithm.

1. Consider as input the matrix $A \in GL_n(\mathbb{R})$. Check that $\det A \neq 0$, in which case the algorithm continues.
2. Denote as $u_1, \dots, u_n$ the column vectors of the matrix A .
3. Apply the Gram-Schmidt orthogonalization process:

$$\begin{cases} v_1 = u_1 \\ v_2 = u_2 - \text{pr}_{v_1} u_2 \\ \dots \\ v_n = u_n - \text{pr}_{v_1} u_n - \dots - \text{pr}_{v_{n-1}} u_n, \end{cases} \Leftrightarrow \begin{cases} u_1 = v_1 \\ u_2 = c_{12} v_1 + v_2 \\ \dots \\ u_n = c_{1n} v_1 + \dots + c_{n-1,n} v_{n-1} + v_n, \end{cases}$$

where $\text{pr}_{v_k} u_l = c_{kl} \cdot v_k$, $c_{kl} = \frac{\langle u_l, v_k \rangle}{\langle v_k, v_k \rangle}$, $l = \overline{2, n}$, $k = \overline{1, l-1}$.

4. Re-write the last system as $A = VR$, where V is the matrix having the column vectors

$v_1, \dots, v_n$ and

$$R = \begin{pmatrix} 1 & c_{12} & \dots & c_{1n} \\ 0 & 1 & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_{n-1,n} \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

5. Note that the column vectors of the matrix Q form an orthonormal basis B' of $\mathbb{R}^n$. Hence check that the scalar product of pairwise column-vectors of Q is zero, and the norms of the column-vectors are all equal to 1.
6. Decompose $V = QD$, where $Q = [a_1^{-1}v_1, \dots, a_n^{-1}v_n]$, $D = \text{diag}(a_1, \dots, a_n)$ and $a_k = \|v_k\|_2$, $k = \overline{1, n}$.
7. Note that the orthogonality of $Q \in O(n)$ infers $\det(Q) \in \{\pm 1\}$. In case that $\det(Q) = -1$, the orthonormal basis B' is negative oriented. In this case transform the B' to a positive-oriented orthonormal basis of $\mathbb{R}^n$, in two possible ways: either replace both the first column of Q and D with their opposites; or interchange two columns of Q and interchange as well the corresponding norms from the diagonal of the matrix D . As result, the column vectors of the new matrix Q will form a *positive oriented* orthonormal basis of $\mathbb{R}^n$ and $Q \in SO(n)$.
8. Point out the QDR decomposition $A = QDR$, with $Q \in SO(n)$, $D \in \mathbb{D}$ and $R \in H_R$.

Remarks. a) The algorithm provides as well the QR decomposition $A = Q \cdot \tilde{R}$, with Q special orthogonal (rotation) matrix and $\tilde{R} = DR$ upper triangular.

b) For solving a linear system $AX = b \Leftrightarrow Q\tilde{R}X = b$, use the algorithm and the orthogonality of the matrix $Q \in SO(n) \subset O(n)$ (namely, $QQ^T = I_n \Rightarrow Q^{-1} = Q^T$) in order to re-write the system equivalently as $\tilde{R}X = \tilde{b}$, where $\tilde{b} = Q^T b$; this is a triangular system which can be easily solved by backward substitution, by successively solving and replacing the solutions in the backward sequence of equations $n \rightarrow n-1 \rightarrow \dots \rightarrow 1$.

2.17 Exam quizzes - part II

Consider the matrix $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \in GL_3(\mathbb{R})$.

I. a) Check that all the principal (Jacobi) minors of A are nontrivial.

b) Apply the *Doolittle*, *LDU* and *CROUT* decompositions for A towards the upper/lower Heisenberg Lie groups.

c) Using the first decomposition and backward substitution, solve the system $AX = b$, where $X = (x, y, z)^T$ and $b = (7, 8, 9)^T$.

II. a) For the matrix A from above, check that $\det A \neq 0$.

b) Apply the *QR decomposition* for A towards $SO(3)$ and the upper Heisenberg Lie groups.

c) Using this decomposition, the orthogonality of Q and the backward substitution, solve the system $AX = b$, where $X = (x, y, z)^T$ and $b = (7, 8, 9)^T$.

III. Check the Lie algebra properties for the examples provided on page 1 of the lectures.

- IV. Check the homomorphism properties for the examples provided on page 2 of the lectures.
- V. Check the Lie algebra properties for the examples provided on page 3 of the lectures.
- VI. Show that the inversion defined on page 4 of the lectures is self-dual.
- VII. Check the consequence of the first realization on page 5 of the lectures.
- VIII. Define a Finsler isometry. State several related results.

Список литературы

- [1] Álvarez Paiva J.C., Breck G. What is wrong with the Hausdorff measure in Finsler spaces // arXiv.math.DG/0408413, 2004; // Adv. Math. 204, 2006, pp.647-663.
- [2] Álvarez Paiva J.C., Thompson A.C. Volumes on normed and Finsler spaces. In: Riemann-Finsler Geometry, MSRI Publications, 49, 2004, pp.1-46.
- [3] Anastasiei M., Vacaru S. Fedosov quantization of Lagrange–Finsler and Hamilton–Cartan spaces and Einstein gravity lifts on (co)tangent bundles // arXiv:0710.3079v3 [math-ph] 15 Nov 2008.
- [4] Asanov G.S. Finsler Geometry, Relativity and Gauge Geometry. D.Reidel, Dordrecht, 1985.
- [5] Balan V. CMC and minimal surfaces in Berwald-Moor spaces // Hypercomplex Numbers in Geometry and Physics, 2(6), 3, 2006, pp.113-122.
- [6] Balan V. CMC and minimal surfaces in Finsler spaces with convex indicatrix // Conference on Integral Geometry, Geometric Inequalities, Finsler Geometry and Convexity / Related Topics, 22-29 Dec. 2006, Tianjin, China,
http://www.cim.nankai.edu.cn/activites/conferences/hy061226/lectures/Presentation_Balan.pdf
- [7] Balan V. Moving frames of Riemannian surfaces in GL^3 spaces // Algebras, Groups and Geometries / Hadronic Press, 17, 2000, 3, pp.273-282.
- [8] Balan V. Notable submanifolds in Berwald-Moór spaces // BSG Proceedings 17 / Geometry Balkan Press 2010, to appear.
- [9] Bao D., Chern S.-S., Shen Z. An Introduction to Riemann-Finsler Geometry. Springer-Verlag, 2000.
- [10] Barthel W. Variationsprobleme der Oberflächenfunktion in der Finslerschen Geometrie // Math. Z. 62, 1955, 23-36.
- [11] Bejancu A. Finsler Geometry and Applications. E.Horwood, 1990.
- [12] Bejancu A. Special immersions of Finsler spaces // Stud.Cerc.Mat. 39, 1987, pp. 463-487.
- [13] Bejancu A., Ferran H.R. Geometry of Pseudo-Finsler Submanifolds // Kluwer Acad. Publ., Dordrecht, 2000.
- [14] Bejancu A., Ferran H.R. On the vertical bundle of a pseudo-Finsler manifold // Internat. J. Math. & Math. Sci. 32, 1999, pp.637-642.
- [15] Belletini G., Paolini M. Some aspects of mean curvature flow in presence of nonsmooth anisotropies // In: 3-rd European Congress of Mathematics (ECM), Barcelona, Spain, July 10-14, 2000, (Ed: Casacuberta, Charles & al), Volume II, Basel, Birkhäuser, Progress in Mathematics 202, 2001, pp.245-253.

- [16] Chernov V.M. On defining equations for the elements of associative and commutative algebras and on associated metric forms // In: (Eds: D.G.Pavlov, Gh.Atanasiu, V.Balan) "Space-Time Structure. Algebra and Geometry Lilia Print, Moscow 2007, pp. 189-209.
- [17] Deng S., Hou Z. A simple proof of a Theorem of H.C. Wang // School of Mathematical Sciences and LPMC, Nankai University, Tianjin, China, 2004, scientific report.
- [18] Deng S., Hou Z. Some applications of Lie theory to Finsler geometry // preprint.
- [19] Deng S., Hou Z. The group of isometries of a Finsler space // Pacific Jour. Math. 207, 2002, 1, 149-155
- [20] Dorfmeister J., Balan V. On the DPW method for general Lie groups. The 932-nd A.M.S. Conference, 27-28 March 1998, Kansas State University, Manhattan, Kansas, USA, lecture.
- [21] Garas'ko G.I. Generalized analytic functions and congruences of geodesics. (in Russian) // Hypercomplex Numbers in Geometry and Physics, 2 (2), 2004, pp. 15-23.
- [22] Garas'ko G.I. The extension of the conformal transformation concept. (in Russian) // Hypercomplex Numbers in Geometry and Physics, 1 (3), (2005), 16-25.
- [23] Garas'ko G.I., Pavlov D.G. The geometry of non-degenerate polynumbers (in Russian) // Hypercomplex Numbers in Geometry and Physics, 1 (7), vol. 4, 2007, pp. 3-25.
- [24] He Q., Shen Y.-B. On Bernstein type theorems in Finsler spaces with the volume form induced from the projective sphere bundle // Proc. of the AMS, 134 (3), 2006, pp. 871-880.
- [25] Hassan A.S.M. Infinitesimal Transformations and Changes in Finsler Geometry and Special Finsler Spaces. Ph. D. Thesis under the supervision of Prof.Dr. Nabil Youssef and Prof.Dr. Sayed H. Abed, Cairo, Egypt 2009.
- [26] Hombu H. Die KrAummungstheorie im Finslerschen Raum // J. Fac. Sci. Hokkaido Univ., Ser. I. Math. 5, 1936, pp.67-94.
- [27] Kobayashi S., Nomizu K. Foundations of Differential Geometry. I, II, Interscience Publishers, New York, 1963, 1969.
- [28] Kozma L. On isometries of Finsler manifolds // Int. Conf. Finsler Geometry, Feb. 15-16 2008, Indianapolis, USA, lecture.
- [29] Lebedev S. Properties of spaces associated with commutative-associative H_3 and H_4 algebras // Hypercomplex Numbers in Geometry and Physics 1 (1), 2004, pp. 63-70; (in Russian volume) pp. 68-74.
- [30] Matsumoto M., Shimada H. On Finsler spaces with 1-form metric II. Berwald-Moór's metric $L = (y^1 y^2 \dots y^n)^{1/n}$ // Tensor N.S. 32, 1978, pp.275-278.
- [31] Miron R., Anastasiei M. The Geometry of Vector Bundles. Theory and Applications // Kluwer Acad. Publishers, FTPH, no.59, 1994.
- [32] Moór A. Über die Dualität von Finslerschen und Cartanschen Räumen // Acta Math. 88, 1952, pp.347-370.
- [33] Moór A. Ergänzung zu meiner Arbeit: Über die Dualität von Finslerschen und Cartanschen Räumen // Acta Math. 91, 1954, pp. 187-188.
- [34] Munteanu Gh. Minimal hypersurfaces of a complex Finsler space // Proc. of Conf. "Riemannian Geometry and Applications" Brasov, June 21-25, 2007, preprint.
- [35] Pavlov D.G. Four-dimensional time // Hypercomplex Numbers in Geometry and Physics, 1(1), 2004, pp. 31-39.
- [36] Pavlov D.G. Generalization of scalar product axioms // Hypercomplex Numbers in Geometry and Physics, 1(1), 2004, pp. 5-18.

- [37] Pavlov D.G., Kokarev S.S. Additive angles in the space $\mathcal{H}_3$ // Hypercomplex Numbers in Geometry and Physics 2 (10), vol. 5 ,2008, pp. 25-43.
- [38] Pavlov D.G., Kokarev S.S. Riemannian metrics, having contact to the 3-dimensional Finslerian Berwald-Moór metric // Hypercomplex Numbers in Geometry and Physics 2 (10), vol. 5, 2008, pp.15-24.
- [39] Perminov N.S. The configuratrix and resultant // Hypercomplex Numbers in Geometry and Physics, 1, 6, 2009, pp. 69-72.
- [40] Qun H.,Shen Y. On Bernstein type theorems in Finsler spaces with the volume form induced from the projective sphere bundle // Proc. of the A.M.S. 134, 2005, No. 3, pp. 871-880.
- [41] Rashevsky P.K. Geometrical Theory of Partial Differential Equations. (in Russian), OGIZ, 1947.
- [42] Rashevsky P.K. Polymetric Geometry (in Russian) // in "Proc. Sem. on Vector and Tensor Analysis and Applications in Geometry, Mechanics and Physics"(Ed. V.F. Kagan), vol. 5, M-L, OGIZ, 1941.
- [43] Rashevsky P.K. Riemann Geometry and Tensor Analysis. (in Russian), Nauka Eds, Moscow, 1967.
- [44] Rund H. The Differential Geometry of Finsler Spaces. Springer-Verlag 1959.
- [45] Rund H. A divergence theorem for Finsler metrics // Monatsh. Math. 79 (1975), 233-252.
- [46] Rund H. A Uber die parallelverschiebung in Finslerschen Raumen // Math. Z. 54 (1951), 115-128.
- [47] Vacaru S. Finsler and Lagrange Geometries in Einstein and String Gravity // Int. J. Geom. Methods. Mod. Phys. (IJGMMP) 5, 2008, pp.473-511; arXiv: 0801.4958 [gr-qc]
- [48] Vacaru S. Clifford and Riemann- Finsler Structures in Geometric Mechanics and Gravity // In: "Selected Works by S. Vacaru, P. Stavrinou, E. Gaburov and D. Gonta; Differential Geometry - Dynamical Systems, Monograph 7, Geometry Balkan Press, 2006, www.mathem.pub.ro/dgds/mono/va-t.pdf and [gr-qc/0508023](http://arxiv.org/abs/0801.4958).
- [49] Youssef N., Abed S.H., Soleiman A. A global approach to the theory of special Finsler manifolds // J. Math. Kyoto Univ. 48, 4, 2008, pp. 857-893, arXiv: 0704.0053 [math.DG]
- [50] Youssef N., Abed S.H., Soleiman A. Geometric objects associated with the fundamental connections in Finsler Geometry // J. Egypt Math. Soc. 17 (2009), arXiv: 0805.2599v2 [math.DG]
- [51] Youssef N., Abed S.H., Soleiman A. A global approach to the theory of connections in Finsler Geometry // arXiv: 0801.3220 [math.DG]
- [52] Shen Z. Differential Geometry of Sprays and Finsler Spaces. Kluwer Acad. Publishers, 2001.
- [53] Bump D. Lie Groups. Springer-Verlag 2004.
- [54] Dubrovin B.A., Fomenko A.T., Novikov S.P. Modern Geometry - Methods and Applications // Graduate Texts in Mathematics 93, Springer-Verlag, 1992.
- [55] Duzhin S.V., Chebotarevski B.D. Transformation Groups // A.M.S. Student Mathematical Library 25, 2004.
- [56] Gökeler M., Schücker T. Differential Geometry, Gauge Theories and Gravity // Cambridge Monographs on Math. Phys. 1987.
- [57] Gorbatshevich V.V., Onischik A.L., Vinberg B.E. Foundations of Lie Theory and Lie Transformation Groups. Springer-Verlag 1997.
- [58] Kobayashi S. Transformation Groups in Differential Geometry. Springer-Verlag 1972.
- [59] Monastyrsky M. Topology of Gauge Fields and Condensed Matter. Plenum Press 1993.
- [60] Stephani H., Kramer D., MacCallum M., Hoenselaers C.,Herlt E. Exact Solutions to Einstein's Field Equations // Cambridge Monographs of Math. Phys. 2003.

- [61] Vacaru S. <http://www.slac.stanford.edu/spires/find/hep/www?rawcmd=f+a+vacaru>
http://arxiv.org/find/grp_physics,grp_math/1/au:+Vacaru/0/1/0/all/0/1
<http://www.scribd.com/people/view/1455460-sergiu>
- [62] Vacaru S., Parametric nonholonomic frame transforms and exact solutions in Gravity // Int. J. Geom. Methods. Mod. Phys. (IJGMMP) 4, 2007, pp.1285-1334; // arXiv: 0704.3986 [gr-qc]

Elements of Complex Analysis

Nicoleta Brnzei

Transilvania University, Brasov, Romania.

E-mail: nico.brnzei@rdslink.ro

The text below is a review of the basic notions and proof techniques in complex analysis, designed for mathematicians and physicists. A special attention is paid to applications to physics and topics such as complex potentials associated to plane vector fields or conformal transformations, as well as to geometric interpretations. The reader is supposed to be acquainted to the basic notions of real analysis and topology.

*"The shortest route between two truths in the real domain
passes through the complex domain."
(Jacques Salomon Hadamard)*

1 Complex Numbers

1.1 Definition and Operations

A *complex number* is an ordered pair $z = (x, y)$ of real numbers. The set

$$\mathbb{C} = \mathbb{R}^2 = \{z \mid z = (x, y), \quad x, y \in \mathbb{R}\}$$

of complex numbers, endowed with the operations

$$\begin{aligned} + & : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, \quad (x, y) + (u, v) = (x + u, y + v) \\ \cdot & : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, \quad (x, y) \cdot (u, v) = (xu - yv, xv + yu) \end{aligned} \tag{1}$$

is a commutative field.

In particular, for pairs $(x, 0) \in \mathbb{C}$, we have:

$$(x, 0) + (u, 0) = (x + u, 0), \quad (x, 0) \cdot (u, 0) = (xu, 0), \quad \forall x, u \in \mathbb{R}.$$

Consequently, the set $\{(x, 0) \mid x \in \mathbb{R}\}$ is a subfield of $(\mathbb{C}, +, \cdot)$, which is isomorphic to $(\mathbb{R}, +, \cdot)$. This isomorphism gives us the right to identify the pair $(x, 0)$ with the real number x and re-denote it as x :

$$x := (x, 0), \quad \forall x \in \mathbb{R}. \tag{2}$$

Multiplying a complex number by a real number a comes down to a rescaling of the components of the latter:

$$a(x, y) = (a, 0)(x, y) = (ax, ay). \tag{3}$$

The number

$$i := (0, 1),$$

is called the *imaginary unit* and has the property:

$$i^2 = -1$$

(where, by (2), -1 means obviously the pair $(-1, 0)$).

With these notations, for any $z = (x, y) \in \mathbb{C}$, we can write $(x, y) = (x, 0) + (0, y) = x(1, 0) + y(0, 1)$, hence we get the *Cartesian* (or *rectangular*, or *algebraic*) form of the complex number z

$$z = x + iy.$$

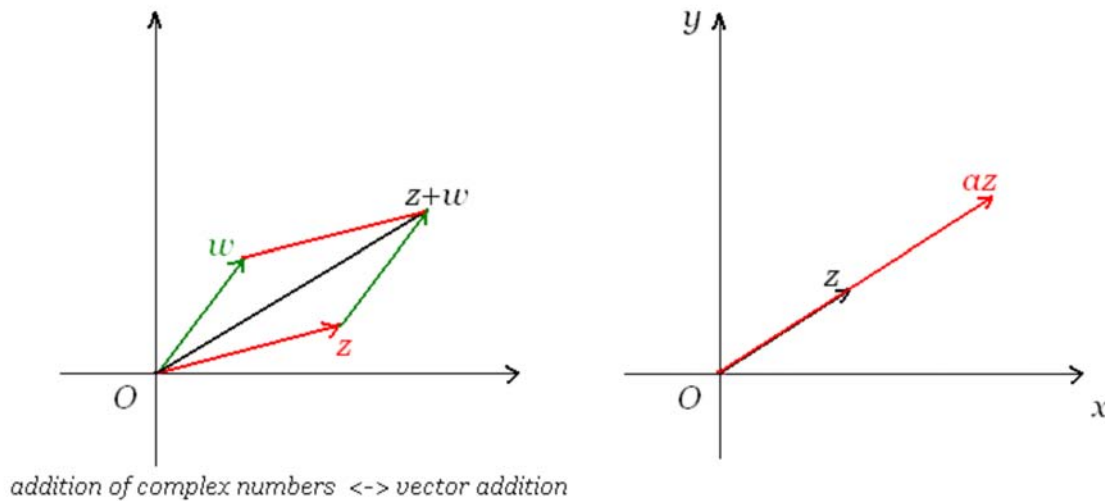
The number $Re z = x$ is called the *real part* of z , while $Im z = y$ is called its *imaginary part*.

Let us now do a little bit of geometry by means of complex numbers.

To the pair $(x, y) \in \mathbb{R}^2$ (equivalently: to the complex number $z = x + iy$) we can naturally attach the vector $\vec{v}(x, y)$ in the Euclidean plane endowed with an orthonormal reference frame xOy (or the point $M(x, y)$ in this plane). So we get the correspondence

$$z = x + iy \leftrightarrow \vec{v}(x, y) \leftrightarrow M(x, y).$$

A brief look at rules (1) and (3) tells us that addition of complex numbers corresponds to the usual addition of free vectors (by the parallelogram rule), while multiplication of complex numbers by real ones corresponds to scalar multiplication of free vectors in plane. Moreover, together with



the addition (1) and the scalar multiplication (3), the set $\mathbb{C}$ is a vector field (of dimension 2) over $\mathbb{R}$.

The *modulus* of the complex number $z = x + iy$ is the nonnegative real number

$$|z| = \sqrt{x^2 + y^2}. \quad (4)$$

The modulus of z can be geometrically interpreted as the Euclidean *norm* of the vector $\vec{v}(x, y)$ (equivalently, as the distance from the origin to point $M(x, y)$). It can be easily checked that the modulus $|| : \mathbb{C} \rightarrow \mathbb{R}^+$ obeys the properties

1. $|z| \geq 0, \forall z \in \mathbb{C}$ and $|z| = 0 \Leftrightarrow z = 0$;
2. $|zw| = |z||w|, \forall z, w \in \mathbb{C}$;
3. the triangle inequality: $|z + w| \leq |z| + |w|, \forall z, w \in \mathbb{C}$.

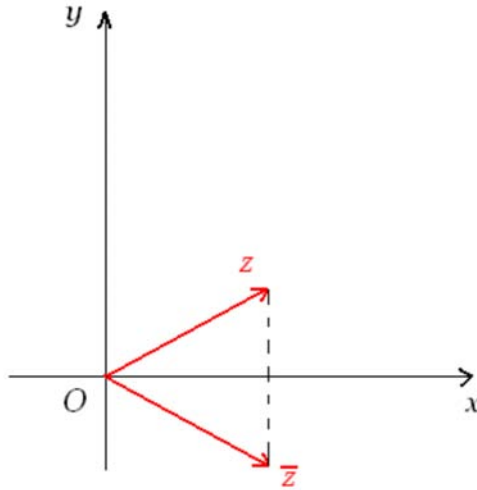
Properties 1 and 3 are the properties of a norm, while 2 is more general than the corresponding property of norms. It also holds: $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}, \forall z, w \in \mathbb{C}, w \neq 0$.

In particular, for real numbers $z = x + i0 \in \mathbb{R}$, definition (4) yields $|x| = \sqrt{x^2}$, which coincides with the definition of the modulus (absolute value) of a real number. Hence, (4) extends to $\mathbb{C}$ the definition of the modulus of real numbers.

The *conjugate* of $z = x + iy \in \mathbb{C}$ is defined as

$$\bar{z} = x - iy.$$

Geometrically, conjugation corresponds to reflection in the Ox -axis:



Properties of the conjugate:

1. $z\bar{z} = |z|^2, \forall z \in \mathbb{C}$;
2. $\overline{z \pm w} = \bar{z} \pm \bar{w}, \forall z, w \in \mathbb{C}$;
3. $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}, \forall z, w \in \mathbb{C}, w \neq 0$.

Exercises:

1. Prove properties 1-3 of the conjugate of a complex number.
2. Prove that, for all $z, w \in \mathbb{C}$, there holds $||z| - |w|| \leq |z - w|$.
3. Draw the set of points $z = (x, y)$ in plane with the properties: a) $|z - 2 + 3i| = 2$;
b) $\operatorname{Re}(\bar{z} - i) = 4$; c) $|z + 1| + |z - 1| = 4$; d) $|z + 1| - |z - 1| = 4$; e) $|z - 1 + 2i| = |z + 3 + i|$.

1.2 Polar Form

For $z = x + iy$, by denoting

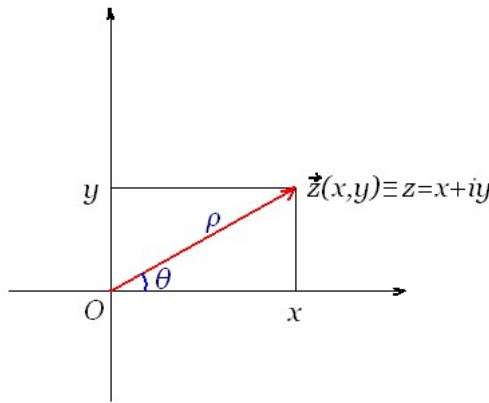
$$\rho = |z|, \quad \theta = \operatorname{Arg}(z) \in (-\pi, \pi],$$

the *modulus* and the *principal argument*, respectively, that is,

$$\begin{cases} \rho = \sqrt{x^2 + y^2} \\ \tan \theta = \frac{y}{x} \end{cases} \leftrightarrow \begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases},$$

the complex number z can be written in *polar form* as:

$$z = \rho(\cos \theta + i \sin \theta). \quad (5)$$



The relation above also holds for any $\theta \in \arg(z) = \{\operatorname{Arg}(z) + 2k\pi \mid k \in \mathbb{Z}\}$. The set

$$\arg(z) = \{\operatorname{Arg}(z) + 2k\pi \mid k \in \mathbb{Z}\}$$

is called the *extended argument* of z .

The polar form is most suitable for multiplication and raising to integer powers of complex numbers. Namely, when multiplying $z_1 = \rho_1(\cos \theta_1 + i \sin \theta_1)$ with $z_2 = \rho_2(\cos \theta_2 + i \sin \theta_2)$ moduli are multiplied, while arguments are to be added:

$$z_1 z_2 = z_1 \rho_1 \rho_2 (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)).$$

Particularly, for the n -th power of a complex number (5), there holds *Moirve's formula*:

$$z^n = \rho^n (\cos n\theta + i \sin n\theta), \quad \forall n \in \mathbb{Z}.$$

Geometric interpretation:

Multiplication of a complex number w by z in (5) corresponds to a homothety of ratio ρ and a rotation of angle θ around the origin applied to w . If $\rho = |z| = 1$, then wz will have as its image in plane the point obtained by rotating the image of w around O by an angle equal to $\operatorname{Arg}(z)$.

The equation $z^n = w$ has the distinct solutions

$$z_k = \sqrt[n]{R} \left(\cos \frac{\Theta + 2k\pi}{n} + i \sin \frac{\Theta + 2k\pi}{n} \right), \quad R = |w|, \quad \Theta = \operatorname{Arg}(w), \quad k = 1, \dots, n, \quad (6)$$

that is, a complex number w has n distinct n -th roots $z_1, \dots, z_n$.

Exercises:

1. Write in polar form: $1 + i, -2, 3i, 1 + i\sqrt{3}$.
2. Calculate: $(1 + i)^{100}$.
3. Solve $z^3 = 1, z^4 = 16i$.

So far, we have seen that: translation, homothety, rotation, reflections into the coordinate axes, can be described in terms of operations with complex numbers. There is one more geometric transformation which has an elegant description by means of complex numbers, namely, *inversion*.

For a point M in plane, its *inverse* with respect to the circle $\mathcal{C}(O, R)$ is defined as the point M' situated on the line OM , such that $\overrightarrow{OM} \cdot \overrightarrow{OM'} = R^2$. In terms of complex numbers, if M corresponds to $z \in \mathbb{C}$ and M' , to $z' \in \mathbb{C}$, then

$$z' = \frac{R^2}{\bar{z}}.$$

Proof. Let $z = |z|(\cos \theta + i \sin \theta)$. Relation $\overrightarrow{OM} \cdot \overrightarrow{OM'} = R^2$ is equivalent to $|z||z'| = R^2$, that is, $|z'| = \frac{R^2}{|z|}$. The collinearity of $\overrightarrow{OM}$ and $\overrightarrow{OM'}$ means that z and z' have the same argument θ . We get that $z' = \frac{R^2}{|z|}(\cos \theta + i \sin \theta) = \frac{R^2}{|z|(\cos(-\theta) + i \sin(-\theta))} = \frac{R^2}{\bar{z}}$. ■

2 Complex Functions of a Real Variable

Let $I \subset \mathbb{R}$ be an interval. Functions $\gamma : I \rightarrow \mathbb{C}, t \mapsto x(t) + iy(t)$, are obviously identified with real vector-valued functions $t \mapsto (x(t), y(t))$ defined over I . Hence, the notions of limit, continuity, differentiability and integrability reduce to the corresponding notions from analysis in $\mathbb{R}^n$. For instance, the derivative of γ is

$$\gamma'(t) = x'(t) + iy'(t), \quad \forall t \in I$$

and the definite integral over $[a, b] \subset I$,

$$\int_a^b \gamma(t) dt = \int_a^b x(t) dt + i \int_a^b y(t) dt.$$

Geometrically, functions $\gamma : I \rightarrow \mathbb{C}$ of a real variable describe plane curves.

Examples:

1. The function $\gamma : [-1, 1] \rightarrow \mathbb{C}, \gamma(t) = t + it^2$ represents the arc of the parabola $y = x^2$ corresponding to $x \in [-1, 1]$.
2. Let

$$e^{i\theta} := \cos \theta + i \sin \theta, \quad \theta \in \mathbb{R} \tag{7}$$

(which describes the unit circle $\mathcal{C}(0, 1)$). Then, there hold the properties:

$$e^{i(\theta+\phi)} = e^{i\theta} \cdot e^{i\phi}; \quad e^{i(\theta-\phi)} = \frac{e^{i\theta}}{e^{i\phi}}; \quad (e^{i\theta})^n = e^{in\theta}, \quad \forall \theta, \phi \in \mathbb{R}, \quad n \in \mathbb{Z}.$$

In the following, we shall present a solution of a classical problem, [3], by means of the above defined exponential $e^{i\theta}$, $\theta \in \mathbb{R}$.

Example 1: Let a body of mass M be fixed and a body of mass m orbiting around M ; let us suppose that M is situated at the origin $O(0,0)$ and m has the position vector $z = z(t) \in \mathbb{C}$, where the parameter $t \in I \subset \mathbb{R}$ designates time.

Let us determine the trajectory $z = z(t)$ of m . The gravitational force exerted by M upon m is described by

$$F = \frac{GMm - z(t)}{|z(t)|^2 |z(t)|}$$

(the minus sign corresponds to the orientation towards M of the force).

But, on the other side, according to Newton's law, F is equal to mass multiplied by acceleration $F = mz''(t)$. By equating the two expressions for F , we get:

$$z'' = -\frac{GMz}{|z|^3}. \quad (8)$$

The polar form of z can be written as $z = re^{i\theta}$, where $r = r(t)$ and $\theta = \theta(t)$. Performing the derivations, we get:

$$z' = r'e^{i\theta} + rie^{i\theta}\theta', \quad z'' = r''e^{i\theta} + 2ir'\theta'e^{i\theta} - r(\theta')^2e^{i\theta} + ir\theta''e^{i\theta}.$$

Taking $e^{i\theta}$ as a common factor and replacing into (8), we get:

$$(r'' + 2ir'\theta' - r(\theta')^2 + ir\theta'')e^{i\theta} = -\frac{GM}{r^2}e^{i\theta}.$$

Simplifying by $e^{i\theta}$ and equating the real and imaginary parts of both sides of the remaining equation, we get:

$$\begin{cases} r'' - r(\theta')^2 = -\frac{GM}{r^2} \\ 2\theta'r' + r\theta'' = 0. \end{cases}$$

The latter equation is equivalent to $(r^2\theta')' = 0$, which leads to

$$r^2\theta' = \alpha = \text{const} \Rightarrow \theta' = \frac{\alpha}{r^2}.$$

Replacing into the first equation of the system, we get

$$r'' - \frac{\alpha^2}{r^3} = -\frac{GM}{r^2}. \quad (9)$$

Equation (9) becomes easier to solve if we perform a change of variable from t (time) to θ . Namely, we have

$$r' = \frac{dr}{d\theta}\theta' = \frac{dr}{d\theta} \frac{\alpha}{r^2} = -\alpha \frac{d}{d\theta} \left(\frac{1}{r} \right), \quad r'' = -\alpha \theta' \frac{d^2}{d\theta^2} \left(\frac{1}{r} \right) = -\frac{\alpha^2}{r^2} \frac{d^2}{d\theta^2} \left(\frac{1}{r} \right).$$

Denoting $s := \frac{1}{r}$ and replacing into (9),

$$-\alpha^2 s^2 \frac{d^2 s}{d\theta^2} - \alpha^2 s^3 = -GM s^2,$$

which is,

$$\frac{d^2s}{d\theta^2} + s = \frac{GM}{\alpha^2}.$$

The latter is a linear inhomogeneous equation with constant coefficients. The general solution is $s = c_1 \cos \theta + c_2 \sin \theta + \frac{GM}{\alpha^2} =: A \cos(\theta + \phi) + \frac{GM}{\alpha^2}$. This is,

$$r = \frac{1}{A \cos(\theta + \phi) + \frac{GM}{\alpha^2}},$$

or, equivalently,

$$r = \frac{\alpha^2}{1 + \varepsilon \cos(\theta + \phi)},$$

$\varepsilon = \frac{A\alpha^2}{GM}$, which is nothing but the polar equation of a conic of eccentricity ε .

3 Functions of a Complex Variable

3.1 Limit. Continuity. Differentiability

In the following, by "domain we shall always mean an open and connected set.

Let $D \subset \mathbb{C}$ be a domain and $f : D \rightarrow \mathbb{C}$. We shall examine functions

$$f : D \rightarrow \mathbb{C}, \quad f(z) \equiv f(x, y) = u(x, y) + iv(x, y).$$

Actually, they are functions defined over some region in plane, and having their values in plane. The real functions $u, v : D \rightarrow \mathbb{R}$ are called the real and the imaginary part of f respectively:

$$u = \operatorname{Re} f, \quad v = \operatorname{Im} f.$$

Example: The function $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = z^2 = (x + iy)^2$ can be written as $f(z) = (x^2 - y^2) + 2ixy$, hence $u(x, y) = x^2 - y^2$, $v(x, y) = 2xy$.

As a normed space, $\mathbb{C}$ is actually $\mathbb{R}^2$ endowed with the Euclidean norm, hence the notion of limit on $\mathbb{C}$ coincides with that one in $\mathbb{R}^2$ with the norm $\|(x, y)\| = \sqrt{x^2 + y^2} (= |x + iy|_{\mathbb{C}})$. That is, a neighborhood of some $z_0 \in \mathbb{C}$ is a small disk $|z - z_0| < \varepsilon$ ($\varepsilon > 0$) centered at z_0 .

- We say that $L \in \mathbb{C}$ is the *limit* of $f : D \rightarrow \mathbb{C}$ at $z_0 \in D'$ (z_0 is an accumulation point of D) and we write $\lim_{z \rightarrow z_0} f(z) = L \in \mathbb{C}$, if

$$\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) : \forall z \in D, |z - z_0| < \delta \Rightarrow |f(z) - L| < \varepsilon,$$

where $|\cdot|$ denotes the complex modulus $|x + iy| = \sqrt{x^2 + y^2}$.

We are going to discuss a little bit later what $\lim_{z \rightarrow z_0} f(z) = \infty$ means, for the moment we only take into account the case when the limit is "finite": $L \in \mathbb{C}$.

- $f : D \rightarrow \mathbb{C}$ is said to be *continuous* at $z_0 \in D$ if there exists the limit $\lim_{z \rightarrow z_0} f(z)$ and it is equal to $f(z_0)$: $\lim_{z \rightarrow z_0} f(z) = f(z_0)$.

The properties of the complex limit are similar to those of functions defined on subsets in $\mathbb{R}^n$:

- 1) $\lim_{z \rightarrow z_0} (f(z) \pm g(z)) = \lim_{z \rightarrow z_0} f(z) \pm \lim_{z \rightarrow z_0} g(z)$;
- 2) $\lim_{z \rightarrow z_0} (f(z) \cdot g(z)) = \lim_{z \rightarrow z_0} f(z) \cdot \lim_{z \rightarrow z_0} g(z)$;
- 3) $\lim_{z \rightarrow z_0} \frac{f(z)}{g(z)} = \frac{\lim_{z \rightarrow z_0} f(z)}{\lim_{z \rightarrow z_0} g(z)}$, provided that $\lim_{z \rightarrow z_0} g(z) \neq 0$.

4) Also, the limit of a complex function at a point is the sum of the limits of its real and imaginary parts respectively at that point: if $f(z) = u(x, y) + iv(x, y)$, then

$$\lim_{z \rightarrow z_0} f(z) = \lim_{(x,y) \rightarrow (x_0,y_0)} u(x, y) + i \cdot \lim_{(x,y) \rightarrow (x_0,y_0)} v(x, y),$$

(where $z = x + iy$, $z_0 = x_0 + iy_0$).

The set $\mathbb{C}$ can be compactified by adding a single point, denoted by ∞ :

$$\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}.$$

A neighborhood of ∞ is a set

$$V_\infty = \{z \in \mathbb{C} \mid |z| > M\}.$$

That is, if D is a set of complex numbers and $f : D \rightarrow \mathbb{C}$, then

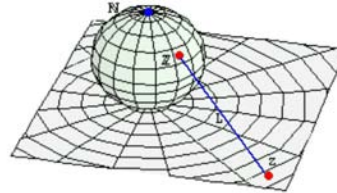
1. If z_0 is an accumulation point for D , we say that $\lim_{z \rightarrow z_0} f(z) = \infty$, if for any $M > 0$, there exists some $\delta > 0$ such that $|z - z_0| < \delta \Rightarrow |f(z)| > M$.
2. We say that $\lim_{z \rightarrow \infty} f(z) = L \in \mathbb{C}$, if for any $\varepsilon > 0$, there exists some $\delta > 0$ such that $|z| > \delta \Rightarrow |f(z) - L| < \varepsilon$.
3. We say that $\lim_{z \rightarrow \infty} f(z) = \infty$, if for any $M > 0$, there exists some $N > 0$ such that $|z| > N \Rightarrow |f(z)| > M$. (In cases 2 and 3, ∞ has to be an "accumulation point" for D).

Rules of computation with ∞ are immediate:

- 1) $a + \infty = \infty$, $\forall a \in \mathbb{C}$;
- 2) $a \cdot \infty = \infty$, $\forall a \in \bar{\mathbb{C}} \setminus \{0\}$;
- 3) $\frac{a}{\infty} = 0$, $\frac{a}{0} = \infty$, $\forall a \in \mathbb{C} \setminus \{0\}$

The quantities $\infty + \infty$, $\frac{\infty}{\infty}$, $0 \cdot \infty$ are indeterminates.

An intuitive way of visualizing infinity in $\mathbb{C}$ is the *Riemann sphere*:



Let the complex plane be endowed with a Cartesian frame xOy , and a sphere (S) be tangent to the plane at O , and N the opposite point to O on the sphere. To point $z = x + iy \equiv (x, y)$, we

make correspond the point on Z the sphere, obtained by intersecting the sphere with the straight line Nz . The correspondence

$$z \mapsto Z : \mathbb{C} \rightarrow (S) \setminus \{N\}$$

is bijective. To point N on the sphere, there naturally corresponds the point $\infty \in \bar{\mathbb{C}}$, thus extending the above correspondence to a bijective function $f : \bar{\mathbb{C}} \rightarrow (S)$.

The notion of differentiability of a complex function requires a special attention.

Definition 1 A function $f : D \rightarrow \mathbb{C}$ is said to be differentiable at $z_0 \in D$ if there exists in $\mathbb{C}$ the limit

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

If f is differentiable at z_0 , the number $f'(z_0) \in \mathbb{C}$ is called the derivative of f at z_0 .

Example 2 1. For $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = z^2$, we get $f'(z_0) = \lim_{z \rightarrow z_0} \frac{z^2 - z_0^2}{z - z_0} = \lim_{z \rightarrow z_0} (z + z_0) = 2z_0$, for any $z_0 \in \mathbb{C}$, hence f is differentiable all over $\mathbb{C}$ and there holds a similar rule of derivation to the one of real functions, namely, $(z^2)' = 2z$.

2. Not the same can be said about the conjugation $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = \bar{z}$. For arbitrary $z_0 = x_0 + iy_0 \in \mathbb{C}$, we have $f'(z_0) = \lim_{z \rightarrow z_0} \frac{\bar{z} - \bar{z}_0}{z - z_0}$. Amplifying by $\bar{z} - \bar{z}_0$, it follows

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{(\bar{z} - \bar{z}_0)^2}{|z - z_0|^2} = \lim_{z \rightarrow z_0} \frac{(\overline{z - z_0})^2}{|z - z_0|^2}.$$

Let us consider $w = z - z_0 = x + ix = (1 + i)x$. Then, the above limit is

$$f'(z_0) = \lim_{x \rightarrow 0} \frac{(1 - i)^2 x^2}{2x^2} = -i.$$

On the other side, we can take as well $w = z - z_0 = y - iy$ and then,

$$f'(z_0) = \lim_{y \rightarrow 0} \frac{(1 + i)^2 y^2}{2y^2} = i.$$

That is, the derivative of $f(z) = \bar{z}$ is not uniquely defined, at any $z_0 \in \mathbb{C}$! In other words, f is not differentiable anywhere in $\mathbb{C}$.

Having a brief look at Definition 1, it immediately follows that the rules of derivation remain the same as for real functions. If f and g are differentiable at some point $z_0 \in \mathbb{C}$, then at z_0 there hold: $(f \pm g)' = f' \pm g'$, $(cf(z))' = cf'(z)$, $\forall c \in \mathbb{C}$ - constant, $(f \cdot g)' = f'g + fg'$, $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$ (provided that $g(z_0) \neq 0$). Also, if g is differentiable at z_0 and f is differentiable at $g(z_0)$, then $(f \circ g)'(z_0) = f'(g(z_0))g'(z_0)$.

There exists a specific terminology related to differentiability of complex functions. Namely:

- If f has a derivative at some $z_0 \in \mathbb{C}$, then f is said to be *differentiable* at z_0 ;
- if f is differentiable on a neighborhood of z_0 , then we say that f is *holomorphic* (or *analytic*) at z_0 ;
- if f is analytic at every $z_0 \in \mathbb{C}$, then we call it an *entire function*.

3.2 Cauchy-Riemann Equations

Looking back to Example 2 in the previous subsection, a natural question arises: can we find some *simple* and easy-to-use necessary and sufficient conditions of differentiability for a complex function? The answer is given by the famous Cauchy-Riemann equations.

Let us suppose that $f : D \rightarrow \mathbb{C}$ is differentiable at some point z_0 belonging to the domain D . The derivative of f at z_0 is

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}.$$

We shall follow a similar track to what we have done in Example 2 above: choose two different particular $\Delta z \rightarrow 0$, and impose that they lead to the same value $f'(z_0)$.

So, let us choose for the beginning, $\Delta z = \Delta x + 0i$. Then.

$$f'(z_0) = \lim_{\Delta x \rightarrow 0} \left(\frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} + i \frac{v(x_0 + \Delta x, y_0) - v(x_0, y_0)}{\Delta x} \right),$$

which is nothing but

$$f'(z_0) = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0). \quad (10)$$

Now, let us take $\Delta z = 0 + i\Delta y$. We get

$$f'(z_0) = \lim_{\Delta y \rightarrow 0} \left(\frac{u(x_0, y_0 + \Delta y) - u(x_0, y_0)}{i\Delta y} + i \frac{v(x_0, y_0 + \Delta y) - v(x_0, y_0)}{i\Delta y} \right),$$

which leads, together with $\frac{1}{i} = -i$, to

$$f'(z_0) = \frac{\partial v}{\partial y}(x_0, y_0) - i \frac{\partial u}{\partial y}(x_0, y_0). \quad (11)$$

In order that $f'(z_0)$ should be uniquely defined, expressions 10 and 11 have to be equal. By equating their real and imaginary parts respectively, we obtain

Theorem 3 *Let $f : D \rightarrow \mathbb{C}$ and $z_0 \in D$. A necessary condition for f to be differentiable at z_0 is that at z_0 $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ are differentiable and there hold the Cauchy-Riemann equations:*

$$\begin{cases} \frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \\ \frac{\partial v}{\partial x}(x_0, y_0) = -\frac{\partial u}{\partial y}(x_0, y_0) \end{cases}. \quad (12)$$

The Cauchy-Riemann equations are generally not sufficient for differentiability.

Still, under supplementary restrictions, they become also sufficient. Namely:

Theorem 4 *If for the function $f : D \rightarrow \mathbb{C}$, Cauchy-Riemann equations (12) hold at $z_0 \in \mathbb{C}$ and the functions $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ admit continuous first order derivatives at z_0 , then f is differentiable at z_0 .*

Proof. Let us suppose that equations (12) hold at z_0 , and $u, v \in \mathcal{C}^1(z_0)$.

Then,

$$\frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0)}{\Delta x + i\Delta y} + i \frac{v(x_0 + \Delta x, y_0 + \Delta y) - v(x_0, y_0)}{\Delta x + i\Delta y}.$$

The first numerator in the right hand side is

$$u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0) = [u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0 + \Delta y)] -$$

$$- [u(x_0, y_0 + \Delta y) - u(x_0, y_0)] = \Delta x \frac{\partial u}{\partial x}(\xi_1, y_0 + \Delta y) + \Delta y \frac{\partial u}{\partial y}(x_0, \xi_2).$$

By the continuity of the partial derivatives of u , this yields $\Delta x [\frac{\partial u}{\partial x}(x_0, y_0) + \varepsilon_1] + \Delta y [\frac{\partial u}{\partial y}(x_0, y_0) + \varepsilon_2]$, with $\varepsilon_1, \varepsilon_2 \rightarrow 0$ as $\xi_1 \rightarrow x_0, \xi_2 \rightarrow y_0$. By Cauchy-Riemann equations, it becomes:

$$\Delta x \frac{\partial u}{\partial x}(x_0, y_0) - \Delta y \frac{\partial v}{\partial x} + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y.$$

Similarly, the second numerator (multiplied by i) is:

$$iv(x_0 + \Delta x, y_0 + \Delta y) - iv(x_0, y_0) = i[\Delta x \frac{\partial v}{\partial x}(x_0, y_0) + \Delta y \frac{\partial v}{\partial y}(x_0, y_0) + \varepsilon_3 \Delta x + \varepsilon_4 \Delta y],$$

with $\varepsilon_3, \varepsilon_4 \rightarrow 0$ as $\Delta x \rightarrow 0, \Delta y \rightarrow 0$.

We get then $f(z_0 + \Delta z) - f(z_0) = \frac{\partial u}{\partial x}(x_0, y_0)(\Delta x + i\Delta y) + i \frac{\partial v}{\partial x}(x_0, y_0)(\Delta x + i\Delta y) + \varepsilon$, where $\varepsilon = \varepsilon_1 \Delta x + \varepsilon_2 \Delta y + i(\varepsilon_3 \Delta x + \varepsilon_4 \Delta y)$ obviously tends to 0 even when divided by $\Delta x + i\Delta y$

Hence,

$$\lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0)$$

is uniquely defined, no matter how we choose the small variation Δz . That is, f is differentiable at z_0 . ■

As a useful remark, if $f = u + iv$ is differentiable at $z_0 \in \mathbb{C}$, then its derivative coincides with its partial derivative with respect to x :

$$f'(z) = \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0).$$

Also, following a similar technique as for real functions, one can prove:

Proposition 5 *If $f : D \rightarrow \mathbb{C}$ is differentiable at $z_0 \in \mathbb{C}$, then it is also continuous at z_0 .*

Exercises:

1) Find the points $z_0 = x_0 + iy_0 \in \mathbb{C}$ where the function $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = \bar{z}^2$ is differentiable. Is f analytic at those points?

2) Same question for $f(z) = x^2 + (1 - i)y^2$.

3.3 Harmonic Functions

A first idea of the value of holomorphic functions for practical applications is their relation to harmonic functions.

Definition 6 A real-valued function $\phi : D \rightarrow \mathbb{R}$, $(x, y) \mapsto \phi(x, y)$ defined over a domain D in $\mathbb{R}^2$ and of class at least 2 on D , is called *harmonic*, if it satisfies Laplace's equation

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0.$$

Applications of harmonic functions are numerous, and they include problems in electrostatics, thermodynamics, fluid mechanics etc.

Theorem 7 If $f = u + iv$ is holomorphic on a domain D and $u = \operatorname{Re} f, v = \operatorname{Im} f$ are of class $\mathcal{C}^2$ on D , then u and v are harmonic functions on D .

The proof of this statement is immediate, from Cauchy-Riemann equations and Schwartz's theorem:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &\stackrel{C-R}{=} \frac{\partial^2 v}{\partial x \partial y} \stackrel{\text{Schwartz}}{=} \frac{\partial^2 v}{\partial y \partial x} \stackrel{C-R}{=} -\frac{\partial^2 u}{\partial y^2} \Rightarrow \\ \Rightarrow \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} &= 0. \end{aligned}$$

Functions u and v which constitute the real and the imaginary part of a holomorphic function respectively are called *harmonic conjugates* to each other (i.e., they are both harmonic and they are related by Cauchy-Riemann equations).

If we also impose that the domain D should be *simply connected* (intuitively: "with no holes in it"), then the converse is also true:

Theorem 8 If u is a harmonic function on a simply connected domain $D \subset \mathbb{C}$, then there exists a holomorphic function $f : D \rightarrow \mathbb{C}$ such that

$$u = \operatorname{Re}(f)$$

and a holomorphic function $g : D \rightarrow \mathbb{C}$ such that

$$u = \operatorname{Im}(g).$$

Proof. Let u be harmonic. We look for a harmonic conjugate v of u , that is, for $v : D \rightarrow \mathbb{R}$ such that

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Let: $\omega = -\frac{\partial u}{\partial y}dx + \frac{\partial u}{\partial x}dy$. The harmonicity of u tells us that actually ω is a closed 1-form on $D \subset \mathbb{R}^2$. Since the domain is simply connected, it is also exact, hence there exists $\psi : D \rightarrow \mathbb{R}$ with $d\psi = \omega$ (namely, ψ is the value of the path-independent integral $\psi(x, y) = \int_{(x_0, y_0)}^{(x, y)} \omega$, for $(x_0, y_0) \in D$ arbitrary and fixed). That is,

$$\frac{\partial \psi}{\partial x} = -\frac{\partial u}{\partial y}, \quad \frac{\partial \psi}{\partial y} = \frac{\partial u}{\partial x}.$$

Setting $v := \psi$ and

$$f := u + iv,$$

(f is obviously holomorphic in D), the proof is completed. The existence of a holomorphic function g such that $u = \operatorname{Im} g$ is proved similarly. ■

4 Elementary Functions

We shall analyze in the following those complex functions, which provide analogues of the elementary functions from real calculus.

4.1 Polynomial and Rational Functions

They are defined by $f : D \rightarrow \mathbb{C}$, $f(z) = \frac{a_n z^n + a_{n-1} z^{n-1} + \dots + a_0}{b_m z^m + b_{m-1} z^{m-1} + \dots + b_0}$, where $D \subset \mathbb{C}$ is a domain which does not contain the roots of the denominator.

It is easy to prove that polynomial and rational functions are holomorphic on their domains of definition and their derivatives are to be computed by the same rules as in the real case. For instance, $(z^n)' = n z^{n-1}$, $\forall n \in \mathbb{Z} \setminus \{0\}$ etc.

4.2 Exponential Function

As we have seen above, the real exponential function can be naturally extended to $\mathbb{C}$ by setting

$$\exp(z) = e^x(\cos y + i \sin y), \quad \forall z = x + iy \in \mathbb{C}.$$

Let us examine the properties of this function.

First, the complex exponential function is *periodic*, having its main period equal to $2\pi i$:

$$\exp(z + 2\pi i) = \exp(z), \quad \forall z \in \mathbb{C}.$$

By a straightforward computation, we can convince ourselves that:

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2), \quad \forall z_1, z_2 \in \mathbb{C}.$$

The exponential function is an entire function, and its derivative is:

$$(\exp(z))' = \exp(z),$$

that is, it is calculated by a similar rule to that one referring to the real exponential function. We can prove the above, by separating real and imaginary parts of $\exp(z)$:

$$u(x, y) = e^x \cos y; \quad v(x, y) = e^x \sin y$$

(which are of class $\mathcal{C}^\infty$ all over $\mathbb{R}^2$), checking the Cauchy-Riemann equations and applying Theorem 4. Its derivative is:

$$(\exp(z))' = u'_x + i v'_x = e^x \cos y + i e^x \sin y = \exp(z).$$

Exercises:

- 1) Calculate $\exp(1 + i)$, $\exp(i\pi)$, $\exp(2i\pi)$.
- 2) Prove that: $\exp(\bar{z}) = \overline{\exp(z)}$, $|\exp(z)| = e^x$ for $z = x + iy \in \mathbb{C}$.
- 3) Find the image through the exponential function of the following sets:
 - a) the line $A = \{z \in \mathbb{C} \mid \operatorname{Re} z = x_0 - \text{const.}\}$;
 - b) the left half-plane $D_1 = \{z \in \mathbb{C} \mid \operatorname{Re} z < 0\}$;
 - c) the right half-plane $D_2 = \{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$.

4.3 Trigonometric Functions

The sinus and cosinus of a complex number are defined by

$$\cos z = \frac{\exp(iz) + \exp(-iz)}{2}, \quad \sin z = \frac{\exp(iz) - \exp(-iz)}{2i}. \quad (13)$$

This definition might seem a little bit strange at first sight. Still, let's examine what it gives for some particular cases.

- For $z = x + 0i \in \mathbb{R}$, we have

$$\begin{aligned} \cos x &= \frac{e^{ix} + e^{-ix}}{2} = \frac{1}{2}(\cos x + i \sin x + \cos x - i \sin x) = (\cos x)_{\mathbb{R}} \\ \sin x &= \frac{e^{ix} - e^{-ix}}{2i} = \frac{1}{2i}(\cos x + i \sin x - \cos x + i \sin x) = (\sin x)_{\mathbb{R}}, \end{aligned}$$

which coincide with the definitions of $\cos$ and $\sin$ as real functions. That is, (13) really extends to $\mathbb{C}$ the functions $\sin$ and $\cos$. So, it makes sense to call (13), $\cos$ and $\sin$. Let us go further and see some other properties of these functions.

- For $z = 0 + iy$ (imaginary number), we get

$$\begin{cases} \cos(iy) = \frac{e^{-y} + e^y}{2} = \cosh(y) \\ \sin(iy) = \frac{1}{2i}(e^{-y} - e^y) = i \sinh(y) \end{cases}, \quad (14)$$

which gives the link between the usual ("circular") sinus and cosinus and their hiperbolic counterparts.

Conversely, we have

$$\cosh(iy) = \cos y; \sinh(iy) = i \sin y. \quad (15)$$

And finally, the newly defined sinus and cosinus obey absolutely similar properties to those defined on $\mathbb{R}$:

1. $\cos^2 z + \sin^2 z = 1, \quad \forall z \in \mathbb{C};$
2. $\cos(z \pm w) = \cos z \cos w \mp \sin z \sin w, \quad \forall z, w \in \mathbb{C};$
3. $\sin(z \pm w) = \sin z \cos w \pm \cos z \sin w, \quad \forall z, w \in \mathbb{C};$
4. $\sin(x + iy) = \sin x \cosh(y) + i \cos x \sinh(y);$
 $\cos(x + iy) = \cos x \cosh(y) - i \sin x \sinh(y), \quad \forall x, y \in \mathbb{R}.$
5. The complex sinus and cosinus are entire functions. Their derivatives are

$$(\cos z)' = -\sin z, \quad (\sin z)' = \cos z,$$

just as in the real case.

Exercises:

1. Prove relation (15) and properties 1-5 above.
2. Show that the equation $\cos z = 2$ has solutions in $\mathbb{C}$.
3. Solve the equation $\sin z = 0$.
4. Prove that $\cos(z + 2k\pi) = \cos z, \sin(z + 2k\pi) = \sin z$ for every $z \in \mathbb{C}$.
5. Show that $|\sin z|^2 = \cos^2 x + \sinh^2(y), \quad \forall z = x + iy \in \mathbb{C}.$

4.4 Complex Logarithm

The real complication begins when talking of complex logarithms. So far, all elementary functions were "well defined" meaning that images $f(z)$ were unique for every z in the domain of f . But let us think of the solutions w of the equation

$$\exp(w) = z,$$

for $z \in \mathbb{C}$ arbitrary and fixed.

Obviously, since the complex exponential is a periodic function, such a solution won't be unique.

Indeed, let $w = u + iv$ and $z = \rho(\cos \theta + i \sin \theta)$. Then

$$e^u(\cos v + i \sin v) = \rho(\cos \theta + i \sin \theta),$$

which yields the solutions $u = \ln \rho = \ln |z|$, $v = \theta + 2k\pi = \text{Arg}(z) + 2k\pi$, $k \in \mathbb{Z}$. We get the complex logarithm $\log z$ as the set of solutions

$$w_k = \log_k z = \ln |z| + i(\text{Arg}(z) + 2k\pi), \quad k \in \mathbb{Z},$$

which makes sense for any $z \in \mathbb{C} \setminus \{0\}$.

As defined above, the complex logarithm is a so-called "multi-valued function" actually, it takes infinitely many distinct values, which differ from each other by a multiple of $2\pi i$. In order to obtain a regular, one-valued function, we can set $k := 0$, thus getting

$$\text{Log}(z) = \ln |z| + i\text{Arg}(z), \quad \forall z \neq 0,$$

where $\text{Arg}(z)$ is the principal argument of z .

The function $\text{Log}(z)$ is called the *principal branch of the logarithm*. The relation between the principal branch and the k -th branch of the logarithm is

$$\log_k(z) = \text{Log}(z) + 2k\pi i.$$

Properties:

1. $\exp(\log_k(z)) = z$, for any $k \in \mathbb{Z}$ and $z \in \mathbb{C}$.
2. Conversely, the value of $\log_k(\exp(z))$ is not necessarily z , since

$$\log_k(\exp(z)) = z + 2k\pi i, \quad k \in \mathbb{Z}.$$

3. The principal logarithm $\text{Log}(z)$ is discontinuous on the real negative axis $\{z_0 \mid z_0 = x + 0i, x < 0\}$. This can be checked by considering the limits

$$\begin{aligned} \lim_{z \rightarrow z_0, z \in S^+} \text{Log}(z) &= \ln |z_0| + i\pi, \\ \lim_{z \rightarrow z_0, z \in S^-} \text{Log}(z) &= \ln |z_0| - i\pi, \end{aligned}$$

where S^+ denotes the upper half-plane $S^+ = \{z \in \mathbb{C} \mid \text{Im} z > 0\}$ and S^- , the negative one $S^- = \{z \in \mathbb{C} \mid \text{Im} z < 0\}$.

Outside the real negative axis, the function Log is continuous. Hence, it makes sense to look for a derivative at those points:

$$\begin{aligned} Log'(z_0) &= \lim_{z \rightarrow z_0} \frac{Log(z) - Log(z_0)}{z - z_0} = \lim_{w \rightarrow w_0} \frac{w - w_0}{\exp(w) - \exp(w_0)} = \\ &= \frac{1}{(\exp w)'|_{w=w_0}} = \frac{1}{\exp(w_0)}. \end{aligned}$$

Thus: Log is holomorphic outside the real negative axis and

$$Log'(z) = \frac{1}{z}.$$

A similar result holds true for any branch $\log_k$, namely:

$$(\log_k(z))' = \frac{1}{z}. \quad (16)$$

Exercises:

1. Calculate: $Log(1+i)$, $\log(i)$, $\log(-1)$.
2. Check whether $Log(zw) = Log(z) + Log(w)$. Find the appropriate formula.
3. Prove relation (16).
4. Determine the images through the principal branch Log of the logarithm of the following sets:

$$4.1 \quad A := \{z \in \mathbb{C} \mid Arg(z) = \theta_0 - const.\};$$

$$4.2 \quad B := \{z \in \mathbb{C} \mid |z| < 1\}.$$

4.5 Complex Powers

They are defined as

$$z^c = \exp(c \log z). \quad (17)$$

Since they involve complex logarithms, complex powers are generally multi-valued. Let us check whether the above definition is consistent with what we know concerning integer and powers and n -th roots of $z \in \mathbb{C}$.

1. For $c = n \in \mathbb{Z}$, we get

$$\begin{aligned} z^n &= \exp(n \log z) = \exp[n \ln |z| + in(Arg(z) + 2k\pi)] = \\ &= |z|^n [\cos(nArg(z) + 2nk\pi) + i \sin(nArg(z) + 2nk\pi)]. \end{aligned}$$

With $\theta = Arg(z)$, $\rho = |z|$, this yields

$$z^n = \rho^n (\cos n\theta + i \sin n\theta),$$

which is nothing but Moivre's formula.

2. By a similar reasoning, rational powers $z^{\frac{p}{q}}$ are given by

$$z^{\frac{p}{q}} = |z|^{\frac{p}{q}} \left[\cos\left(\frac{p}{q}(\phi + 2k\pi)\right) + i \sin\left(\frac{p}{q}(\phi + 2k\pi)\right) \right], \quad \phi = \text{Arg}(z).$$

In particular, for $\frac{p}{q} = \frac{1}{n}$, $n \in \mathbb{N}$, we get the distinct n -th roots of z as

$$z^{\frac{1}{n}} = |z|^{\frac{1}{n}} \left(\cos \frac{\phi + 2k\pi}{n} + i \sin \frac{\phi + 2k\pi}{n} \right), \quad k = 0, \dots, n-1.$$

A special attention is required by the powers of e . Namely, we have for $c \in \mathbb{C}$,

$$e^c = \exp(c \log e) = \exp(c(\text{Log}(e) + 2k\pi i)) = \exp(c \text{Log}(e)) \exp(c 2k\pi i).$$

1. We have $\text{Log}(e) = 1$, hence $\exp(c \text{Log}(e)) = \exp(c)$, but, if c is not an integer number, the second factor is multi-valued. This is why we avoided denoting $e^z = \exp(z)$ - while $\exp(z)$ is single-valued, e^z is not.

Still, it is a common **convention** to denote:

$$e^z \stackrel{\text{notation}}{=} \exp(z) = e^{c \text{Log}(e)}.$$

Unless otherwise stated, by e^z we shall mean from now on, $\exp(z)$.

Exercises:

1. Calculate: $(1-i)^{\frac{1}{3}}$, $(1+i\sqrt{3})^{1+i}$, i^i , 1^i ;
2. Compare the sets of values $\log(i^{\frac{1}{2}})$ and $\frac{1}{2} \log(i)$.

4.6 Branches of Multi-Valued Functions. Riemann Surfaces

Riemann surfaces are generalizations of the complex plane, created in order to transform multi-valued functions such as $z^{\frac{1}{n}}$ or $\log(z)$ into single-valued and continuous functions.

Let $w = f(z)$ define a multi-valued complex function. A *branch* of f is a single-valued function f_0 which is continuous on some domain in D and assigns to each $z \in D$ one of the values of $f(z)$.

One of the most classical examples concern the branches of the n -th root function $w = z^{\frac{1}{n}}$, $n \in \mathbb{N}$:

$$w_k = |z|^{\frac{1}{n}} \left(\cos \frac{\phi + 2k\pi}{n} + i \sin \frac{\phi + 2k\pi}{n} \right), \quad k = 0, \dots, n-1.$$

Let us take $n = 2$. We get two branches defined over $\mathbb{C}$ of the function $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = z^{\frac{1}{2}}$:

$$f_1(z) = \sqrt{|z|} e^{i \frac{\text{Arg}(z)}{2}}, \quad f_2(z) = \sqrt{|z|} e^{i(\frac{\text{Arg}(z)}{2} + \pi)}.$$

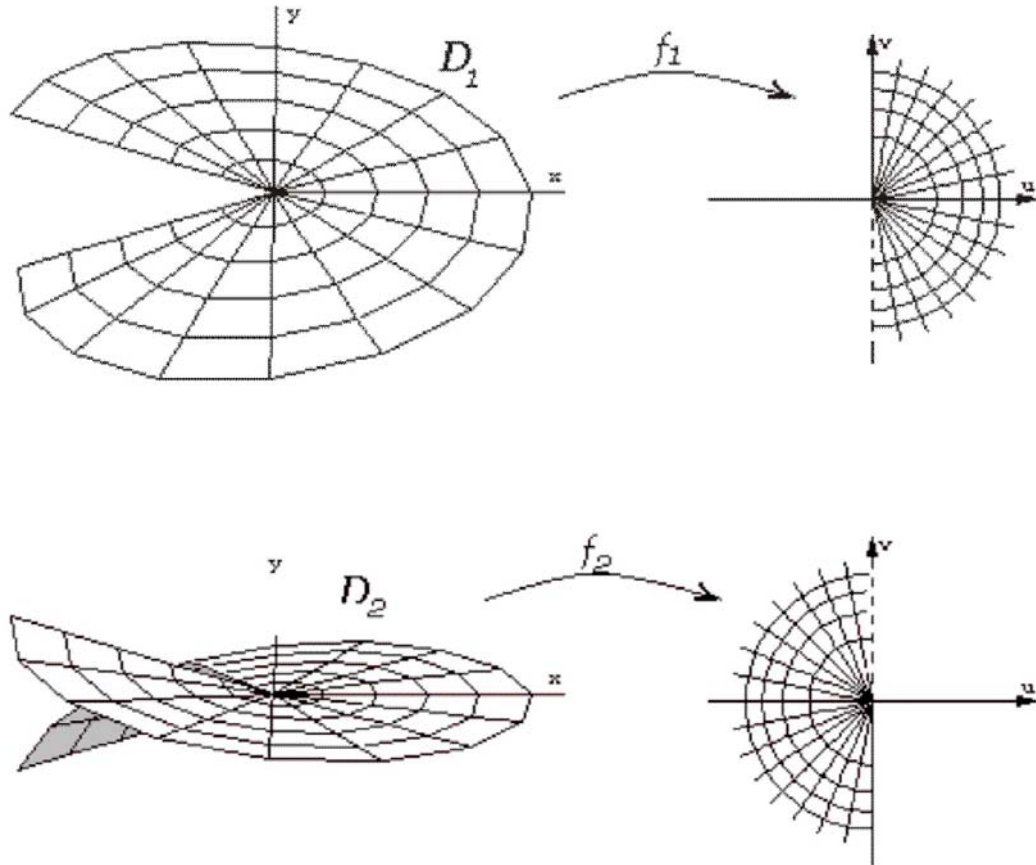
It is easy to prove (by taking limits over the positive and negative half-planes respectively) that f_1 and f_2 are discontinuous along the real negative axis $\text{Arg}(z) = \pi$ and continuous all over outside this set. Let us perform the following geometric construction.

We represent the image of $\mathbb{C}$ by f_1 and f_2 respectively (in the picture, we have just represented the closed unit disk $|z| \leq 1$).

Let us denote by D_1 and D_2 the domains for f_1 and f_2 (which are identical copies of the unit disk $|z| \leq 1$ - in the following, ρ ranges from 0 to 1):

- $D_1 = \{z = \rho(\cos \theta + i \sin \theta) \in \mathbb{C} \mid \theta \in (-\pi, \pi]\} \Rightarrow f_1(D_1) = \{w \mid \text{Arg}(w) \in (-\frac{\pi}{2}, \frac{\pi}{2}]\}$; hence, the image of D_1 by f_1 is the right half-plane and the positive Oy-axis.
- $D_2 = \{z = \rho(\cos \theta + i \sin \theta) \in \mathbb{C} \mid \theta \in (-\pi, \pi]\} \Rightarrow f_2(D_2) = \{w \mid \text{Arg}(w) \in (\frac{\pi}{2}, \pi] \cup (-\pi, -\frac{\pi}{2}]\}$. The image of D_2 through f_2 is the left half-plane together with the negative Oy-axis.

Then, let us "cut" the disks along the negative Ox-axis, as in the following pictures:



Now, let us stick together D_1 and D_2 along the cut, this way: the upper edge of D_1 is stuck to the lower edge of D_2 and the lower edge of D_1 , to the upper edge of D_2

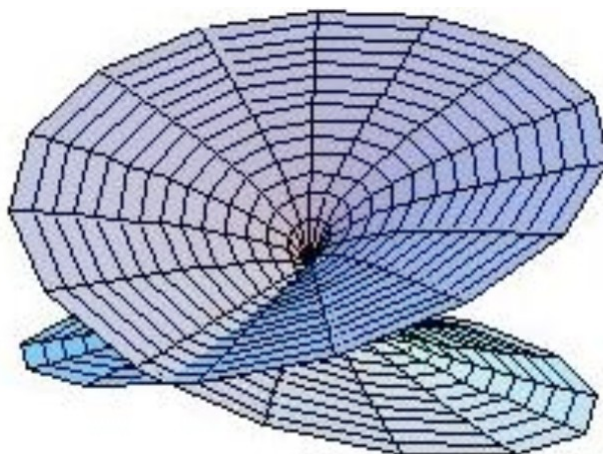
Then, the function

$$f : D_1 \cup D_2, f(z) = \begin{cases} f_1(z) & \text{if } z \in D_1 \\ f_2(z) & \text{if } z \in D_2 \end{cases}$$

is continuous on its domain.

The image is the following: as $\text{Arg}(z)$ reaches $\frac{\pi}{2}$, the value $f_1(z)$ "passes into" $f_2(z)$, and as $\text{Arg}(z)$ grows to $-\frac{\pi}{2}$, the value $f_2(z)$ "passes into" $f_1(z)$.

The obtained surface is called the *Riemann surface* attached to the 2-valued function $z^{\frac{1}{2}}$.

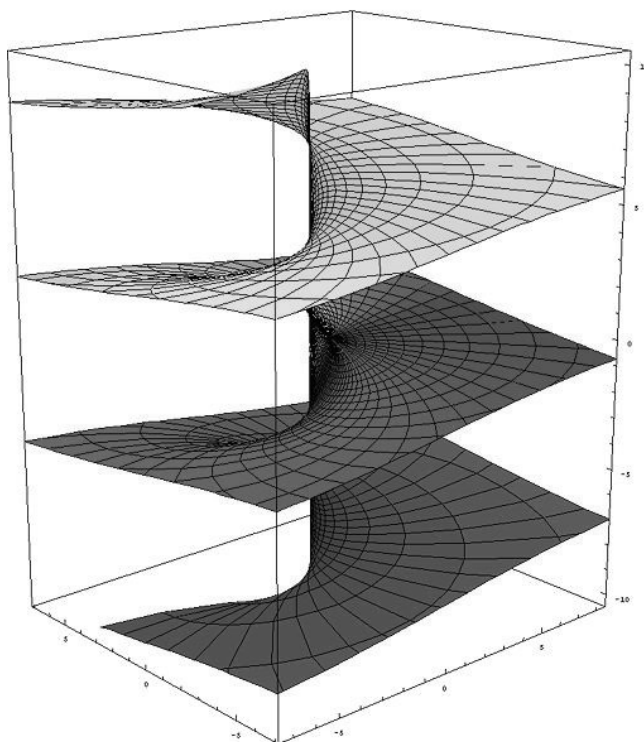


As a remark, we can perform the cut along any half-line $\text{Arg}(z) = \alpha$, the construction is similar.

For the logarithm $\log z$, since it has infinitely many values, we need to stick together infinitely many copies of $\mathbb{C}$, cut along the negative real axis: $D_k = \{z = \rho(\cos \theta + i \sin \theta) \in \mathbb{C} \mid \theta \in (-\pi, \pi]\}, k \in \mathbb{Z}$. We get a function:

$$f : \bigcup_{k \in \mathbb{Z}} D_k \rightarrow \mathbb{C}, f(z) = \log_k(z) \text{ for } z \in D_k.$$

Intuitively, when $\theta = \text{Arg}(z)$ grows to π , the branch $\log_k(z)$ "transform into" the next one $\log_{k+1}(z)$. The situation is different from the above one, since no branch $\log_k(z)$ would ever transform back into one of the previous branches.



5 Conformal Transformations

5.1 Definition and Properties

Conformal mappings, together with holomorphic ones, are the most important classes of complex functions from the point of view of applications.

Let $D, \Delta \subset \mathbb{R}$ be two domains. We say that $f : D \rightarrow \Delta$ is a *conformal map* (or that Δ is a *conformal image* of D) if:

1. f is bijective;
2. f and f^{-1} are continuous on D and Δ respectively;
3. if C_1 and C_2 are $\mathcal{C}^1$ -smooth curves in D , $z_0 \in C_1 \cap C_2$ is nonsingular for both C_1 and C_2 , then the angle between C_1 and C_2 (measured at z_0) is the same as the angle between the images $f(C_1)$ and $f(C_2)$ (measured at $f(z_0)$). In other words, f is *angle-preserving*.

For holomorphic maps, there exists a nice criterion of conformality:

Theorem 9 (of characterization): Let $D \subset \mathbb{C}$ be a domain and $f : D \rightarrow \Delta$, with $\Delta = f(D)$. If f is holomorphic and injective in D then f is a conformal map.

Remark 10 If $f = u + iv$ is holomorphic and injective in D , then its derivative $f'(z)$ is nonzero all over D . This is easy to check, having in mind that

$$\frac{D(u, v)}{D(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \stackrel{C=R}{=} \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 = |f'(z)| \neq 0.$$

Let us now prove the theorem

Proof. 1) f bijective - obvious.

2) Since f is holomorphic, it follows, on one side, that it is continuous in D . On the other side, it also follows that the inverse f^{-1} is holomorphic, and hence it is continuous, too.

3) *angle-preserving property*: Let $(C_1) : z = \gamma_1(t)$ and let point z_0 correspond to the value $t_1 : z_0 = \gamma_1(t_1)$. Its tangent vector at z_0 is then $\gamma'_1(t_1) =: r_1 e^{i\phi_1}$ (since z_0 is nonsingular for (C_1) , the tangent vector $\gamma'_1(t_1)$ does not vanish, hence the argument $\phi_1 = \text{Arg}(\gamma'_1(t_1))$ is defined).

Similarly, let (C_2) be described by $(C_2) : z = \gamma_2(\tilde{t})$. If z_0 corresponds to the value $\tilde{t} = t_2$ of the parameter, then the tangent vector to (C_2) at z_0 is of the form $\gamma'_2(t_2) =: r_2 e^{i(\phi_1 + \theta)}$, where θ is the angle between the two tangents.

Let us calculate now the angle between the images $f(C_1)$ and $f(C_2)$.

Let $f'(z_0) = \rho e^{i\alpha}$. Then the tangent vector to $f(C_1) : Z = \Gamma_1(t)$ at $f(z_0)$ is

$$\Gamma'_1(z_0) = f'(z_0)\gamma'_1(t_1) = \rho r_1 e^{i(\alpha + \phi_1)}.$$

Similarly, the tangent vector to $f(C_2) : Z = \Gamma_2(\tilde{t})$ at the same point is

$$\Gamma'_2(z_0) = f'(z_0)\gamma'_2(t_2) = \rho r_2 e^{i(\alpha + \theta + \phi_1)}.$$

The angle between the two tangent vectors is given by the difference of the arguments

$$\widehat{(f(C_1), f(C_2))} = \alpha + \theta + \phi_1 - \alpha - \phi_1 = \theta,$$

which is the same as the angle between C_1 and C_2 . ■

The most spectacular and important result concerning conformal mappings is *Riemann's Theorem*:

Theorem 11 (Riemann): *If U is a simply connected domain strictly contained in $\mathbb{C}$, then there exists a holomorphic conformal mapping from U to the unit open disk $D = \{z \in \mathbb{C} \mid |z| < 1\}$.*

Proof sketch: What we shall do in the following is a sketch of a proof - a complete and rigorous proof requires solving a Dirichlet problem for the disk, which will be presented later. We shall, still, assume that it is known that a Dirichlet problem for the disk has a unique solution.

Let us build a conformal mapping $f : U \rightarrow D$ such that:

- $f(z_0) = 0$, where z_0 is an arbitrary and fixed point of U and
- f maps the boundary of U into the boundary of $D : f(\partial U) = \partial D$.

In order to fulfill the first requirement, let us build f such that it contains $z - z_0$ as a factor:

$$f(z) := (z - z_0) \exp(g(z)), \quad (18)$$

where the holomorphic function $g = g(z)$ is to be determined. The notation $\exp(g(z))$ (instead of simply $g(z)$) will be transparent several lines below.

Let us denote $g(z) := u(x, y) + iv(x, y)$ (where $z = x + iy$). Then, $|\exp(g(z))| = e^{u(x, y)}$. The request $f(\partial U) = \partial D$ is equivalent to

$$|f(z)| = 1, \quad \forall z \in \partial U.$$

Passing to modules in (18) yields then

$$1 = |z - z_0| e^u,$$

which leads immediately to

$$u = \operatorname{Re}(g(z)) = -\ln |z - z_0|.$$

What we got is a Dirichlet problem on the unit disk: since u is the real part of a holomorphic function, it is harmonic. Consequently, knowing its values on the boundary ∂U , we can uniquely determine its values inside U (as we shall prove in the section devoted to applications of the Cauchy formula).

Thus, we got a harmonic function $u : U \rightarrow \mathbb{R}$. Taking its harmonic conjugate v , we get a holomorphic function $g := u + iv : U \rightarrow \mathbb{C}$, which obeys (18), q.e.d.

An important property of conformal transformations (which will be used in the followong) is the fact that they preserve harmonicity of functions.

Theorem 12 *Let $\phi : G \rightarrow \mathbb{R}$, $(u, v) \mapsto \phi(u, v)$ be a harmonic function in the simply connected domain $G \subset \mathbb{R}^2 \equiv \mathbb{C}$, and $f : D \rightarrow G$, $z = x + iy \mapsto u(x, y) + iv(x, y)$ (where $D \subset \mathbb{C}$ is another domain) be a holomorphic conformal mapping. Then, the composition $g = \phi \circ f : (x, y) \in D \mapsto \phi(u(x, y), v(x, y)) \in \mathbb{R}$ is harmonic in D .*

Proof. Let ϕ be harmonic in G . Then there exists a holomorphic function $\Phi : G \rightarrow \mathbb{C}$ such that $\phi = \operatorname{Re}\Phi$.

We have $g = \phi \circ f = \operatorname{Re}(\Phi \circ f)$ and, since $\Phi \circ f$ is holomorphic, we obtain that g is holomorphic, too. ■

5.2 Mobius Transformations

An important class of conformal transformations is provided by *Mobius transformations*. They are defined by

$$f(z) = \frac{az + b}{cz + d}. \quad (19)$$

We can immediately see that:

- If $c = 0$ (and $d \neq 0$) then we can define $f : \mathbb{C} \rightarrow \mathbb{C}$ and it is an entire function.
- If $c \neq 0$, then f is holomorphic on $\mathbb{C} \setminus \{-\frac{d}{c}\}$.

In the following, we shall take into account only the case when $c \neq 0$ (since $c = 0$ provides polynomial functions, which have already been discussed). That is, we consider

$$f : \mathbb{C} \setminus \{-\frac{d}{c}\} \rightarrow \mathbb{C} \setminus \{\frac{a}{c}\}. \quad (20)$$

Let us see some properties of functions (19), (20).

Proposition 13 *If $ad - bc \neq 0$, then functions*

$$f : \mathbb{C} \setminus \{-\frac{d}{c}\} \rightarrow \mathbb{C} \setminus \{\frac{a}{c}\}, \quad f(z) = \frac{az + b}{cz + d}$$

are bijective.

Proof. Let us suppose $f(z_1) = f(z_2)$, which is, $\frac{az_1 + b}{cz_1 + d} = \frac{az_2 + b}{cz_2 + d}$. By equating “crosswise” products, we are led to $(bc - ad)(z_2 - z_1) = 0$. By hypothesis, $bc - ad$ does not vanish. Consequently, $z_1 = z_2$ and f is injective.

Further, the equation $u = \frac{az + b}{cz + d}$ leads to $z = \frac{b - du}{-a + cu}$, which is defined for any $u \neq \frac{a}{c}$, hence f is also surjective. ■

An immediate (but not unimportant) consequence of the above and of the holomorphy of Mobius transformations is

Proposition 14 *Mobius transformations (19), (20) are conformal mappings.*

Proposition 15 *Any Mobius transformation can be written as a composition of translations, homotheties, reflections into axes and inversions.*

Proof. Let us apply a little algebraic trick, namely, add and subtract an $\frac{a}{c}$ in the expression (19). We have

$$f(z) = \left(\frac{az + b}{cz + d} - \frac{a}{c}\right) + \frac{a}{c} = \frac{a}{c} + \frac{acz + bc - acz - ad}{c(cz + d)} = \frac{a}{c} + \frac{bc - ad}{c(cz + d)}.$$

That is, we can write f as

$$f(z) = \frac{a}{c} + \frac{bc - ad}{c^2} \frac{1}{z + \frac{d}{c}}.$$

Let us see what we obtained. The mapping $z \xrightarrow{T_1} z_1 = z + \frac{d}{c}$ is a translation. Further, $z_1 \xrightarrow{T_2} z_2 := \frac{1}{z_1} (= \overline{\frac{1}{\bar{z}_1}})$ is an inversion composed with a reflection into Ox . Multiplying $\frac{bc - ad}{c^2}$ to the result provides a homothety, and adding $\frac{a}{c}$ - a translation. ■

The above property has an important geometric consequence:

Theorem 16 *If f is a Mobius transformation, then f maps lines into lines or circles, and circles into lines or circles.*

Proof. The equation of a circle or of a line can be written as

$$m(x^2 + y^2) + nx + py + q = 0, \quad (21)$$

$m, n, p, q \in \mathbb{R}$ (where $m = 0$ provides lines and $m \neq 0$, circles).

Taking into account the above Proposition, it is enough to prove the statement for the inversion

$$f(z) = \frac{1}{\bar{z}},$$

(since it is well-known that translations, homotheties and reflections into axes map lines into lines and circles into circles).

Let $w = \frac{1}{\bar{z}}$. Then, $z = \frac{1}{\bar{w}}$ can be written as

$$z = \frac{w}{|w|^2}.$$

Let $z = x + iy$ and $w = u + iv$, $x, y, u, v \in \mathbb{R}$. The above equality leads to

$$x = \frac{u}{u^2 + v^2}, \quad y = \frac{v}{u^2 + v^2}.$$

Also, we have $x^2 + y^2 = |z|^2 = \frac{1}{|w|^2} = \frac{1}{u^2 + v^2}$. Replacing into (21), we get

$$\frac{m}{u^2 + v^2} + \frac{nu + pv}{u^2 + v^2} + q = 0,$$

or

$$q(u^2 + v^2) + nu + pv + m = 0. \quad (22)$$

The above also represents a circle or a line. More precisely:

- A line through the origin ($m = 0, q = 0$) transforms into a line which passes through the origin.
- A line which does not contain the origin ($m = 0, q \neq 0$) transforms into a circle which passes through the origin.
- A circle through $O(0, 0)$ ($m \neq 0, q = 0$) is transformed into a line which does not contain O .

- A circle which does not pass through O ($m \neq 0, q \neq 0$) transforms into a circle which does not pass through O .

■

Actually, the above provides an elegant proof of the properties of the inversion.

Let us now extend the Mobius transform $f(z) = \frac{az+b}{cz+d}$ in (19), (20) to $\bar{\mathbb{C}}$, by setting

$$f(-\frac{d}{c}) = \infty, \quad f(\infty) = \frac{a}{c}.$$

The above defined function $f : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ remains bijective.

Definition 17 *If $z, z_1, z_2, z_3 \in \bar{\mathbb{C}}$ with z_1, z_2, z_3 distinct, then the quantity*

$$[z, z_1, z_2, z_3] = \frac{z - z_1}{z - z_3} : \frac{z_2 - z_1}{z_2 - z_3}$$

is called the cross ratio of z, z_1, z_2, z_3 .

The cross ratio is a Mobius transformation with the properties

$$f(z_1) = 0, \quad f(z_2) = 1, \quad f(z_3) = \infty. \quad (23)$$

Moreover, it is the unique Mobius transformation which obeys (23). In order to prove the uniqueness, let us suppose that

$$h(z) = \frac{az+b}{cz+d}$$

is another Mobius transformation with these properties. Then, from $h(z_1) = 0$, we get $b = -az_1$, and hence $az+b = a(z-z_1)$. Also, from $h(z_3) = \infty$, we have $cz_3+d=0$, which yields $d = -cz_3$ and consequently $cz+d = c(z-z_3)$. We have so far

$$h(z) = \frac{a(z-z_1)}{c(z-z_3)}.$$

By using $h(z_2) = 1$, we also obtain that $\frac{a}{c} = \frac{z_2-z_3}{z_2-z_1}$ which proves the statement.

Further, taking into account bijectivity, we can build Mobius transformations which map any three distinct points in $\bar{\mathbb{C}}$ into any three distinct points in $\bar{\mathbb{C}}$. Namely, if we want to map z_1, z_2, z_3 into w_1, w_2, w_3 respectively, we build $f_1 : \{z_i\} \rightarrow \{0, 1, \infty\}$, $f_2 : \{w_i\} \rightarrow \{0, 1, \infty\}$; then, the composition $f_2^{-1} \circ f_1$ is a Mobius transformation which maps z_1, z_2, z_3 into w_1, w_2, w_3 respectively. The uniqueness is immediate. Thus, we got an important result:

Theorem 18 *If z_1, z_2, z_3 three distinct points in $\bar{\mathbb{C}}$ and $w_1, w_2, w_3 \in \bar{\mathbb{C}}$ are also distinct, then there uniquely exists a Mobius transformation $f : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}$ with*

$$f(z_1) = w_1, \quad f(z_2) = w_2, \quad f(z_3) = w_3.$$

6 Plane Vector Fields. Complex Potential

6.1 Plane Vector Fields

Let us consider functions of type

$$\vec{F}(x, y) = M(x, y) \vec{i} + N(x, y) \vec{j}$$

defined on some domain $D \subset \mathbb{R}^2$. To each pair of values (x_0, y_0) for which both M and N are defined, such a function assigns a vector $\vec{F}(x_0, y_0)$ in the plane. $\vec{F}$ is therefore called a *vector function* of two variables.

To visualize the function $F(x, y)$, at each point (x_0, y_0) in the domain we place the corresponding vector $\vec{F}(x_0, y_0)$ such that its tail is at (x_0, y_0) . Thus each point of the domain is the tail end of a vector, and what we get is called a *plane vector field*. This vector field gives a picture of the vector function $F(x, y)$.

Conversely, given a vector field in a region of the xy -plane, it determines a vector function as above, by expressing each vector of the field in terms of its $\vec{i}$ and $\vec{j}$ components. We will use the same symbol $\vec{F}$ to denote both the field and the function, saying "the vector field $\vec{F}$ ".

Basically, a vector field **in the Euclidean 3D-space** depends on four coordinates: time and the coordinates of the point x, y, z :

$$\vec{F} = \vec{F}(t, x, y, z) = F_x(t, x, y, z) \vec{i} + F_y(t, x, y, z) \vec{j} + F_z(t, x, y, z) \vec{k}.$$

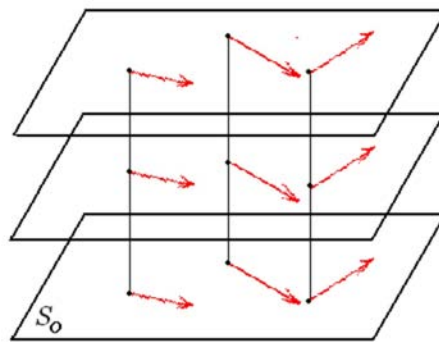
Physically, **plane vector fields** defined above correspond to the following situation:

- they are *stationary* (do not change with time):

$$\vec{F} = \vec{F}(x, y, z);$$

- they are *plane-parallel*: all the vectors of the field are parallel to the same plane S and moreover, they have constant values along all perpendicular lines to S . Thus, by considering $S = xOy$, we get

$$\vec{F} = \vec{F}(x, y) \Rightarrow F = F_x + iF_y.$$



a plane-parallel vector field

Circulation is the amount of force that pushes along a closed boundary or path. It's the total "push" you get when going along a closed path. Rigorously speaking,

$$\text{Circulation} = \oint_C \vec{F} \cdot d\vec{r}$$

Curl is the circulation per unit area; if we are to evaluate curl at a point, we make the area tend to 0 ("shrink" it around the point):

$$rot \vec{F} = \lim_{\sigma \rightarrow 0} \frac{1}{\sigma} \oint_{\partial D} \vec{F} \cdot d\vec{r}.$$

where $\sigma = area(D)$. The above gives

$$rot(\vec{F}) = \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \vec{k}.$$

The $rot(\vec{F})$ measures the tendency of $\vec{F}$ to rotate at a given point.

- If

$$rot(\vec{F}) = 0,$$

then $\vec{F}$ is called *irrotational*. In this case, $\vec{F}$ is a *conservative (or gradient)* vector field, which means that there exists a function $\phi : D \rightarrow \mathbb{R}$, whose gradient is $\vec{F}$

$$\vec{F} = \nabla \phi,$$

Flux (of a plane vector field) is the amount of “something” (electric field, water, heat etc.) passing through the border of a domain.

$$\text{Flux} = \oint_{\partial D} (\vec{F} \cdot \vec{n}) ds$$

Divergence is the amount of flux entering or leaving a **point** - that is, it measures the magnitude of a source or sink at a given point of $\vec{F}$. It is given by

$$div \vec{F} = \lim_{\sigma \rightarrow 0} \frac{1}{\sigma} \oint_{\partial D} (\vec{F} \cdot \vec{n}) ds,$$

where $\vec{n}$ is the normal of the contour ∂D . This is:

$$div \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y}.$$

If

$$div \vec{F} = 0 \text{ on } D,$$

the field is called *solenoidal* on D .

6.2 Complex Potential

Let $\vec{F} = F_x(x, y) \vec{i} + F_y(x, y) \vec{j}$ be a plane vector field over some simply connected domain $D \subset \mathbb{R}^2$. The method of complex potentials consists in identifying a complex function $\Omega(z) = \phi(x, y) + i\psi(x, y)$ (called the *complex potential*) which contains essential information on the behavior of the vector field $\vec{F}$. Namely:

- at points where both $\text{div} \vec{F}$ and $\text{rot} \vec{F}$ vanish, the complex potential Ω is holomorphic. Eventual singularities of Ω indicate "special" points such as sources, sinks, or vortices;

- Lines $\psi = \text{const.}$ are the lines of flux (streamlines, lines of force) of the vector field, while $\phi = \text{const.}$ are its *equipotential lines*.

- If $\vec{F}$ is *solenoidal*,

$$\text{div} \vec{F} = 0 \quad \text{on } D,$$

then $\frac{\partial F_x}{\partial x} = -\frac{\partial F_y}{\partial y}$, which means that there exists a function $\psi : D \rightarrow \mathbb{R}$ such that

$$F_x dy - F_y dx = d\psi \Rightarrow \psi = \int_{z_0}^z -F_y dx + F_x dy,$$

where the integral is independent of the path joining z_0 and z . This function ψ has the property

$$\frac{\partial \psi}{\partial x} = -F_y, \quad \frac{\partial \psi}{\partial y} = F_x \quad (24)$$

- If $\vec{F}$ is *irrotational*,

$$\text{rot}(\vec{F}) = 0,$$

then there exists a function $\phi : D \rightarrow \mathbb{R}$, whose gradient is $\vec{F}$

$$\vec{F} = \nabla \phi,$$

equivalently:

$$\frac{\partial \phi}{\partial x} = F_x, \quad \frac{\partial \phi}{\partial y} = F_y. \quad (25)$$

More precisely, we have $d\phi = F_x dx + F_y dy$, which is,

$$\phi = \int_{z_0}^z F_x dx + F_y dy.$$

Briefly, relations (24) and (25) are actually

$$\begin{cases} \frac{\partial \phi}{\partial x} = F_x = \frac{\partial \psi}{\partial y} \\ \frac{\partial \phi}{\partial y} = F_y = -\frac{\partial \psi}{\partial x} \end{cases} \quad .$$

which is, nothing but the Cauchy-Riemann conditions for

$$\Omega(z) = \phi(x, y) + i\psi(x, y). \quad (26)$$

That is, for $\vec{F}$ - both irrotational and solenoidal, we have built a **holomorphic** function $\Omega : D \rightarrow \mathbb{C}$, called the *complex potential* of the field $\vec{F}$. Its real part

$$\phi = \operatorname{Re}(\Omega)$$

is called the *velocity potential*, while

$$\psi = \operatorname{Im}(\Omega)$$

is called the *stream function* of $\vec{F}$.

The lines

$$\psi = \text{constant}$$

are the *field lines* of $\vec{F}$, since their tangent vectors are colinear, at each point, to the vectors $\vec{F}(x, y)$:

$$m_{tg}(\psi = \text{const.}) = \frac{-\frac{\partial \psi}{\partial x}}{\frac{\partial \psi}{\partial y}} = \frac{F_y}{F_x} = m(\vec{F}).$$

The lines

$$\phi = \text{constant}$$

are called *equipotential lines*.

There holds:

Proposition 19 *Equipotential lines intersect streamlines at right angles.*

Proof. At an arbitrary and fixed point (x, y) in plane, the slope of the tangent to the streamline is $m_\psi = \frac{F_y}{F_x}$, while the slope of the tangent to the equipotential line $\phi = \text{constant}$ is $m_\phi = -\frac{\phi'_x}{\phi'_y} = -\frac{F'_x}{F'_y}$, and the product of the two slopes is constantly equal to -1 . ■

The derivative of the complex potential Ω is

$$w = \Omega' = \frac{\partial \phi}{\partial x} + i \frac{\partial \psi}{\partial x}$$

which is,

$$w = F_x - iF_y. \quad (27)$$

The function $w = w(z)$ (defined over D) is called the *complex velocity* attached to $\vec{F}$. Actually, w is the conjugate of the complex number $F = F_x + iF_y$ attached to $\vec{F}$:

$$w = \bar{F}.$$

Exercises:

1. Find the field lines, the equipotential lines and the complex velocity for the complex potentials:

$$a) \Omega_1(z) = z^2, \quad b) \Omega_2(z) = \frac{1}{z^2}.$$

2. The equipotential lines of some plane field are $x^3 - 3xy^2 = a$, $a > 0$. Find: the potential function, the stream function and the flow lines of the field.

6.3 Applications

6.3.1 In Electrostatics

- **For the electric field** $\vec{E}(E_x, E_y)$, it can be proven that:

- $\vec{E}$ is always irrotational: $\text{rot}(\vec{E}) = 0$.

1. - If, at a point $z \equiv (x, y)$, the density of charge is 0, then it is also solenoidal: $\text{div}(\vec{E}) = 0$.
In this case, the *potential function (electrostatic potential)*, traditionally taken with a minus sign (in comparison to the above said), is ϕ , with

$$-E_x = \frac{\partial \phi}{\partial x}, \quad -E_y = \frac{\partial \phi}{\partial y} \quad (\Leftrightarrow \vec{E} = -\nabla \phi).$$

Its harmonic conjugate (the "*stream function*") is in this case ψ , given by

$$-E_x = \frac{\partial \psi}{\partial y}, \quad E_y = \frac{\partial \psi}{\partial x},$$

and the complex potential Ω is defined as

$$\Omega(z) = \phi(x, y) + i\psi(x, y).$$

The "complex velocity" attached to Ω is

$$w = \Omega'(z) = \frac{\partial \phi}{\partial x} + i \frac{\partial \psi}{\partial x} = -E_x + iE_y = -\bar{E}.$$

This is, F is holomorphic and we have

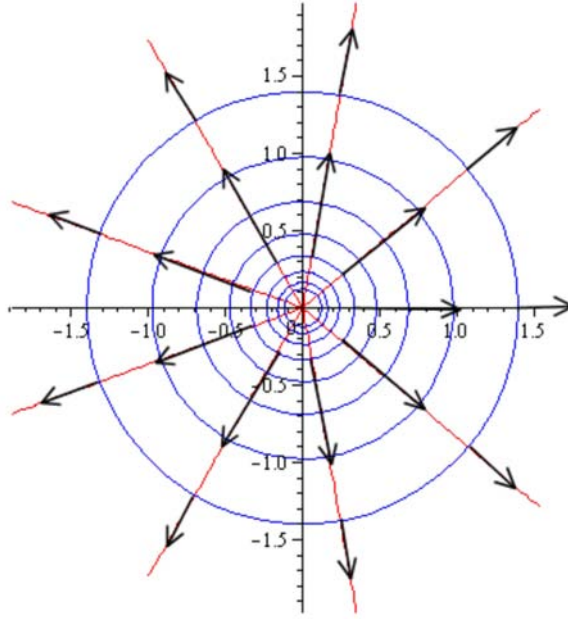
$$E = -\overline{\Omega'(z)}.$$

- Now, let us determine the **electric field attached to a punctual charge** q situated at O :

$$\vec{E} = \frac{2qx}{x^2 + y^2} \vec{i} + \frac{2qy}{x^2 + y^2} \vec{j}$$

(where q is a constant); we attach to $\vec{E}$ the complex number

$$E = \frac{2qx}{x^2 + y^2} + i \frac{2qy}{x^2 + y^2} = \frac{2q}{r} \frac{z}{|z|} \Rightarrow E = \frac{2qz}{|z|^2}.$$



We get:

$$\begin{aligned}\phi &= -\int_{z_0}^z E_x dx + E_y dy = -2q \int_{z_0}^z \frac{x dx + y dy}{x^2 + y^2} = -q \ln(x^2 + y^2) + c = \\ &= 2q \ln\left(\frac{1}{|z|}\right) + c.\end{aligned}$$

and the stream function ψ is

$$\begin{aligned}\psi &= \int_{z_0}^z E_y dx - E_x dy = 2q \int_{z_0}^z \frac{y dx - x dy}{x^2 + y^2} = -2q \cdot \text{Arc tan}\left(\frac{y}{x}\right) + c = \\ &= 2q \cdot \text{Arg}\left(\frac{1}{z}\right) + c = 2q \arg \frac{1}{z} \pm 2n\pi q + c,\end{aligned}$$

where the integral was taken on an arbitrary path from $z_0 = 1$ to z , not containing the origin O .

Remark: ψ above is a multi-valued function (which is explainable, if we take into account the fact that $\mathbb{C} \setminus \{0\}$ is not simply connected!).

The complex potential Ω is

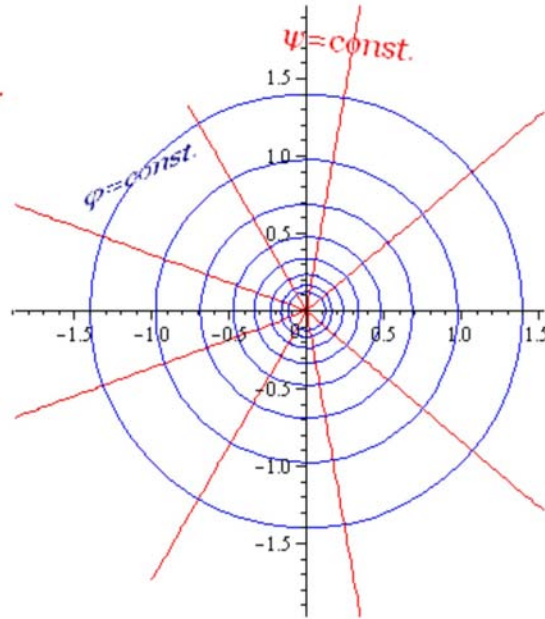
$$\Omega = \phi + i\psi = 2q\left(\ln\left(\frac{1}{|z|}\right) + \text{Arg}\left(\frac{1}{z}\right)\right) + c \equiv 2qLn\frac{1}{z}$$

(where we have taken, for commodity, $c = 0$). We notice that the origin O (the point where the charge is situated) is a singular point for Ω .

Equipotential lines $\phi = K_1$ are circles with the center at O

$$|z| = \text{const.},$$

while the field lines (*lines of flux*) $\psi = K_2$ are lines through O :



- For the field generated by n punctual charges $q_1, \dots, q_n$ situated at the points $z_1, z_2, \dots, z_n$ respectively:

$$E = \sum_{k=1}^n \frac{2q(z - z_k)}{|z - z_k|^2}.$$

In particular:

- For a system of **two charges of different signs** $+q$ and $-q$, $E = 2q\left(\frac{z - z_1}{|z - z_1|^2} - \frac{z - z_2}{|z - z_2|^2}\right)$, which leads to the complex potential

$$\Omega = -2q(\ln(z - z_1) - \ln(z - z_2)) = 2q \ln \frac{z - z_2}{z - z_1}.$$

Equipotential lines $\phi = \operatorname{Re}(\Omega) = \operatorname{const.}$ are

$$\left| \frac{z - z_2}{z - z_1} \right| = \operatorname{const.}$$

(Apollonius' circles w.r.t. z_1 and z_2).

The lines of flux $\psi = \operatorname{const.}$ are

$$\operatorname{Im}(F) = 2q \operatorname{Arg} \left(\frac{z - z_2}{z - z_1} \right) = \operatorname{const.}$$

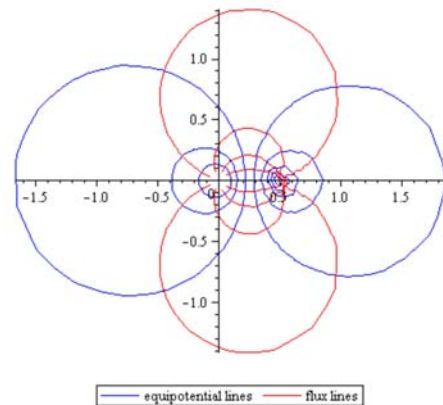
- which are circles through z_1 and z_2 .

- For a system of **two charges of the same sign and size** q :

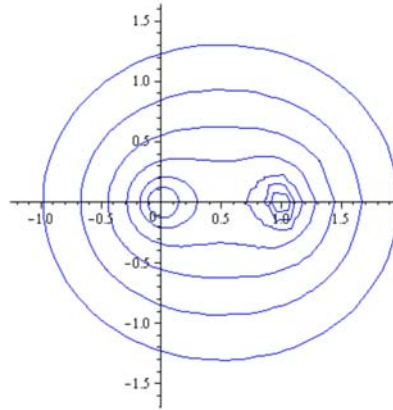
$$E = 2q\left(\frac{z - z_1}{|z - z_1|^2} + \frac{z - z_2}{|z - z_2|^2}\right) \Rightarrow F(z) = -2q \ln((z - z_1)(z - z_2))$$

Equipotential lines $\phi = \operatorname{Re}(\Omega) = \operatorname{const.}$ are

$$|(z - z_1)(z - z_2)| = c = \operatorname{const.}$$



(a *lemniscate*).



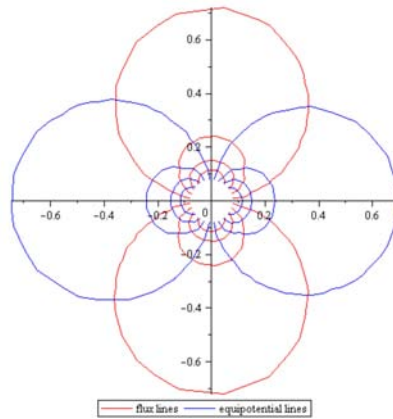
- **Dipoles:** For charges $\pm q$ situated at $z_1 = 0$ and $z_2 = -h$, we have, as above, the complex potential

$$-2q \operatorname{Ln} \left(\frac{z+h}{z} \right) = -2p \operatorname{Ln} \left(\frac{z+h}{z} \right)^{1/h},$$

with $p = qh$. By taking the limit with $h \rightarrow 0$ (and choosing q such that $p = \text{const}$), we get

$$F(z) = -2p \lim_{h \rightarrow 0} \operatorname{Ln} \left(\frac{z+h}{z} \right)^{1/h} = -2p \operatorname{Ln}(e^{1/z}) = -2p \frac{1}{z}.$$

Equipotential lines $\phi = \operatorname{Re}(\Omega) = \text{const}$ are circles through $z = 0$ (with the centers on Ox), while lines of flux $\psi = \operatorname{Im}(\Omega) = \text{const}$ are also circles through $z = 0$, but having their centers on Oy .



6.3.2 In Fluid Mechanics

The condition of absence from a domain D of sources/sinks and vortices is

$$\operatorname{div} \vec{V} = 0, \operatorname{rot} \vec{V} = 0,$$

which leads to the complex potential

$$\Omega = \phi(x, y) + i\psi(x, y),$$

where $\vec{V} = \nabla\phi$, this is, we have the *potential function*

$$\phi = \int_{z_0}^z V_x dx + V_y dy$$

and the *stream function*

$$\psi = \int_{z_0}^z -V_y dx + V_x dy.$$

The function ψ helps to express the *flux* of the fluid through some curve $C \subset D$:

$$(\vec{V}, \vec{n})ds = d\psi.$$

Hence, the flux is

$$Q = \int_C (\vec{V}, \vec{n})ds = \int_C d\psi = \psi(x_2, y_2) - \psi(x_1, y_1).$$

If the domain is not simply connected, then both ϕ and ψ can be multi-valued functions.

The complex velocity is $\mathbb{V} = \Omega'(z) = \frac{\partial\phi}{\partial x} + i\frac{\partial\psi}{\partial x} = V_x - iV_y$, which is,

$$\mathbb{V} = \bar{V}.$$

Examples:

- The velocity vector field $\vec{V}$ of a **fluid whose movement is constant** in direction and size:

$$\vec{V} = V_x \vec{i} + V_y \vec{j}.$$

(where V_x, V_y - constant).

1. In this case, we get, by integration,

$$\begin{aligned}\phi &= V_x x + V_y y + c_1 \\ \psi &= -V_y x + V_x y + c_2.\end{aligned}$$

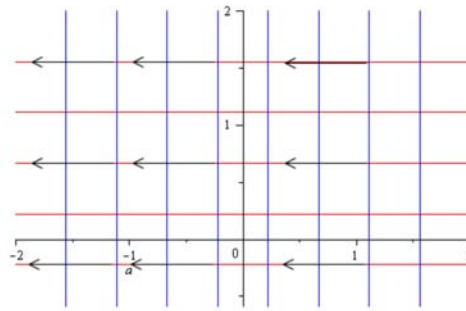
The complex potential is

$$\Omega = \phi + i\psi = V_x x + V_y y + i(-V_y x + V_x y) = x(V_x - iV_y) + iy(V_x - iV_y) = \bar{V}z.$$

A **source** or to a **sink** is characterized in fluid mechanics by

$$V = \frac{Q}{2\pi} \frac{z}{|z|^2},$$

(where $Q > 0$ corresponds to a source, and $Q < 0$ - to a sink).



- For a punctual **source**, we have

$$V_x = \frac{Q}{2\pi} \frac{x}{x^2 + y^2}, \quad V_y = \frac{Q}{2\pi} \frac{y}{x^2 + y^2},$$

which leads to the complex potential

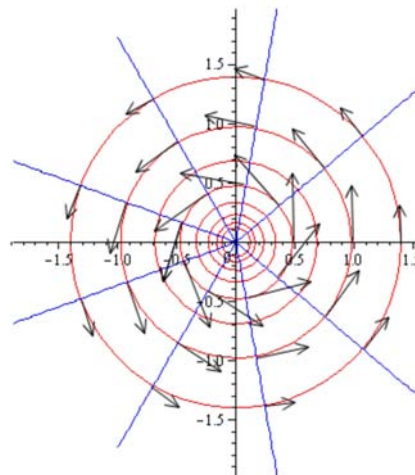
$$\Omega = \frac{Q}{2\pi} \operatorname{Ln}(z).$$

Example 20 • The movement of a fluid, characterized by

$$V = \frac{\Gamma i}{2\pi} \frac{z}{|z|^2} = -\frac{\Gamma y}{2\pi(x^2 + y^2)} + i \frac{\Gamma x}{2\pi(x^2 + y^2)}$$

corresponds to a **vortex**. For $\Gamma > 0$, the fluid rotates counterclockwise, and for $\Gamma < 0$, it rotates clockwise.

$$\Omega = \frac{\Gamma}{2\pi i} \operatorname{Ln}(z).$$



- For a **dipole** in fluid mechanics, we have

$$\Omega = -\frac{M}{2\pi} \frac{1}{z},$$

where $M = Qh$ is the momentum of the dipole.

6.3.3 In Thermodynamics

The vector of heat flow of a punctual source is modeled by:

$$Q = \frac{q}{2\pi} \frac{z}{|z|^2},$$

where $q > 0$ is a constant.

In this case, the role of the potential function ϕ is played by the temperature v :

$$\vec{Q} = -k\nabla v.$$

The harmonic conjugate u of v is the function of heat flow u :

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = -\frac{1}{k}Q_y, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} = \frac{1}{k}Q_x$$

The vector $\vec{Q}$ is

$$\vec{Q} = -k\nabla v = -k\frac{\partial v}{\partial x} - ik\frac{\partial v}{\partial y} = -ki\left(\frac{\partial u}{\partial x} - i\frac{\partial v}{\partial x}\right) = -ki\overline{\Omega'(z)}.$$

6.4 The Use of Conformal Transformations

The method is the following: in order to find the complex potential Ω corresponding to a vector field on some domain D , we find a conformal transformation f of D onto some more convenient domain Δ , where the corresponding complex potential $\tilde{\Omega}$ can be easily found. Then, the potential Ω is obtained as the composition of the potential on Δ with f : $\Omega = \tilde{\Omega} \circ f$.

$$\begin{array}{ccc} D & \xrightarrow{f} & \Delta \\ & \searrow \Omega & \downarrow \tilde{\Omega} \\ & & \mathbb{C} \end{array}$$

Let us have some examples from fluid mechanics.

- Find the complex potential, the streamlines and the equipotential lines for a fluid flowing around a (straight-angled) corner,
 - a) on its outer side; b) on its inner side.

a) The outer side of the corner can be modeled, for instance, as the union of the first three quadrants in plane

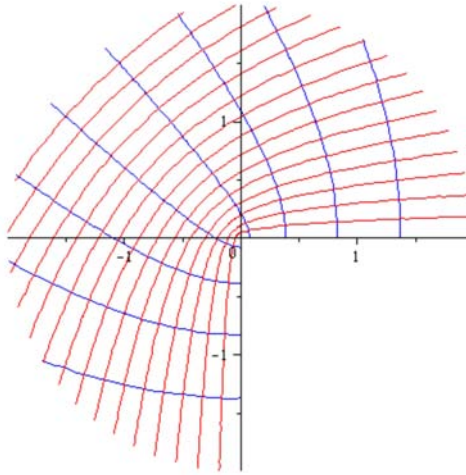
$$D = \{z \in \mathbb{C} \mid \text{Arg}(z) \in (0, \frac{3\pi}{2})\}.$$

The idea is to map this domain onto the upper half-plane

$$\Delta = \{w \in \mathbb{C} \mid \text{Arg}(w) \in (0, \pi)\}.$$

We admit that in the absence of sources, sinks, vortices or obstacles, a fluid has a flow which is parallel to the Ox axis, hence the corresponding potential is

$$\tilde{\Omega}(w) = Aw, \quad A > 0.$$



A conformal map $f : D \rightarrow \Delta$ is $f(z) = z^{2/3} =: w$, hence

$$\Omega(z) = \tilde{\Omega}(f(z)) = Az^{2/3}, \quad A > 0.$$

The streamlines and equipotential lines are indicated in the picture below (for $A := 1$):

b) The "inner side" of the corner

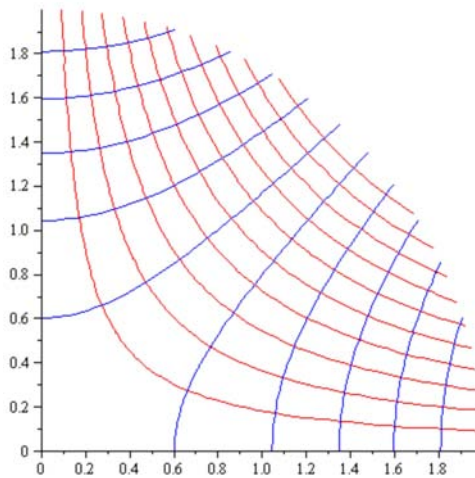
$$D = \{z \in \mathbb{C} \mid \text{Arg}(z) \in (0, \frac{\pi}{2})\}.$$

is to mapped onto the upper half-plane $\Delta = \{w \in \mathbb{C} \mid \text{Arg}(w) \in (0, \pi)\}$ by

$$f : D \rightarrow \Delta, \quad f(z) = z^2 =: w$$

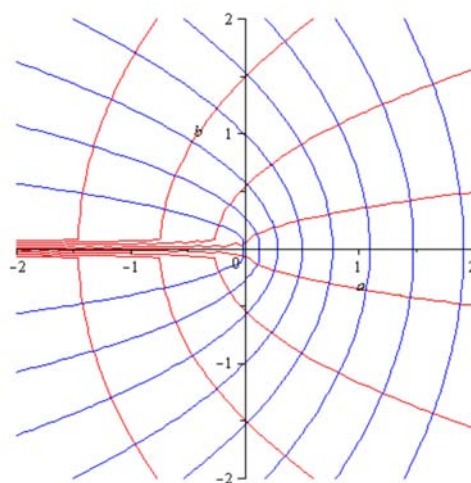
On Δ , by the same idea as above, we have, $\tilde{\Omega}(w) = Aw$, $A > 0$, hence

$$\Omega(z) = \tilde{\Omega}(f(z)) = Az^2, \quad A > 0.$$

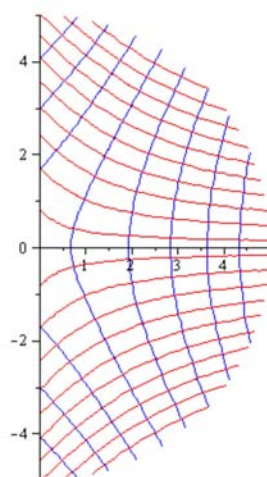


6.5 Other Examples

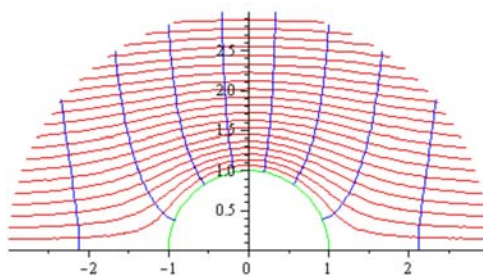
1) For $\Omega(z) = \sqrt{z}$



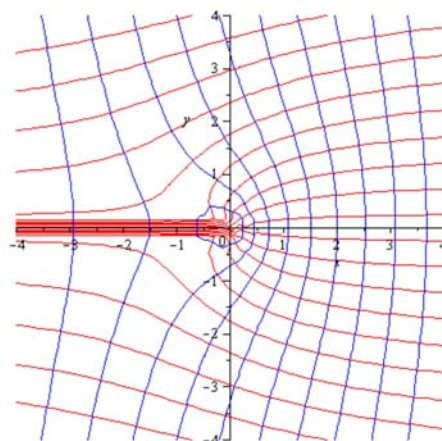
2) Wedge flow: $\Omega(z) = Az^{3/2}$



3) Flow around a half-circle: $\Omega(z) = A\left(z + \frac{1}{z}\right)$:

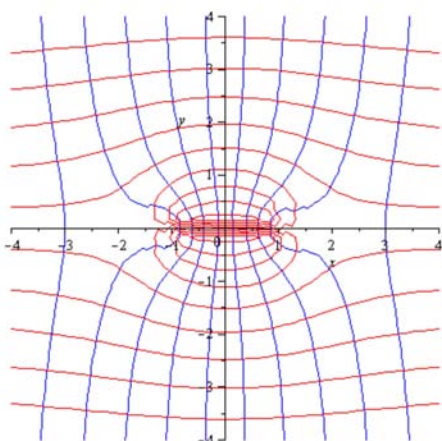


4) **Rankine half-body:** This flow results from combining a uniform flow ($\Omega = -Az$) with a source located at the origin ($\Omega = -\mu \ln z$). The resulting complex potential is: $\Omega = -Uz - \mu \ln(z)$;

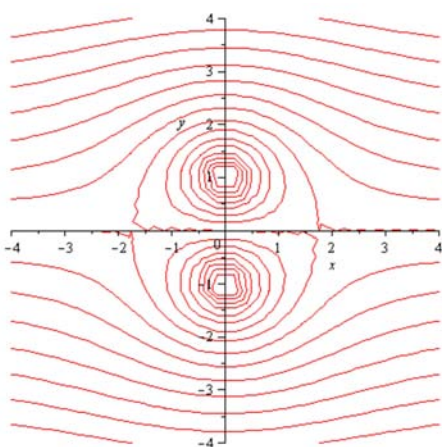


5. The **Rankine oval** - which results from combining a uniform flow with a source at $x = -a$ and a sink at $x = a$. Thus, the complex potential is given by:

- $\Omega := -Uz - \mu \ln(z + a) + \mu \ln(z - a)$



6. The **Kelvin oval** -obtained by combining two vortices, a clockwise vortex at $y = a$ and a counterclockwise vortex at $y = -a$, both of strenght $\mu = \frac{\Gamma}{2\pi}$, with a uniform flow of magnitude $\Gamma = 4\pi aU$



•

7 Integration

7.1 Definition and Basic Properties

As we have already seen, for functions $\phi : [a, b] \rightarrow \mathbb{C}$ of a real variable $t \in [a, b] \subset \mathbb{R}$, the integral is simply

$$\int_a^b \phi(t) dt = \int_a^b \operatorname{Re}(\phi(t)) dt + i \int_a^b \operatorname{Im}(\phi(t)) dt. \quad (28)$$

For functions $f : D \rightarrow \mathbb{C}$ defined over a subset $D \subset \mathbb{C}$, since there is no unique path between two given points in plane, integration is to be performed along curves $\gamma : [a, b] \rightarrow D$. Namely,

Definition 21 Let $\gamma : [a, b] \rightarrow \mathbb{C}$ be a smooth curve and f is a complex function which is continuous along γ . The integral of f on γ is

$$\int_{\gamma} f := \int_{\gamma} f(z) dz = \int_a^b f(\gamma(t)) \gamma'(t) dt. \quad (29)$$

The definition naturally extends to piecewise smooth curves. If $\gamma : [a, b] \rightarrow \mathbb{C}$ is piecewise smooth, then the integral of f along γ is defined as the sum of the integrals of f along the "pieces" $[a, c_1], [c_1, c_2], \dots, [c_{n-1}, b]$ of the interval $[a, b]$ along which γ is smooth:

$$\int_{\gamma} f = \int_a^{c_1} f(\gamma(t)) \gamma'(t) dt + \int_{c_1}^{c_2} f(\gamma(t)) \gamma'(t) dt + \dots + \int_{c_{n-1}}^b f(\gamma(t)) \gamma'(t) dt.$$

Let us make a remark. If the smooth curve γ is described by

$$\gamma : z = \gamma(t) \equiv x(t) + iy(t),$$

then the term $\gamma'(t) dt$ is nothing but $dz = dx + idy$, and the integral can be re-expressed as

$$\int_{\gamma} f(z) dz = \int_{\gamma} (u + iv)(dx + idy) = \int_{\gamma} u dx - v dy + i \int_{\gamma} v dx - u dy.$$

Example: Let $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = \bar{z}^2 = (x^2 - y^2) - 2ixy$.

a) Let us calculate the integral of f over the line segment $\gamma(t) = t + it$, $t \in [0, 1]$.

We have $x = t$, $y = t$, and $\gamma'(t) = 1 + i$. Replacing into (29), we get

$$\begin{aligned} \int_{\gamma} f &= \int_0^1 f(\gamma(t)) \gamma'(t) dt = \int_0^1 [(t^2 - t^2) - 2it^2](1 + i) dt = \\ &= -2i(1 + i) \frac{t^3}{3} \Big|_0^1 = \frac{2}{3}(1 - i). \end{aligned}$$

b) The integral of f along the parabola $\gamma(t) = t + it^2$, $t \in [0, 1]$ is computed in the same manner. Namely, $x = t$, $y = t^2$, $\gamma'(t) = 1 + 2it$, and the integral becomes

$$\int_{\gamma} f = \int_0^1 [(t^2 - t^4) - 2it^3](1 + 2it)dt = \int_0^1 t^2 + 3t^4 - 2it^5 dt,$$

which is the integral of a complex function defined over an interval in $\mathbb{R}$ and hence goes by (28). We get

$$\int_{\gamma} f = \frac{t^3}{3} + \frac{3t^5}{5} - 2i \frac{t^6}{6} \Big|_0^1 = \frac{1}{3} + \frac{3}{5} - \frac{1}{3}i.$$

Some basic properties of the complex integral are listed below:

1. $\int_{\gamma} \alpha f + \beta g = \alpha \int_{\gamma} f + \beta \int_{\gamma} g$;
2. Changing orientation of the curve leads to a change of sign of the integral. If γ^- is the curve obtained by "going conversely" along $\gamma : [a, b] \rightarrow \mathbb{C}$, namely: $\gamma^-(t) = \gamma(a + b - t)$, $\forall t \in [a, b]$, then

$$\int_{\gamma^-} f = - \int_{\gamma} f.$$

3. The complex integral is additive with respect to the integration path:

$$\int_{\gamma_1 \cup \gamma_2} f = \int_{\gamma_1} f + \int_{\gamma_2} f,$$

where $\gamma_1 \cup \gamma_2$ is the path obtained by joining the paths γ_1 and γ_2 with the property that the ending point of γ_1 coincides with the starting point of γ_2 .

4. $\left| \int_{\gamma} f \right| \leq \max_{z \in \gamma} |f(z)| \cdot \text{length}(\gamma)$, where $\text{length}(\gamma) = \int_a^b |\gamma'(t)| dt = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt$.

Proof. 1) - obvious.

2) We have

$$\begin{aligned} \int_{\gamma^-} f &= \int_a^b f(\gamma(a + b - t)) \gamma'(a + b - t) dt = \\ &\stackrel{t^* = a+b-t}{=} \int_b^a f(\gamma(t^*)) \gamma'(t^*) dt^* = - \int_a^b f(\gamma(t^*)) \gamma'(t^*) dt^* = - \int_{\gamma} f. \end{aligned}$$

3) Let $\gamma = \gamma_1 \cup \gamma_2$, and $\gamma_1, \gamma_2 : [0, 1] \rightarrow \mathbb{C}$ (we chose for simplicity $[0, 1]$ as the interval of definition of γ_1, γ_2 - the proof can be made analogously for the general case). Then γ can be parametrized as $\gamma : [0, 2] \rightarrow \mathbb{C}$, $\gamma(t) = \begin{cases} \gamma_1(t), & t \in [0, 1] \\ \gamma_2(t - 1), & t \in [1, 2] \end{cases}$ and the statement is proved by using the additivity of the Riemann integral with respect to the interval.

$$4) \left| \int_{\gamma} f \right| = \left| \int_a^b f(\gamma(t)) \gamma'(t) dt \right| \leq \int_a^b |f(\gamma(t))| |\gamma'(t)| dt \leq \max_{z \in \gamma} |f(z)| \int_a^b |\gamma'(t)| dt = \max_{z \in \gamma} |f(z)| \cdot \text{length}(\gamma). \blacksquare$$

7.2 Antiderivatives

Let $D \subset \mathbb{C}$ be a domain and $f : D \rightarrow \mathbb{C}$. We say that f admits an antiderivative (a primitive) if there exists some $F : D \rightarrow \mathbb{C}$ such that $F' = f$.

If f has an antiderivative, then its integral does not depend on the path joining two fixed points in plane..

Theorem 22 *If $\gamma : [a, b] \rightarrow D$ is smooth, f is continuous on D and $F' = f$ over D , then*

$$\int_{\gamma} f = F(\gamma(b)) - F(\gamma(a)).$$

In particular, if f has an antiderivative, then the integral $\int_{\gamma} f$ does not depend on the path γ joining $p = \gamma(a)$ and $q = \gamma(b)$.

Proof. We have $\int_{\gamma} f = \int_a^b f(\gamma(t)) \gamma'(t) dt = F(\gamma(t)) \Big|_{t=a}^{t=b} . \blacksquare$

Remark: Having a look at the example in the previous section, we can notice that the integral $\int_{\gamma} \bar{z}^2 dz$ depends on the integration path, hence $\bar{z}^2$ does not admit a primitive.

Also, an immediate consequence of the above theorem is

Corollary 23 *If γ is a smooth closed curve in D and the continuous function f admits a primitive on D , then the integral of f along γ vanishes:*

$$\int_{\gamma} f = 0.$$

7.3 Cauchy's Theorem

Let again D denote a domain in $\mathbb{C}$.

Definition 24 *Two closed curves $\gamma_1, \gamma_2 : [0, 1] \rightarrow D$ are called D -homotopic (and we denote $\gamma_1 \sim_D \gamma_2$) if there exists a continuous function $h : [0, 1] \times [0, 1] \rightarrow D$ such that;*

1. $h(t, 0) = \gamma_1(t), \quad \forall t \in [0, 1];$
2. $h(t, 1) = \gamma_2(t), \quad \forall t \in [0, 1];$
3. $h(0, s) = h(1, s), \quad \forall s \in [0, 1]$ (all curves $h(\cdot, s)$ are closed).

The function h is called a continuous deformation of γ_1 into γ_2 .

One of the most important results in complex analysis refers to the values of the integral of a holomorphic function over homotopic paths. Namely,

Theorem 25 (Cauchy): If $f : D \rightarrow \mathbb{C}$ is holomorphic on D and $\gamma_1 \sim_D \gamma_2$, then

$$\int_{\gamma_1} f = \int_{\gamma_2} f.$$

Proof. We will make the proof in the case when the homotopy h is of class at least 2. The result holds true also in the absence of this supplementary restriction, but the proof exceeds the purposes of this course.

Let $I_s := \int_{\gamma_s} f$, where $\gamma_s : [0, 1] \rightarrow \mathbb{C}$ is $\gamma_s(t) = h(t, s)$. Then $I(s) = \int_0^1 f(h(t, s)) \frac{\partial h}{\partial t}(t, s) dt$ and its derivative with respect to s is

$$\begin{aligned} \frac{dI}{ds}(s) &= \int_0^1 \frac{\partial}{\partial s} [f(h(t, s)) \frac{\partial h}{\partial t}(t, s)] dt = \\ &= \int_0^1 f'(h(t, s)) \frac{\partial h}{\partial s} \frac{\partial h}{\partial t} + f(h(t, s)) \frac{\partial^2 h}{\partial s \partial t} dt \stackrel{c^2}{=} \int_0^1 \frac{\partial}{\partial t} [f(h(t, s)) \frac{\partial h}{\partial s}(t, s)] dt = \\ &= f(h(t, s)) \frac{\partial h}{\partial s}(t, s) \Big|_{t=0}^{t=1} = 0, \end{aligned}$$

where the last equality is due to the fact that $h(0, s) = h(1, s)$ for all $s \in [0, 1]$ (which also implies $\frac{\partial h}{\partial s}(0, s) \equiv \frac{\partial h}{\partial s}(1, s)$).

Therefore, we have $\frac{dI}{ds} = 0$, which means that $I = I(s)$ is a constant. In particular, $I(0) = I(1)$, which completes the proof. ■

Corollary 26 If the path γ is D -homotopic to a point (in other words, D -contractible), then the integral of any holomorphic function on γ is zero: $\int_{\gamma} f = 0$.

Corollary 27 If f is an entire function and γ is closed, then $\int_{\gamma} f = 0$.

7.4 Cauchy's Integral Formula

A spectacular result concerning holomorphic functions is that their values over some closed contour uniquely determine their values inside that contour. These values are given by Cauchy's integral formula.

Let us present for the beginning a particular case, which will be then used for proving the general one.

Theorem 28 (Cauchy's formula for a circle): Let C_R be the circle of center w and radius R in the complex plane (oriented counterclockwise) and suppose f is holomorphic on the interior $D = \text{int}(C_R)$. Then the value of f at the center w is

$$f(w) = \frac{1}{2\pi i} \int_{C_R} \frac{f(z)}{z - w} dz.$$

Proof. We have that $C_R \sim_D \setminus \{w\} C_r$, where $r \rightarrow 0$, and the function $g(z) = \frac{f(z)}{z-w}$ is holomorphic in any open set contained in $D \setminus \{w\}$. Hence, according to Cauchy's theorem,

$$\int_{C_R} \frac{f(z)}{z-w} dz = \int_{C_r} \frac{f(z)}{z-w} dz.$$

We shall prove in the following that the difference $\int_{C_r} \frac{f(z)}{z-w} dz - 2\pi i f(w)$ tends to zero as r tends to zero..

The circle C_r can be parametrized as

$$z = w + re^{i\theta}, \quad \theta \in [0, 2\pi].$$

We have

$$\int_{C_r} \frac{1}{z-w} dz = \int_0^{2\pi} \frac{ire^{i\theta}}{re^{i\theta}} d\theta = 2\pi i.$$

Then, the modulus of the considered difference is

$$\begin{aligned} & \left| \int_{C_r} \frac{f(z)}{z-w} dz - 2\pi i f(w) \right| = \left| \int_{C_r} \frac{f(z)}{z-w} - \frac{f(w)}{z-w} dz \right| \leq \\ & \leq \max_{C_r} \left| \frac{f(z) - f(w)}{z-w} \right| \cdot 2\pi r = \max_{C_r} \frac{|f(z) - f(w)|}{r} \cdot 2\pi r = \\ & = 2\pi \max_{C_r} |f(z) - f(w)| \xrightarrow{r \rightarrow 0} 0. \end{aligned}$$

That is, $\int_{C_R} \frac{f(z)}{z-w} dz = \int_{C_r} \frac{f(z)}{z-w} dz = 2\pi i f(w)$, q.e.d. ■

Let us now generalize the result.

A closed oriented curve γ is called *positively oriented* if its "interior" is situated on the left side of an observer moving in the sense indicated by the orientation of the curve.

The existence of the "inside" of a closed contour is stated by the *Jordan Curve Theorem*: If c is a simple closed curve in the plane $\mathbb{C}$, then the complement of the image of c consists of two distinct connected components. One of these components is bounded (the interior) and the other is unbounded (the exterior). The image of c is the boundary of each component.

Theorem 29 (*Cauchy's integral formula*): Let f be holomorphic on some domain D , and γ a simple, smooth, positively oriented, closed and D -contractible curve in D . If w is an arbitrary point in the domain enclosed by γ , $w \in \text{int}(\gamma)$, then

$$f(w) = \frac{1}{2\pi i} \int_{C_r} \frac{f(z)}{z-w} dz.$$

The statement is obvious, by combining the previous theorem with Cauchy’s theorem.

For further use, let us re-express Cauchy’s formula for the circle in a convenient manner. If $w \in \mathbb{C}$ is the center of the circle $C_r := \mathcal{C}(w, r)$, then along this circle we have $z = w + re^{it}$, $t \in [0, 2\pi]$. We get that

$$f(w) = \frac{1}{2\pi i} \int_{C_r} \frac{f(z)}{z - w} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(w + re^{it})}{re^{it}} ire^{it} dt = \frac{1}{2\pi} \int_0^{2\pi} f(w + re^{it}) dt.$$

Splitting the above into its real and imaginary parts, $f(w) = u(w) + iv(w)$, we have

$$u(w) = \frac{1}{2\pi} \int_0^{2\pi} u(w + re^{it}) dt, \quad v(w) = \frac{1}{2\pi} \int_0^{2\pi} v(w + re^{it}) dt. \quad (30)$$

One of the most beautiful applications of the Cauchy’s formula is to the computation of complex integrals. Let us have an example.

Example: Let us calculate $\int_{\gamma} \frac{\cos z}{z} dz$, where γ is the unit circle $\mathcal{C}(0, 1)$.

The function $f(z) = \cos z$ is holomorphic on the interior of γ . Moreover, $w = 0$ belongs to the interior of γ , hence we can apply Cauchy’s formula, thus getting

$$\int_{\gamma} \frac{\cos z}{z - 0} dz = 2\pi i \cos 0 = 2\pi i.$$

Also, Cauchy’s formula can be extended, allowing us to calculate not only the values $f(w)$ of a function inside a contour, but also its derivatives.

Theorem 30 (*Cauchy’s formula for derivatives*): Let f be holomorphic in some domain $D \subset \mathbb{C}$, γ a positively oriented, simple, closed, smooth and D -contractible curve and w , an arbitrary point in the interior of γ . Then:

$$\begin{aligned} f'(w) &= \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - w)^2} dz, \\ f''(w) &= \frac{1}{\pi i} \int_{\gamma} \frac{f(z)}{(z - w)^3} dz, \\ &\dots \\ f^{(k)}(w) &= \frac{k!}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - w)^{k+1}} dz, \quad k \in \mathbb{N}^*. \end{aligned}$$

The proof can be made inductively, by subsequent deriving over w .

An immediate consequence of the above is:

Corollary 31 If $f : D \rightarrow \mathbb{C}$ is holomorphic in the domain $D \subset \mathbb{C}$, then it is of class C^∞ on D (it is derivable infinitely many times).

Also, the above formula can be used in order to compute integrals.

Example: Calculate $\int_{|z|=1} \frac{\sin z}{z^2} dz$.

We have that $f(z) = \sin z$ is holomorphic on the disk $D := \{z \in \mathbb{C} \mid |z| < 1\}$, and $w = 0$ belongs to D . Hence,

$$\int_{|z|=1} \frac{\sin z}{(z-0)^2} dz = 2\pi i f'(0) = 2\pi i \cos 0 = 2\pi i.$$

7.5 Applications of Cauchy's Formula

A host of important mathematical results can be easily proved by making use of Cauchy's formula. Let us see several of them.

Theorem 32 (*Fundamental Theorem of Algebra*): *Every non-constant polynomial with coefficients in $\mathbb{C}$ has a root in $\mathbb{C}$.*

Proof. Let us suppose that the polynomial $p \in \mathbb{C}[X]$ has no roots in $\mathbb{C}$. That is, $p(z) \neq 0$ for all $z \in \mathbb{C}$. Then, the function $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) = \frac{1}{p(z)}$ is entire.

Let C_R denote the circle of center $w = 0$ and radius $R > 0$. By Cauchy's formula, we have

$$\frac{1}{p(0)} = \frac{1}{2\pi i} \int_{C_R} \frac{\frac{1}{p(z)}}{z} dz.$$

Let us make now $R \rightarrow \infty$. We get

$$\left| \frac{1}{p(0)} \right| = \left| \frac{1}{2\pi i} \int_{C_R} \frac{1}{zp(z)} dz \right| \leq \left| \frac{1}{2\pi i} \right| \max_{C_R} \left(\frac{1}{|zp(z)|} \right) \cdot 2\pi R.$$

Let p be of degree n , hence $p(z) = a_n z^n + \dots + a_1 z + a_0$, $\forall z \in \mathbb{C}$. For large values of R , the values of p along the circle $C_R : |z| = R$ are well approximated by the term of maximal degree. Hence, for R big enough, we can state that, for instance,

$$|zp(z)| \in \left[\frac{1}{2} |a_n z^{n+1}|, 2 |a_n z^{n+1}| \right].$$

We get that

$$\left| \frac{1}{p(0)} \right| \leq \frac{2\pi R}{2\pi} \frac{2}{|a_n z^{n+1}|} = \frac{2R}{|a_n| R^{n+1}} \xrightarrow{R \rightarrow \infty} 0.$$

We are led to $\frac{1}{p(0)} = 0$, which is absurd. Consequently, p has to admit at least a root. ■

Another important consequence of Cauchy's formula is

Theorem 33 (*Liouville*): *Any bounded entire function is constant.*

Proof. Let $f : \mathbb{C} \rightarrow \mathbb{C}$ be entire and bounded, $|f(z)| \leq M, \forall z \in \mathbb{C}$.

By Cauchy’s formula for the first derivative, we have, for arbitrary $w \in \mathbb{C}$,

$$f'(w) = \frac{1}{2\pi i} \int_{C_R} \frac{f(z)}{(z-w)^2} dz,$$

where C_R is the circle centered at w and of radius R :

$$C_R = \{z \in \mathbb{C} \mid z = w + R e^{it}, \quad t \in [0, 2\pi]\}.$$

Taking the modulus of the above, we are led to

$$|f'(w)| = \frac{1}{2\pi} \left| \int_{C_R} \frac{f(z)}{(z-w)^2} dz \right| \leq \frac{1}{2\pi} \max_{C_R} \frac{|f(z)|}{|(z-w)^2|} \cdot 2\pi R.$$

But, along C_R , the denominator $|(z-w)^2|$ is constantly equal to R^2 , hence

$$|f'(w)| \leq \frac{2\pi R}{2\pi R^2} \max_{C_R} |f(z)| \leq \frac{M}{R},$$

which tends to 0 as R tends to infinity.

This means that $f'(w) = 0$, and since w is arbitrary, the function f is a constant. ■

We have up to now left as open the question of the relation between holomorphic functions and functions which admit antiderivatives. The answer is given by

Theorem 34 *If a function $f : D \rightarrow \mathbb{C}$ is holomorphic in the simply connected domain D , then f has an antiderivative in D .*

Proof. Let

$$F(z) = \int_{\gamma} f(w) dw, \quad \forall z \in D$$

where γ is an arbitrary path joining a fixed point $z_0 \in D$ to point z . By splitting $f = u + iv$, $w = x + iy$, the integral is written as

$$F(z) = \int_{\gamma} f(w) dw = \int_{\gamma} u dx - v dy + i \int_{\gamma} v dx - u dy.$$

By Cauchy-Riemann equations, it follows that the two integrals are path independent, hence $F(z)$ is well defined. Let us denote

$$U = \int_{z_0}^z u dx - v dy, \quad V = \int_{z_0}^z v dx - u dy.$$

That is, we have $F(z) = U + iV$ and

$$\begin{aligned} dU &= udx - vdy \Rightarrow \frac{\partial U}{\partial x} = u \\ dV &= vdx - udy \Rightarrow \frac{\partial V}{\partial x} = v. \end{aligned}$$

We get $F'(z) = \frac{\partial U}{\partial x} + i\frac{\partial V}{\partial x} = u + iv = f$, which completes the proof. ■

The converse statement is obvious.

Theorem 35 (Morera): *If f is a continuous function on a domain D and $\int_{\gamma} f = 0$ for all smooth closed paths in D , then f is holomorphic in D .*

Proof. We shall show that f admits an antiderivative in D . Let $a \in D$ be fixed and

$$F(z) := \int_{\gamma_z} f,$$

where $z \in D$ and γ_z is an arbitrary path in D , joining a and z (since D is connected, γ_z always exists).

- The function F is well defined (it does not depend on the path γ_z):

Let γ'_z be some other path in D , joining a and z . Then, the difference

$$\int_{\gamma_z} f - \int_{\gamma'_z} f = \int_{\gamma_z} f + \int_{(\gamma'_z)^-} f = \int_{\gamma_z \cup (\gamma'_z)^-} f$$

is zero, by hypothesis. Hence, $\int_{\gamma_z} f = \int_{\gamma'_z} f$.

- There holds $F'(z) = f(z)$, $\forall z \in D$:

We have $F'(z) = \lim_{\Delta z \rightarrow 0} \frac{F(z + \Delta z) - F(z)}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{1}{\Delta z} \int_{\gamma_{\Delta z}} f(w)dw$. Then, it easily follows that

$$|F'(z) - f(z)| \xrightarrow{\Delta z \rightarrow 0} 0,$$

hence F is a primitive of f on D . Since f admits a primitive, it follows that it is also holomorphic on D .

■

Other applications of Cauchy's formula refer to harmonic functions.

Theorem 36 (The maximum principle): *If $K \subset \mathbb{R}^2$ is a compact set, and u is a non-constant harmonic function on K , then u takes its maximum/minimum values on the boundary of K .*

Proof. Let us suppose that the maximum value of u is obtained for some interior point $w \in K$. Let $C = \mathcal{C}(w, r)$, $r > 0$, be a circle contained in K . Then, the value of u at the center w is given by Cauchy’s formula

$$u(w) = \frac{1}{2\pi} \int_0^{2\pi} u(w + re^{it}) dt.$$

Since w is a maximum point, we have $u(z) \leq u(w)$ for all $z \in K$. Let us suppose that for some value $z_0 \in K$, the inequality is strict: $u(z_0) < u(w)$ and choose r such that z lies on the circle C_r . By the continuity of u , we conclude that for a whole interval (t_1, t_2) corresponding to points of C_r around z_0 , there holds $u(z) < u(w)$.

That is,

$$\begin{aligned} 2\pi u(w) &= \int_0^{t_1} u(w + re^{it}) dt + \int_{t_1}^{t_2} u(w + re^{it}) dt + \int_{t_2}^{2\pi} u(w + re^{it}) dt < \\ &< t_1 u(w) + (t_2 - t_1) u(w) + (2\pi - t_2) u(w), \end{aligned}$$

which is nothing but $2\pi u(w) < 2\pi u(w)$, contradiction. Consequently, $w \in \partial K$, q.e.d. ■

A consequence of the above is

Theorem 37 *If two harmonic functions agree on the boundary of a bounded domain D , then they agree inside the domain D .*

The proof is immediate, taking into account that the closure of D is compact and applying the maximum principle for the difference between the two functions.

In other words, the values of a harmonic function inside a bounded domain are uniquely determined by its values on the boundary. Determining a harmonic function on a domain D , by knowing its values on the boundary ∂D is called a *Dirichlet problem*.

The solution of the Dirichlet problem for the disk $|z| \leq \rho$, $\rho > 0$ is given by the celebrated *Poisson’s integral formula*, [3], which can be obtained as a direct consequence of (30):

Theorem 38 *Let us suppose that the values of the harmonic function ϕ are known along the circle $C_\rho : |z| = \rho$, for some $\rho > 0$. Then, the values of ϕ inside the disk $\{z \in \mathbb{C} \mid z = re^{i\theta}, r < \rho, \theta \in [0, 2\pi]\}$ are given by*

$$\phi(r, \theta) = \frac{\rho^2 - r^2}{2\pi} \int_0^{2\pi} \frac{\phi(re^{it})}{\rho^2 + r^2 - 2\rho r \cos(t - \theta)} dt.$$

8 Series. Power Series

8.1 Sequences

A *sequence* of complex numbers is a function $a : \mathbb{N}^* \rightarrow \mathbb{C}$, $a(n) =: z_n$, usually denoted simply by $(z_n)_{n \geq 1}$.

The number $L \in \mathbb{C}$ is the *limit* of the sequence $(z_n)_{n \geq 1}$ (or $(z_n)_{n \geq 1}$ is *convergent* to L) if for any $\varepsilon > 0$ there exists some $N_\varepsilon \in \mathbb{N}^*$ such that for all $n \geq N_\varepsilon$, $|z_n - L| < \varepsilon$.

Some properties of limits of convergent complex sequences:

- $\lim(z_n \pm w_n) = \lim z_n \pm \lim w_n$;
- $\lim(z_n w_n) = \lim z_n \cdot \lim w_n$;
- $\lim\left(\frac{z_n}{w_n}\right) = \frac{\lim z_n}{\lim w_n}$, if $\lim w_n \neq 0$.
- If $z_n = x_n + iy_n$ for all $n \geq 1$, then

$$\lim z_n = L \Leftrightarrow \exists \lim x_n = \operatorname{Re} L, \lim y_n = \operatorname{Im} L.$$

- The sequence $(z_n)_{n \geq 1}$ is convergent if and only if it is fundamental, i.e., or any $\varepsilon > 0$ there exists some $N_\varepsilon \in \mathbb{N}^*$ such that for all $m, n \geq N_\varepsilon$, $|z_n - z_m| < \varepsilon$.

We say that $\lim z_n = \infty$ if $\forall M > 0 \exists N_\varepsilon \in \mathbb{N}^*$ such that for any $n \geq N_\varepsilon$, $|z_n| > M$. The rules regarding the limit of the sum, difference, product and ratio of two sequences remain valid, with the exception of the indeterminates: $\infty - \infty$, $\frac{\infty}{\infty}$, $0 \cdot \infty$.

For sequences of functions, there is a special type of convergence, namely, *uniform convergence*. Let $f_n : D \rightarrow \mathbb{C}$, $n \in \mathbb{N}^*$, be functions defined over a domain $D \subset \mathbb{C}$. We say that

- f_n converges pointwise to $f : D \rightarrow \mathbb{C}$, if

$$\forall \varepsilon > 0 \forall z \in D \exists N = N(\varepsilon, z) \forall n \geq N : |f_n(z) - f(z)| < \varepsilon.$$

- f_n converges uniformly to $f : D \rightarrow \mathbb{C}$, if

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) \forall z \in D \forall n \geq N : |f_n(z) - f(z)| < \varepsilon.$$

In the case of pointwise convergence, the rank N depends on the point z , while in the case of the uniform one, the rank N is common for all $z \in D$, and only depends on ε .

The importance of uniform convergence is given by:

Proposition 39 Let D be a domain in $\mathbb{C}$, $f_n : D \rightarrow \mathbb{C}$, $n \in \mathbb{N}^*$ be a sequence of functions which uniformly converges to a function $f : D \rightarrow \mathbb{C}$.

- If f_n are continuous on D for all $n \in \mathbb{N}^*$, then f is continuous on D .
- For any curve $\gamma : [a, b] \rightarrow D$, the integrals $\int_\gamma f_n$ uniformly converge to $\int_\gamma f$.
- If f_n are holomorphic on D , then f is also holomorphic on D .

Proof. a) Let $\varepsilon > 0$. By the uniform convergence of f_n , we know that

$$\exists N = N(\varepsilon) \forall z \in D \forall n \geq N : |f_n(z) - f(z)| < \frac{\varepsilon}{3}.$$

Let now $z_0 \in D$ be fixed. Since f_n are continuous, there exist some $\delta_n = \delta_n(\varepsilon)$ such that $|z - z_0| < \delta_n \Rightarrow |f_n(z) - f_n(z_0)| < \frac{\varepsilon}{3}$.

Let us choose and fix $n \geq N(\varepsilon)$, $\delta := \delta_n$ and $|z - z_0| < \delta$. Then

$$\begin{aligned} |f(z) - f(z_0)| &\leq |f(z) - f_n(z)| + |f_n(z) - f_n(z_0)| + |f_n(z_0) - f(z_0)| < \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Hence, $|z - z_0| < \delta$ implies $|f(z) - f(z_0)| < \varepsilon$, q.e.d.

b) is immediate, using the inequality $\left| \int_{\gamma} f_n(z) - f(z) dz \right| \leq \max_{\gamma} |f_n(z) - f(z)| \cdot \text{length}(\gamma)$.

c) Since f_n are holomorphic, their integrals over any close contour γ in D vanish: $\int_{\gamma} f_n(z) = 0$, $\forall n \geq 1$.

Passing to the limit and using b), we get that $\int_{\gamma} f(z) = 0$. Taking into account that γ is arbitrary, by Morera's theorem, we deduce that f is also holomorphic. ■

8.2 Series

Series are sequences $(s_n)_{n \geq 1}$, where s_n is of the form $s_n = a_1 + a_2 + \dots + a_n$ for some other sequence $(a_n)_{n \geq 1}$ in $\mathbb{C}$. A series is convergent if it has a (finite) limit in $\mathbb{C}$. We agree to denote both the series $(s_n)_{n \geq 1}$ and its limit by

$$\sum_{n=1}^{\infty} a_n. \quad (31)$$

A similar result to the one in real analysis holds, namely:

Proposition 40 *If the series (31) is convergent, then the general term a_n tends to 0.*

Exercise: Prove this Proposition.

The converse is obviously not true (counterexamples can be found even in the particular case of series with real terms).

Let us consider now *series of functions* $\sum_{n=1}^{\infty} f_n$, where $f_n = f_n(z)$ are defined over some domain $D \subset \mathbb{C}$.

A simple criterion of convergence is obtained by comparing the moduli $|f_n(z)|$ with the terms of a series with real nonnegative terms.

Theorem 41 (*Weierstrass M-test*): *Let $M_j \in \mathbb{R}$, $M_j \geq 0$, for all $j \geq J$, $J \in \mathbb{N}^*$ such that the series $\sum_{j=1}^{\infty} M_j$ is convergent. If for any $z \in D$ and any $j \geq J$, $|f_j(z)| \leq M_j$, then the series $\sum_{j=1}^{\infty} f_j$ converges **uniformly** on D .*

Proof. We show that the sequence $s_n := \sum_{j=1}^n f_j$ of partial sums is fundamental.

Since $\sum_{j=1}^{\infty} M_j$ is convergent, it is also fundamental. That is, for any $\varepsilon > 0$, we can find some

$$N = N(\varepsilon) \text{ with } m, n \geq N(\varepsilon) \Rightarrow \sum_{j=m}^n M_j < \varepsilon.$$

Then, for all $m, n \geq N(\varepsilon)$ and for all z in D , we have $|s_n(z) - s_m(z)| = \left| \sum_{j=1}^{\infty} f_j \right| \leq \sum_{j=1}^{\infty} |f_j| \leq$

$\sum_{j=1}^{\infty} M_j < \varepsilon$, which means that the sequence $s_n(z)$ is fundamental. Consequently, it is also convergent, and moreover, the rank $N = N(\varepsilon)$ does not depend on z , that is, convergence is uniform. ■

Exercises:

1. Show that $\sum_{n=1}^{\infty} a_n$ is convergent if and only if the series $\sum_{n=1}^{\infty} \operatorname{Re} a_n$ and $\sum_{n=1}^{\infty} \operatorname{Im} a_n$ are both convergent.
2. Prove that $\sum_{n=1}^{\infty} \left(\frac{1}{z}\right)^n$ is uniformly convergent on the set $|z| > 5$.

A special type of series of functions are *power series*, namely, series defined by

$$s_n(z) = c_0 + c_1(z - z_0) + \dots + c_n(z - z_0)^n, \quad (32)$$

briefly, $\sum_{j=1}^{\infty} c_j(z - z_0)^j$. We say that (32) is a power series *centered at* z_0 .

Let us investigate for the beginning a particular case: the *geometric series*

$$\sum_{j=1}^{\infty} z^j. \quad (33)$$

Its partial sums are $s_n = 1 + z + \dots + z^n = \frac{1 - z^{n+1}}{1 - z}$, (for $z \neq 1$).

- For $|z| < 1$, we have $|z^{n+1}| \rightarrow 0$ hence the series converges to $\sum_{j=1}^{\infty} z^j = \frac{1}{1 - z}$. Moreover, taking $M \in [|z|, 1)$ and applying the Weierstrass M-test with $M_j := M^j$, we obtain that the convergence is uniform.
- For $|z| = 1$, we have $s_n = n$ and the series is divergent.
- For $|z| > 1$, the series is again divergent since its general term z^n does not tend to zero.

The situation in the above example can be generalized to arbitrary power series.

Theorem 42 Let $s_n = \sum_{j=1}^n c_j(z - z_0)^j$, and

$$\begin{aligned} \lambda &= \limsup \sqrt[j]{|c_j|} = \lim_{n \rightarrow \infty} \{\sup \sqrt[j]{|c_j|} \mid j \geq n\}, \\ R &= \frac{1}{\lambda}. \end{aligned}$$

The series $(s_n)_{n \geq 1}$ uniformly converges in the disk $|z - z_0| < R$ and diverges for $|z - z_0| > R$.

The number $R = \frac{1}{\lambda}$ is called the *radius of convergence* of the series, while $D : |z - z_0| < R$ is called its *disk of convergence*.

Remark 43 If the sequence $\sqrt[j]{|c_j|}$ has a limit, then

$$\lambda = \lim_{j \rightarrow \infty} \sqrt[j]{|c_j|}, \quad R = \frac{1}{\lambda}.$$

Proof. For $|z - z_0| > R$, let $k \in \mathbb{R}$ such that $|z - z_0| > \frac{1}{k} > R (\Rightarrow k < \lambda)$. There exist infinitely many terms $c_j |z - z_0|^j > (k |z - z_0|)^j > 1$, hence the general term $c_j (z - z_0)^j$ does not tend to zero, that is, the series is divergent.

If $|z - z_0| \leq \rho < R$, then there exists $k \in \mathbb{R}$ such that

$$\lambda = \frac{1}{R} < k < \frac{1}{\rho}.$$

That is, for j large enough, we have $|c_j| < k$. Then

$$|c_j (z - z_0)^j| < k^j \rho^j = (k\rho)^j < 1.$$

Applying the M-test for the series $\sum_{j=1}^n c_j (z - z_0)^j$, with $M_j := (k\rho)^j$, we get that the series is uniformly convergent. ■

Integration and differentiation of power series:

Using the Proposition at the beginning of this section, we can see that inside the disk of convergence, the sum

$$S(z) := \sum_{j=1}^{\infty} c_j (z - z_0)^j \quad (34)$$

is a holomorphic function. This is easy to see, if we take into account that each partial sum $s_n = \sum_{j=1}^n c_j (z - z_0)^j$ is a polynomial (hence, holomorphic) function and the convergence is uniform. Since it is holomorphic, it admits an antiderivative, and it is natural to ask how this antiderivative looks like.

Proposition 44 *Inside its disk D of convergence, the power series (34), eventually multiplied by a function g , can be integrated term by term along any smooth curve $\gamma : [a, b] \rightarrow D$ along which g is continuous:*

$$\int_{\gamma} g(z) S(z) dz = \sum_{j=1}^n c_j \int_{\gamma} g(z) (z - z_0)^j dz. \quad (35)$$

Proof. Let $\varepsilon > 0$ and $N = N(\varepsilon)$ such that $\left| S(z) - \sum_{j=1}^n c_j (z - z_0)^j \right| < \varepsilon$. Also, let $M = \max_{z \in \gamma} |g(z)|$.

The number M exists, since g is continuous, and the image $\gamma([a, b])$ is bounded.

Then, we have

$$\begin{aligned} \left| \int_{\gamma} g(z) S(z) - g(z) \sum_{j=1}^n c_j (z - z_0)^j dz \right| &= \left| \int_{\gamma} g(z) [S(z) - \sum_{j=1}^n c_j (z - z_0)^j] dz \right| \leq \\ &\leq M\varepsilon \cdot \text{length}(\gamma) \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

which proves the statement. ■

Proposition 45 *Inside its disk D of convergence, the power series (34) can be differentiated term by term:*

$$S'(z) = \sum_{j=1}^{\infty} j c_j (z - z_0)^{j-1}.$$

Proof. Cauchy's formula for the derivative $S'(z)$ says that

$$S'(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{S(w)}{(w - z)^2} dw,$$

where γ is a closed contour in D , such that z lies in the interior of γ . Let us apply the previous result, with $g(w) = \frac{1}{(w - z)^2}$ (which is continuous along γ !):

$$S'(z) = \frac{1}{2\pi i} \sum_{j=1}^{\infty} c_j \int_{\gamma} \frac{1}{(w - z)^2} (w - z_0)^j dw.$$

But $\frac{1}{2\pi i} \int_{\gamma} \frac{1}{(w - z)^2} (w - z_0)^j$ is again, by Cauchy's formula for the first derivative,

$$\frac{1}{2\pi i} \int_{\gamma} \frac{(w - z_0)^j}{[(w - z_0) - (z - z_0)]^2} dw = \frac{d}{dz} (z - z_0)^j = j(z - z_0)^{j-1}.$$

Replacing into the above, we get

$$S'(z) = \sum_{j=1}^{\infty} j c_j (z - z_0)^{j-1}.$$

■

8.3 Taylor Series

Let the function f be holomorphic on the disk $D : |z - z_0| < r$, $z \in D$ and $C = \mathcal{C}(z_0, \rho)$, with $\rho \in (0, r)$ be a circle contained in D . Then, for all $s \in C$ and z in the interior of $C : |z - z_0| < \rho$, we have

$$\frac{1}{s - z} = \frac{1}{(s - z_0) - (z - z_0)} = \frac{1}{s - z_0} \cdot \frac{1}{1 - \frac{z - z_0}{s - z_0}}.$$

But, $\left| \frac{z - z_0}{s - z_0} \right| < 1$, and we can write

$$\frac{1}{1 - \frac{z - z_0}{s - z_0}} = \sum_{n=0}^{\infty} \left(\frac{z - z_0}{s - z_0} \right)^n \Rightarrow \frac{1}{s - z} = \frac{1}{s - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{s - z_0} \right)^n,$$

where the convergence of the series is uniform.

Let us apply now Cauchy's formula for $f(z)$. We have

$$\begin{aligned} 2\pi i f(z) &= \int_C \frac{f(s)}{s-z} ds = \int_C \frac{f(s)}{s-z_0} \sum_{n=0}^{\infty} \left(\frac{z-z_0}{s-z_0} \right)^n ds = \\ &= \sum_{n=0}^{\infty} (z-z_0)^n \int_C \frac{f(s)}{(s-z_0)^{n+1}} ds. \end{aligned}$$

The integral $\int_C \frac{f(s)}{(s-z_0)^{n+1}} ds$ is none but $\frac{2\pi i}{n!} f^{(n)}(z_0)$, hence we can state:

Theorem 46 (Taylor's formula): Let f be holomorphic on some disk $D : |z - z_0| < r$, where $r > 0$. Then, for all $z \in D$, there holds

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n,$$

where the convergence of the series is uniform.

As a remark, the Taylor series is convergent on *any* disk on which f is holomorphic, hence, for z_0 fixed, the disk of convergence is the largest disk centered at z_0 inside which f is holomorphic.

Examples:

1) For $f(z) = e^z$ and $z_0 := 0$, we get

$$e^z = \sum_{n=0}^{\infty} \frac{e^{z_0}}{n!} (z - z_0)^n = \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{n!}.$$

The radius of convergence of the series is $R = \infty$.

2) If a holomorphic function $f : D \rightarrow \mathbb{C}$ has the property that $f(z_0) = f'(z_0) = \dots = f^{(k)}(z_0) = 0$ for some $z_0 \in D$ (that is, z_0 is a *zero of order $k+1$ for f*), then $f(z) = \sum_{n=k+1}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$.

That is, there exists a holomorphic function g on D , such that

$$f(z) = (z - z_0)^{k+1} g(z), \quad \forall z \in D.$$

9 Laurent Series. Residues

9.1 Laurent Series

Let us suppose now that f is a holomorphic function in a region $D : R_1 < |z - z_0| < R_2$, where $R_1, R_2 \in [0, \infty]$.

Theorem 47 For any $z \in D$, $f(z)$ can be written as the sum of the convergent series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} b_n (z - z_0)^{-n}, \quad (36)$$

where

$$a_n = \frac{1}{2\pi i} \int_C \frac{f(s)}{(s - z_0)^{n+1}} ds, \quad b_n = \frac{1}{2\pi i} \int_C \frac{f(s)}{(s - z_0)^{-n+1}} ds \quad (37)$$

and C is any circle centered at z_0 and contained in D .

Strictly speaking, $f(z)$ is the sum of two convergent series, but, for the sake of commodity, one usually writes the above in a condensed form

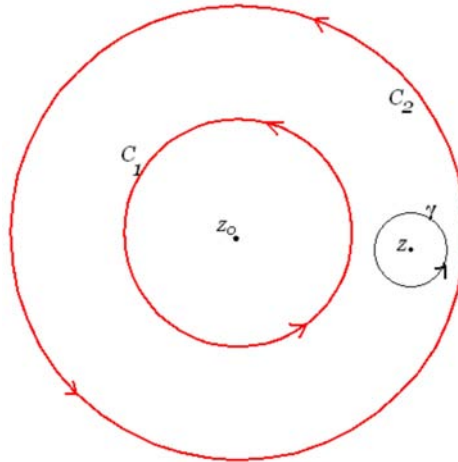
$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n, \quad c_n = \frac{1}{2\pi i} \int_C \frac{f(s)}{(s - z_0)^{n+1}} ds, \quad \forall n \in \mathbb{Z}. \quad (38)$$

The series (36) or, equivalently, (38), is called the *Laurent series of f , centered at z_0* .

Note that: 1) f needs not to be differentiable at z_0 ; 2) in the particular case $R_1 = 0$, meaning that f is holomorphic on a disk centered at z_0 , the integrals which provide b_n vanish, hence the Laurent series coincide with the Taylor series of f at z_0 .

Let us prove now the above theorem.

Let $C_1(z_0, r_1)$, $C_2(z_0, r_2)$ be positively oriented circles centered at z_0 , such that $R_1 < r_1 \leq |z - z_0| \leq r_2 < R_2$ and let Γ be another positively oriented circle centered at z , small enough as not to intersect C_1 , nor C_2 .



We have, by Cauchy's formula

$$2\pi i f(z) = \int_{\Gamma} \frac{f(s)}{s - z} ds,$$

since f is holomorphic inside Γ . On the other side,

$$\int_{C_2} \frac{f(s)}{s - z} ds = \int_{C_1} \frac{f(s)}{s - z} ds + \int_{\Gamma} \frac{f(s)}{s - z} ds$$

which leads to

$$2\pi i f(z) = \int_{C_2} \frac{f(s)}{s - z} ds - \int_{C_1} \frac{f(s)}{s - z} ds.$$

- Let us evaluate the first integral.

For $s \in C_2$, there holds $|s - z_0| > |z - z_0|$, which leads to

$$\frac{1}{s - z} = \frac{1}{(s - z_0) - (z - z_0)} = \frac{1}{s - z_0} \cdot \frac{1}{1 - \frac{z - z_0}{s - z_0}} = \frac{1}{s - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{s - z_0} \right)^n,$$

exactly as we did for the Taylor series. That is,

$$\int_{C_2} \frac{f(s)}{s - z} ds = \int_{C_2} \frac{f(s)}{s - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{s - z_0} \right)^n ds = \sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad (*)$$

where

$$a_n = \int_{C_2} \frac{f(s)}{(s - z_0)^{n+1}} ds. \quad (**)$$

- In order to calculate the second integral, $-\int_{C_1} \frac{f(s)}{s - z} ds$, let us notice that, for $s \in C_1$, we have $|s - z_0| < |z - z_0|$, hence

$$\frac{-1}{s - z} = \frac{1}{z - s} = \frac{1}{z - z_0} \cdot \frac{1}{1 - \frac{s - z_0}{z - z_0}}.$$

From $\left| \frac{s - z_0}{z - z_0} \right| < 1$, we conclude that $\frac{1}{1 - \frac{s - z_0}{z - z_0}} = \sum_{n=0}^{\infty} \left(\frac{s - z_0}{z - z_0} \right)^n$, and therefore

$$\frac{-1}{s - z} = \frac{1}{z - z_0} \sum_{n=0}^{\infty} \left(\frac{s - z_0}{z - z_0} \right)^n = \sum_{n=0}^{\infty} \frac{(s - z_0)^n}{(z - z_0)^{n+1}}.$$

Then,

$$-\int_{C_1} \frac{f(s)}{s - z} ds = \sum_{n=0}^{\infty} \int_{C_1} \frac{f(s)(s - z_0)^n}{(z - z_0)^{n+1}} ds = \sum_{n=1}^{\infty} \frac{1}{(z - z_0)^n} \int_{C_1} f(s)(s - z_0)^{n-1} ds.$$

Taking into account (*) and (**), we have

$$2\pi i f(z) = \sum_{n=0}^{\infty} (z - z_0)^n \int_{C_2} \frac{f(s)}{(s - z_0)^{n+1}} ds + \sum_{n=1}^{\infty} \frac{1}{(z - z_0)^n} \int_{C_1} f(s)(s - z_0)^{n-1} ds.$$

It suffices now to consider instead the integrals over C_1 and C_2 , integrals over some curve C which is D -homotopic to both C_1 and C_2 , and the proof is completed.

In practical applications, determining the coefficients c_n of the Laurent series by calculating the corresponding integrals might turn difficult, hence we usually apply some algebraic tricks and/or use Taylor series, as we shall see below.

Example:

Let us determine the Laurent series centered at 0 of the function

$$f(z) = \frac{1}{z(z-1)}$$

on three different domains.

a) On $D_1 : 0 < |z| < 1$:

On this domain, f is holomorphic outside $z_0 = 0$.

The ratio $\frac{1}{z-1}$ is holomorphic on D_1 , hence it can be written as the sum of a Taylor series (actually, of a geometric series) around 0:

$$\frac{1}{z-1} = -\sum_{n=0}^{\infty} z^n,$$

and we get

$$f(z) = -\frac{1}{z} \sum_{n=0}^{\infty} z^n = -\frac{1}{z} - 1 - z - z^2 - \dots,$$

that is, $c_{-1} = -1$, $c_n = -1$ for $n \geq 0$ and $c_m = 0$ for $m < 0$.

a) On $D_2 : 1 < |z| < \infty$:

On this domain, we have

$$\frac{1}{z-1} = \frac{1}{z} \frac{1}{1 - \frac{1}{z}} = \frac{1}{z} \sum_{n=0}^{\infty} \frac{1}{z^n},$$

since $\left| \frac{1}{z} \right| < 1$, and

$$f(z) = \frac{1}{z(z-1)} = \frac{1}{z^2} \sum_{n=0}^{\infty} \frac{1}{z^n} = \frac{1}{z^2} + \frac{1}{z^3} + \frac{1}{z^4} + \dots$$

Exercise:

Find the Laurent series for the functions

a) $f(z) = \frac{\sin z}{z}$; b) $f(z) = e^{\frac{1}{z}}$, on domains which contain $z_0 = 0$.

9.2 Residues

A point $z_0 \in \mathbb{C}$ is a *singular point* of a function f , if f is not holomorphic at z_0 , but it is holomorphic at some point in each neighborhood of z_0 .

If z_0 is a singular point of f such that f is holomorphic on a punctured disk $0 < |z - z_0| < \varepsilon$, then z_0 is called an *isolated singular point*.

Examples:

1) $f(z) = \frac{1}{z(z-1)}$ has two isolated singular points, namely $z_1 = 0$, $z_2 = 1$.

2) $f(z) - \text{Log}(z)$ has as singular points all the points of the negative Oy -axis.

For *isolated* singular points, there exists the Laurent series centered at z_0 of the function f , on $D : 0 < |z - z_0| < R$, for some $R > 0$

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n. \quad (39)$$

A peculiar importance has the coefficient c_{-1} in the above. Namely, we have

$$c_{-1} = \frac{1}{2\pi i} \int_{\gamma} f(s) ds, \quad (40)$$

where γ is a simple, closed, positively oriented curve contained in D such that z_0 lies inside γ . Applications obviously refer to the calculation of integrals complex functions on closed contours containing isolated singularities.

Definition 48 The coefficient c_{-1} in (39) is called the **residue** of f at z_0 and is denoted by

$$c_{-1} = \operatorname{Res}_{z=z_0}(f).$$

Example: Calculate $\int_{|z|=1} e^{1/z} dz$.

The contour $|z| = 1$ encloses the isolated singularity $z_0 = 0$.

We have $e^w = 1 + w + \frac{w^2}{2!} + \dots + \frac{w^n}{n!} + \dots$

Let us set, for $z \neq 0$, $w = \frac{1}{z}$. We get

$$e^{1/z} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \dots + \frac{1}{n!z^n} + \dots,$$

which is nothing but the Laurent series of $e^{1/z}$ centered at 0, on a domain $0 < |z - z_0| < R$, for some $R > 0$. The residue c_{-1} of f at 0 is $c_{-1} = 1$, hence

$$\int_{|z|=1} e^{1/z} dz = 2\pi i c_{-1} = 2\pi i.$$

Let us suppose now that f is a holomorphic function on the interior of a simple, positively oriented curve C , except some isolated singularities $z_1, \dots, z_n$. Let $C_k = \mathcal{C}(z_k, \varepsilon_k)$ be circles centered at z_k , on which f is holomorphic. Then

$$\int_C f(z) dz = \int_{C_1} f(z) dz + \int_{C_2} f(z) dz + \dots + \int_{C_n} f(z) dz.$$

But, from (40), $\int_{C_k} f(z) dz = 2\pi i \operatorname{Res}_{z=z_k}(f)$. We have thus proved

Theorem 49 (Residue Theorem): Let f be a holomorphic function on the interior of a simple, positively oriented curve C , except some isolated singularities $z_1, \dots, z_n$. Then, the integral of f over C is equal to $2\pi i$ times the sum of residues at singular points inside C :

$$\int_C f(z) dz = 2\pi i \sum_{k=1}^n \text{Res}(f)_{z=z_k}.$$

Exercise: Calculate $\int_{\gamma} \frac{1}{z(z-1)} dz$, where γ is

- a) the disk $D_1 : |z| < \frac{1}{2}$;
- b) the disk $D_2 : |z| < 2$.

9.3 Classification of Isolated Singularities

Let z_0 be an isolated singular point for a function f , and

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n$$

be the Laurent series centered at z_0 .

1. The point z_0 is called a *removable singularity* if $c_j = 0$ for all negative j .

At removable singular points, the Laurent series is actually a Taylor series: $f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n$. By setting $f(z_0) = c_0$, the function becomes holomorphic at z_0 (the singularity can be "removed").

2. If

$$f(z) = c_{-n}(z - z_0)^{-n} + \dots + c_{-1}(z - z_0)^{-1} + c_0 + c_1(z - z_0) + \dots + c_n(z - z_0)^n + \dots$$

(the negative power part of the Laurent series has a finite number n of terms), then z_0 is called a *pole* of order n .

An isolated singularity z_0 is a pole of order n if and only if $(z - z_0)^n f(z)$ is holomorphic on a domain containing z_0 .

3. If the Laurent series of f contains infinitely many terms in its negative part, then z_0 is called an *essential singularity*.

Examples:

- 1) for $f(z) = \frac{\sin z}{z}$, $z \in \mathbb{C} \setminus \{0\}$, the Laurent series centered at 0 is

$$\frac{\sin z}{z} = \frac{1}{z} \left(\frac{z}{1!} - \frac{z^3}{3!} + \dots \right) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots$$

The series is Taylorian, hence $z_0 = 0$ is a removable singularity for $\frac{\sin z}{z}$. Indeed, by setting $f(0) = 1$, the function becomes entire.

2) for $f(z) = \frac{1}{z^3(z-1)}$, the point $z_0 = 0$ is a pole of order 3, since $z^3 f(z)$ is holomorphic on some domain containing z_0 . Point $z_1 = 1$ is a pole of order 1.

3) For $f(z) = e^{1/z} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \dots + \frac{1}{n!z^n} + \dots$, the origin is an essential singularity.

At poles, there exists an elegant "shortcut" for computing residues. Namely, let z_0 be a pole of order n for a function f :

$$f(z) = \frac{c_{-n}}{(z-z_0)^n} + \dots + \frac{c_{-1}}{z-z_0} + c_0 + c_1(z-z_0) + \dots$$

Multiplying both sides of the equality by $(z-z_0)^n$, we get a holomorphic function:

$$\phi(z) := (z-z_0)^n f(z) = c_{-n} + c_{-n+1}(z-z_0) + \dots + c_{-1}(z-z_0)^{n-1} + c_0(z-z_0)^n + \dots$$

The residue c_{-1} of f at z_0 none but the coefficient of $(z-z_0)^{n-1}$ in the Taylor series of ϕ , that is

$$\operatorname{Res}_{z=z_0}(f) = \frac{\phi^{(n-1)}(z_0)}{(n-1)!}.$$

Example:

Find the residue of $f(z) = \frac{e^z}{z^2(z^2+1)}$ at $z_0 = 0$.

We have that $\phi(z) = z^2 f(z)$ is holomorphic on a neighborhood of $z_0 = 0$. That is, $n = 2$ and

$$\operatorname{Res}_{z=0}(f) = \frac{\phi'(0)}{1!}.$$

Actually, we have $\phi(z) = \frac{e^z}{z^2+1}$, hence

$$\phi'(0) = \frac{e^z(z^2+1) - 2ze^z}{(z^2+1)^2} \Big|_{z=0} = 1,$$

which means that $\operatorname{Res}_{z=0}(f) = \frac{\phi'(0)}{1!} = 1$.

Exercise: Calculate the residues of the same function at $z = i$ and $z = -i$.

Список литературы

- [1] Фучс Б.А., Шабат Б.В. Функции комплексного переменного и некоторые их приложения. Наука, Москва, 1964.
- [2] Лаврентьев М.А., Шабат Б.В. Проблемы гидродинамики и их математические модели. Наука, Москва, 1973, <http://hypercomplex.xpsweb.com/page.php?lang=ru&id=144>.
- [3] Cain G. Complex Analysis, 2001, <http://www.math.gatech.edu/~cain/winter99/complex.html>

- [4] Beck M., Marchesi G., Pixton D. A First Course in Complex Analysis. 2002,
<http://www.math.binghamton.edu/dennis/complex.pdf>,
<http://math.sfsu.edu/beck/complex.html>
- [5] Urroz G E. Complex variables applied to two-dimensional potential flows.
[http://www.neng.usu.edu/cee/faculty/gurro/Classes/
Classes_Fall2001/CEE6580/CEE658_FluidsV.pdf](http://www.neng.usu.edu/cee/faculty/gurro/Classes/Classes_Fall2001/CEE6580/CEE658_FluidsV.pdf) .

Современные проблемы ОТО – физика и геометрия

С.В. Сипаров

Санкт-Петербургский государственный университет гражданской авиации, Санкт-Петербург, Россия.

E-mail: sergey@siparov.spb.su

Предлагаемый материал является вспомогательным и представляет собой основу лекций для слушателей Школы по Финслеровой геометрии, т.е. студентов старших курсов математических (реже физических) специальностей университетов. Предполагается, что читатель, как и эти студенты, обладает определенной математической подготовкой, т.е. знаком с формальной стороной вопроса, но, возможно, не обсуждал в подробностях, как и с какой целью некоторые понятия и принципы использовались в естествознании и теоретической физике, а также к каким проблемам это может привести или уже привело. Целью этого изложения является привлечение заинтересованной молодежи в область, где на современном этапе развития науки наблюдается кризис, и обсуждение некоторых возможных путей его преодоления. Это требует по возможности ясного понимания основ теории, одним из которых уже полтора века, а другим – почти сто лет, и которым будет уделено определенное внимание.

1 Введение: взаимоотношение математики и физики

1.1 Наука от Пифагора до Кеплера и Ньютона

Две основные концепции современного естествознания – квантовая механика и теория относительности – формализованы чрезвычайно глубоко. И эта глубина носит не технический, т.е. просто математически сложный, но мировоззренческий, философский характер. Не уделяя внимания этому обстоятельству, невозможно выйти за пределы конкретных задач, какими бы важными они ни были, невозможно предположить, в чем может состоять фундаментальное заблуждение, если вдруг целый класс задач оказывается не решаемым или целый пласт наблюдений начинает противоречить существующей теории. И выдающиеся деятели науки этому обстоятельству внимания уделяли всегда. Странность некоторых математических идей и понятий прорывалась в мир наблюдений и измерений, а результаты последних заставляли строить неведомые дотоле умозрительные конструкции.

Если считать Пифагора одним из родоначальников науки, то можно вспомнить, как от простого “египетского” треугольника он сделал шаг к половине разрезанного по диагонали квадрата и обнаружил, что для измерения этой диагонали не находится *числа* в привычном тогда понимании этого слова. Это открытие так потрясло его, что за разглашение этой тайны, вывод ее за пределы круга соратников-“пифагорейцев” человек подлежал смертной казни. Утверждение “все есть число”, приписываемое Пифагору, имеет глубокий смысл. Различные *числоподобные объекты* (числа натуральные, рациональные, иррациональные, комплексные, двойные, кватернионы и другие) образуют различные *алгебры*, на которых могут быть построены различные *геометрии*, а от геометрии до *мира*, в котором происходят наблюдаемые

и измеряемые *события*, один шаг. И хотя живем мы именно в мире, в основе мироздания оказывается число или иной абстрактный математический объект, например, такой, как сфера – самая совершенная математическая фигура.

Вот и в небе, окружающем Землю, оказываются сферы – у каждой планеты, вращающейся вокруг Земли, своя, а самая большая, всеобъемлющая (между прочим, седьмая, поскольку планет было тогда известно четыре, да еще Солнце с Луной) сфера – это сфера, на которой расположены звезды, образующие не меняющиеся со временем созвездия. Это – абсолют, это – гармония, и никакое наблюдаемое явление не должно поколебать идеал – геометрическое место точек, равноудаленных от одной. Но как быть с наблюдаемым попятным движением планет, перемещающихся среди неизменных рисунков созвездий? Не могут же сферы вращаться не равномерно! И Птолемею приходится помимо “дифферентов” – окружностей, по которым планеты двигаются вокруг Земли, вводить еще “эпициклы” – маленькие окружности, по которым планеты двигаются вокруг точек, как раз и двигающихся по “дифферентам”. А когда были выполнены более точные измерения, потребовалось и все плоскости “дифферентов” и “эпициклов” наклонить под разными углами. Расчет положения планет превратился в зубодробительное упражнение, которое все-таки позволяло найти ответ наиболее упорным исследователям. Лишь через несколько столетий проявилась систематическая ошибка, которая постепенно нарастала со временем.

Однако, когда Коперник “вывернул наизнанку” геоцентрическую систему Птолемея и поместил в центр Солнце, а Земля, как и другие планеты, стала двигаться по круговой орбите, вращаясь вокруг собственной оси, расчет положения планет по этой упрощившейся модели, совсем не требовавшей эпициклов, стал давать еще более плохой результат, чем сложная модель Птолемея. Поэтому революция Коперника еще долго оставалась просто остроумной гипотезой, вызывающей ободрающие улыбки, но вряд ли имеющей отношение к действительности.

В свою очередь Кеплер, анализирувавший результаты двадцатилетних измерений положения планет, выполненных Тихо Браге и им самим (для повышения точности гороскопов), тоже был уверен в мировой гармонии, основанной на математике. Он надеялся, что те окружности, по которым планеты, вероятно, двигаются вокруг Солнца, имеют радиусы, соответствующие концентрическим сферам, в которые последовательно вписаны правильные Платоновы многогранники. И когда он обнаружил, что это не так, а планеты вообще двигаются по эллипсам, то пережил культурный шок, от которого долго не мог оправиться. Но зато эллипсы, измеренные Кеплером, блестяще подошли к расчетам, основанным на гипотезе Коперника, устранили имевшиеся расхождения, и мир, наконец, стал гелиоцентрическим. Закон, связывающий радиусы орбит и периоды обращения планет, обнаруженный Кеплером в результате наблюдений и измерений, впоследствии привел Ньютона к закону гравитации в том виде, в котором мы сейчас его знаем.

С Ньютона началась новая эра в развитии естествознания. Он отказался следовать априорным математическим утверждениям, основанным на “наивном” стремлении к гармонии, сделал своим девизом “гипотез не измышляю” и прочно внедрил в науку метод индукции. И когда измерения Кеплера и полученный им закон, использованные Ньютоном при определении отношений ускорений планет, привели к мысли о существовании дальнедействующего притяжения планет к Солнцу по закону обратных квадратов, Ньютон на двадцать лет положил в стол свой результат, поскольку такая “гипотеза” казалась ему неприемле-

мой. “Мысль о том, чтобы способность возбуждать тяготение могла быть неотъемлемым, внутренне-присущим свойством материи, и чтобы одно тело могло воздействовать на другое через пустоту на расстоянии, без участия чего-то такого, что переносило бы действие и силу от одного к другому – представляется мне столь нелепой, что нет, как я полагаю, человека, способного мыслить философски, кому она пришла бы в голову,” - писал Ньютон в письме к Беркли. И когда, скрепя сердце, он все же дал себе труд ознакомить общественность со своими достижениями, у него не было никаких оснований допустить, что это “воздействие через пустоту” передается не мгновенно, а сама пустота, т.е. то “вместилище”, где все это происходит, не описывается Евклидовой геометрией.

Таким образом, вплоть до XVIII века математика и естествознание, “попеременно лидируя”, участвовали в разработке все более точной картины мира, позволяющей предсказывать будущее. Время было в этой картине той четвертой координатой, взглянуть на различные значения которой с помощью только органов чувств не представлялось возможным. Необходим был инструмент, и им была математика, как и любой инструмент, искажающая картину по не всегда известным правилам.

1.2 Кривизна, силы, поля

В конце XVIII века К.Гаусс стал первым, кто конструктивно подошел к вопросу о применимости геометрии Евклида для описания мира и непосредственно измерил на местности сумму внутренних углов треугольника. Вершины треугольника располагались на вершинах близлежащих гор. Гаусс не обнаружил отклонения геометрии мира от Евклидовой в пределах точности своих измерений.

В начале XIX века Н.Лобачевский принял во внимание и оценил принципиальные возможности астрономических наблюдений, что с неизбежностью привело его к созданию первой неевклидовой геометрии, применимость которой на астрономических масштабах не могла быть исключена.

Во второй половине XIX века Риман разработал дифференциальную геометрию, теорию поверхностей, в которой использовались такие понятия, как кривизна (пропорциональная второй производной смещения по безразмерному параметру) и метрика (связанная со скалярным произведением).

В физике в это время изучались силы Кулона, Лоренца и Ампера, которые, как и гравитация, считались мгновенно дальнodelствующими. Электромагнитные явления были описаны Максвеллом на основании нового подхода – перехода от Ньютоновской теории действующих сил и соответствующих уравнений динамики к теории поля, т.е. к связям между локальными характеристиками среды, к соотношениям, связывающим между собой напряженности и потенциалы сил в точках. Эта формулировка (этот формализм) позволила Максвеллу теоретически предсказать, а затем Герцу экспериментально обнаружить электромагнитные волны, которые на основании опытов Фарадея оказалось возможным отождествить со световыми волнами.

В середине XIX века В. Клиффорд уже последовательно защищал свою мысль о том, что “все проявления физического мира экспериментально неотличимы от соответствующих изменений геометрической кривизны мира”. Простейший пример – разумный червь, способный перемещаться внутри замкнутого одномерного канала, в котором он и живет. Если канал имеет постоянную кривизну (имеет форму окружности), то органы чувств червя ничего не

сообщают червя при его движении. Если же кривизна переменна, то в местах ее изменений форма червя при его движении будет меняться, и его органы чувств испытают воздействие. Таким образом, было установлено, что связь математики (геометрии) и физики, хотя и не носит трансцендентного характера, как было до Ньютона, отказавшегося измышлять гипотезы, но все же является, с одной стороны, взаимной и тесной (неотличимость кривизны и “физических” феноменов), а с другой – достаточно произвольной, зависящей от уровня и глубины наших знаний в настоящий момент. Действительно, Евклидова геометрия пригодна для описания происходящего на Земле, но, как показал Лобачевский, на существенно больших масштабах, не доступных непосредственному наблюдению и исследованию, оснований для выбора именно ее пока нет.

Следуя этим идеям, А.Эйнштейн в начале XX века свел общее, но качественное утверждение Клиффорда к утверждению более узкому, но зато поддающемуся количественному анализу. Основываясь на требовании общей ковариантности физических законов (и постулировав постоянство скорости света и принцип эквивалентности инертной и гравитационной масс), он разработал теорию, утверждающую, что неустранимая в нашем мире гравитация неотличима от проявлений геометрических свойств пространства-времени. Считая, что пространство-время описывается геометрией Римана-Минковского, Эйнштейн в своей основополагающей работе привел пример наблюдаемых эффектов, которые должны иметь место в этом случае и дал их расчет. Опыты показали хорошее соответствие с расчетом, и новая геометрия полноправно вошла в физическую теорию. На макроуровне это позволило, например, уточнить результаты теории гравитации Ньютона в применении к Солнечной системе. На микроуровне в квантовую механику вошла теория Дирака. На мегауровне в космологию вошла теория расширяющейся Вселенной и сопутствующих обстоятельств.

Говоря о данных наблюдений, с которыми должна согласовываться теория, следует указывать масштаб рассмотрения. В астрофизике таких масштабов можно указать три: Солнечная система (окрестность точечной массы), галактика (10^{11} точечных масс, удаленных от остальных) и метagalactika (т.е. вся наблюдаемая Вселенная, состоящая из галактик). Известные достижения общей теории относительности (ОТО), связанные с введением новой геометрии (Римана), позволили связать лабораторную физику и первый из упомянутых масштабов. К концу XX века были накоплены астрофизические данные, которые на уровне современных представлений не удастся объяснить без включения в теорию довольно загадочных новых понятий, например, темной материи или темной энергии, предположительно заключающих в себе 95% всей массы Вселенной, или без пересмотра основ самой теории, включая и геометрию пространства-времени.

1.3 Выводы

1. Принципиальная возможность применения умозрительных математических конструкций для описания мира, воспринимаемого с помощью органов чувств, неочевидна и является постулатом. Каждая из этих процедур – построение конструкций и описание мира – является самодостаточной. Указанная возможность содержит произвол.
2. С одной стороны, известна “непостижимая эффективность математики в естествознании”. С другой стороны, систематизированные данные наблюдений указывают на направления нетривиальных математических исследований. Именно эти обстоятельства являются одновременно мотивацией и регулятором деятельности, известной как теоретиче-

ская физика.

3. Иногда для эквивалентного описания одних и тех же явлений может быть использовано несколько математических подходов. Как правило, сообществом признается лишь первый из них, все остальные отсекаются “бритвой Оккама”.
4. Направления поиска новой теории определяются затруднениями старой.
5. Никакой новой теории, какой бы красивой она ни была и как бы успешно ни справилась с затруднениями старой, не удастся заместить предыдущую физическую теорию, пока не состоятся наблюдения, соответствующие предсказанным ей специфическим результатам.

Дальнейшее изложение построено так. Для того, чтобы понять происхождение проблем ОТО, возникших в настоящее время, и предложить возможные пути их преодоления, проследим за построением классической ОТО, касаясь не столько формальной стороны дела, сколько обсуждения принципов, положенных в ее основание и вытекающих из этого следствий. Затем опишем сами проблемы и обсудим возможные изменения теории, которые можно совершить для объяснения наблюдаемых явлений. Наконец, следует указать и дать расчет явлений, специфических для возможной модификации основ теории.

2 Некоторые геометрические аспекты общей теории относительности Эйнштейна

Общая теория относительности Эйнштейна (ОТО) в оригинальной статье называлась “обобщенная теория относительности и гравитации”. В предшествовавшей работе Эйнштейн фактически проанализировал понятие “измерение” и пришел к выводу, что, пока мы используем световые лучи для проведения измерений с линейками и часами, следует полагать, что

- скорость света равна одной и той же величине, из какой бы (инерциальной) системы отсчета (ИСО) она ни была измерена. (A1)

Этот вывод, известный как постулат Эйнштейна о постоянстве скорости света, позволил дать глубокую физическую интерпретацию следующему математическому обстоятельству. В псевдоевклидовом пространстве Минковского, где квадрат расстояния между точками определяется как

$$ds^2 = dx_1^2 - dx_2^2 - dx_3^2 - dx_4^2, \quad (1)$$

кривую можно задать с помощью так называемого естественного параметра t , который можно связать с координатами с помощью размерного переводного коэффициента. Выбирая в качестве этого параметра “время”, т.е. переводным постоянным коэффициентом будет некоторая “фундаментальная скорость”, и сопоставляя этой “скорости” значение скорости света, постоянное при любых физических измерениях, мы получим возможность интерпретировать окружающий мир как пространственно-временной континуум, в котором могут проявиться математические свойства, характерные для геометрического пространства Минковского. Причина, по которой было выбрано именно пространство Минковского, состояла в интуитивном (привычном) стремлении вычислять расстояние между точками с помощью разностей их координат, используя теорему Пифагора, и сопоставлять его произведению “фундаментальной скорости” на “естественный параметр”. И если считать постулат A1 выполненным,

то во всех ИСО *интервал*, т.е. разность (1) между указанными двумя величинами будет одной и той же, и принцип относительности Галилея превращается в принцип относительности Эйнштейна, и уже не только механическими, но

- никакими опытами невозможно различить покой и равномерное прямолинейное движение системы отсчета. (A2)

Таким образом, физика проникла в геометрию, которая в этой связи перестала быть Евклидовой. А геометрия и сопутствующий математический аппарат приобрели широкие возможности для разнообразных физических интерпретаций. Именно в этой связи преобразования Галилея уступили место преобразованиям Лоренца, а опыт Майкельсона-Морли получил наиболее фундаментальное обоснование из всех, встречавшихся в науке до этого – длительность временного промежутка стала различаться при измерениях в системах отсчета, движущихся относительно друг друга.

Когда впоследствии Эйнштейн перешел к обобщению своих результатов на ускоренные системы отсчета и отождествил инерцию и гравитацию, поскольку не видел способа различить характеризующие их коэффициенты (инертную и гравитационную массы), скорость света или фундаментальная скорость, естественно, осталась постоянной (и конечной), что позволило все глубже увязывать между собой математику и физику. При этом и в этой области обнаружилась “непостижимая эффективность математики в естествознании” (Вебер). Позже Дж.Уилер с полным основанием назвал ОТО геометродинамикой.

2.1 Тензоры

Напомним некоторые понятия, важные для дальнейшего. Основной математический объект ОТО – тензор. Свойства тензора:

- Существует (в виде математического определения, т.е. описания набора свойств)
- Не наблюдаем (как таковой в физическом мире)
- Имеет представителей (компоненты), которые иногда могут быть сопоставлены наблюдаемым и измеряемым величинам
- При изменении локальной системы координат эти компоненты должны изменяться по строго определенным правилам (ко- или контравариантно изменениям ортов базиса), что при анализе измерений и может указывать на существование и присутствие в ситуации тензора
- Использование тензоров для записи законов физики позволяет сформулировать их способом, не зависящим от выбора системы координат, – **ковариантным** образом.

Именно это последнее обстоятельство и привлекает внимание, поскольку у нас нет возможности узнать, движется ли и как та система отсчета, в которой мы работаем и пытаемся предсказать поведение исследуемых объектов.

2.1.1 Фундаментальный тензор

Существует так называемый фундаментальный тензор g_{ik} (дважды ковариантный, т.е. матрица второго ранга), позволяющий “переходить” от i -той контравариантной компоненты тензора к k -той ковариантной компоненте тензора (“опускать значок”). Фундаментальный тензор строится из пар единичных векторов канонического базиса $\vec{e}_i \vec{e}_k = g_{ik}$. Манипулируя

с этой матрицей можно построить и дважды контравариантный фундаментальный тензор, осуществляющий обратный переход (“подъем значка”). Подчеркнем, что ко- и контравариантность являются именно теми фундаментальными свойствами компонент тензора, которые и позволяют осуществить их интерпретацию как величин, наблюдаемых в физическом мире.

2.1.2 Норма, метрика, геометрия

Тензоры обладают численной характеристикой – нормой, – значение которой можно сопоставлять с вводимым определением метрики пространства, т.е. с определением скалярного произведения, иными словами, с выбором геометрии.

Для вычисления нормы вектора в ОТО используется квадратичная метрика, получаемая с помощью фундаментального тензора

$$||x||^2 = g_{\alpha\beta} x^\alpha x^\beta \quad (2)$$

Выбор метрики для вычисления нормы определяет выбор геометрии. См. некоторые подробности в Приложении 1.

2.1.3 Построение тензоров

Тензоры можно конструировать друг из друга по определенным правилам.

Например:

ϕ – скаляр $\Rightarrow A_\mu = \frac{\partial \phi}{\partial x^\mu}$ – ковариантный четырехмерный вектор $\Rightarrow$

$A_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \Gamma_{\mu\nu}^\sigma A_\sigma$ – ковариантный тензор второго ранга (ковариантная производная),

где $\Gamma_{\mu\nu}^\sigma = \frac{1}{2} g^{\sigma\alpha} \left\{ \frac{\partial g_{\mu\alpha}}{\partial x^\nu} + \frac{\partial g_{\nu\alpha}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \right\}$ – так называемые коэффициенты связности (символы Кристоффеля).

Можно убедиться, что сами коэффициенты связности тензорами не являются, т.е. их компоненты при изменении системы координат, хотя и изменяются, но не ковариантным, ни контравариантным образом. Следует иметь в виду, что при использовании тензорного математического аппарата для построения физической теории необходимо следить за тем, какой из рассматриваемых объектов является тензором, а какой нет. В частности, $\Gamma_{\mu\nu}^\sigma A_\sigma$ – уже тензор.

2.2 “Тождества Максвелла” и уравнения Максвелла

Убедимся в “непостижимой эффективности математики в естествознании”, упомянутой выше. Выполним формальные построения

2-1) Пусть φ_ν – компоненты ковариантного 4-вектора. Построим тензор

$$F_{\rho\sigma} = \frac{\partial \varphi_\rho}{\partial x^\sigma} - \frac{\partial \varphi_\sigma}{\partial x^\rho}.$$

Тогда тождественно выполняется

$$\frac{\partial F_{\rho\sigma}}{\partial x^\tau} + \frac{\partial F_{\sigma\tau}}{\partial x^\rho} + \frac{\partial F_{\tau\rho}}{\partial x^\sigma} = 0,$$

где левая часть – антисимметричный тензор 3-го ранга, имеющий 6 независимых компонент. Поэтому можно переписать в явном виде

$$\begin{aligned}\frac{\partial F_{23}}{\partial x^4} + \frac{\partial F_{34}}{\partial x^2} + \frac{\partial F_{42}}{\partial x^3} &= 0 \\ \frac{\partial F_{34}}{\partial x^1} + \frac{\partial F_{41}}{\partial x^3} + \frac{\partial F_{13}}{\partial x^4} &= 0 \\ \frac{\partial F_{41}}{\partial x^2} + \frac{\partial F_{12}}{\partial x^4} + \frac{\partial F_{24}}{\partial x^1} &= 0 \\ \frac{\partial F_{12}}{\partial x^3} + \frac{\partial F_{23}}{\partial x^1} + \frac{\partial F_{31}}{\partial x^2} &= 0\end{aligned}\tag{3}$$

Введем формально новые обозначения

$$\begin{aligned}F_{23} &\equiv B_x; F_{14} \equiv E_x \\ F_{31} &\equiv B_y; F_{24} \equiv E_y \\ F_{12} &\equiv B_z; F_{34} \equiv E_z\end{aligned}\tag{4}$$

Подставляя (4) в (3), получаем в привычных обозначениях векторного анализа тождества, следующие из введенных определений геометрических объектов

$$\begin{aligned}\frac{\partial \vec{B}}{\partial t} + \text{rot } \vec{E} &= 0 \\ \text{div } \vec{B} &= 0\end{aligned}\tag{5}$$

2-2) Пусть I^α - контравариантный вектор. Как и ранее, построим соответствующий ему контравариантный 6-вектор $F^{\alpha\beta}$ и вычислим его дивергенцию $I^\alpha = \frac{1}{\sqrt{-g}} \frac{\partial(\sqrt{-g} F^{\alpha\beta})}{\partial x^\beta}$, где $g = \det g_{\alpha\beta}$. Если $\sqrt{-g} = 1$, то получим

$$I^\alpha = \frac{\partial F^{\alpha\beta}}{\partial x^\beta}\tag{6}$$

Вновь введем формально новые обозначения

$$\begin{aligned}F^{23} &\equiv B_x; F^{14} \equiv -E_x \\ F^{31} &\equiv B_y; F^{24} \equiv -E_y \\ F^{12} &\equiv B_z; F^{34} \equiv -E_z\end{aligned}\tag{7}$$

к которым добавим еще и такие

$$I^1 \equiv j_x; I^2 \equiv j_y; I^3 \equiv j_z; I^4 \equiv \rho\tag{8}$$

Подставляя (2.5-2.6) в (6), получим в привычных обозначениях векторного анализа

$$\begin{aligned}\text{rot } \vec{B} - \frac{\partial \vec{E}}{\partial t} &= \vec{j} \\ \text{div } \vec{E} &= \rho\end{aligned}\tag{9}$$

Формальные уравнения (5) и (9), фактически, – геометрические тождества – отличаются от физических уравнений Максвелла, обобщающих *опытные данные электродинамики*, только тем, что введенным величинам был приписан некоторый физический смысл (сопоставлены

данные измерений) и введены размерные постоянные множители. Можно отметить, что во времена Максвелла (середина XIX в.) обсуждение математических (алгебраических) основ физики было весьма популярно, в частности, введенные Гамильтоном кватернионы были “на слуху” у теоретиков, и Максвелл в своем “Трактате об электричестве” неоднократно подчеркивал смысл и особенности введенных им привычных нам *векторов* – столбцов соответствующих матриц.

2.3 Вариационный принцип, простейший скаляр, геодезическая, принцип эквивалентности

В некоторых частных случаях при перемножении контравариантного и ковариантного векторов (или других соответствующих тензоров одинакового ранга) полученное произведение является скаляром, т.е. объектом, не зависящим от выбора системы координат в связи с определениями ко- и контравариантности (см. также Приложение 1). Вообще говоря, и другие геометрические объекты – вектор, тензор и т.п. – сами по себе не зависят от выбора системы координат, но вот их представители (компоненты) – зависят. Скаляр в этом смысле занимает особое место, и поэтому именно он используется при формулировке так называемого вариационного принципа. В математике вариационный принцип есть основное положение вариационного исчисления – раздела, в котором исследуются свойства экстремалей, а в физике он представляет собой некоторый формальный постулат, оказавшийся чрезвычайно плодотворным. Выбор скаляра для проведения соответствующих манипуляций может быть объектом обсуждения. Эйнштейн полагал, что скаляр должен быть выбран в “простейшей” форме, однако указать критерий необходимой простоты такой, что используемый скаляр приведет к содержательным результатам, затруднительно.

Вариационный принцип устанавливает соответствие между геометрическими объектами и их физическими аналогами по следующей схеме.

Геометрия: кривая, являющаяся экстремалью функционала длины, называется геодезической.

Замечание 1: В Евклидовом пространстве длина определяется функционалом

$$l = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$$

Замечание 2: Экстремаль функционала $S = b \int_{t_1}^{t_2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) dt$ совпадает с экстремалью функционала l .

Физика: выбор $b = m/2$, где m – масса, доставляет выражению $\frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ физический смысл.

Принцип наименьшего действия (Гамильтон):

Принцип наименьшего действия представляет собой постулат, согласно которому

- движения (механической) системы совпадают с экстремальми функционала “действия”

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt, \text{ где } L \text{ есть разность кинетической и потенциальной энергий системы.}$$

Следствие: варьирование функционала действия приводит к уравнениям Эйлера-Лагранжа, которые представляют собой уравнения динамики

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

Более подробно принцип наименьшего действия, его связь с так называемыми “законами физики” и соответствующими уравнениями обсуждаются в Приложении 2.

Принцип эквивалентности

Приступая к исследованию проблемы гравитации, Эйнштейн намеревался исследовать точку зрения Маха на происхождение и природу инертной массы. Мах полагал, что инерция обусловлена влиянием (гравитационным притяжением) “удаленных звезд”. Однако, обнаружив впечатляющее сходство уравнения геодезической с уравнением движения в гравитационном поле, Эйнштейн отошел от обсуждения вопроса о природе массы, поскольку оказалось, что он может и не иметь большого значения, если положить инертную и гравитационную массу эквивалентными и равными по величине. С точки зрения физики это означало, что не существует такого опыта, с помощью которого можно обнаружить различие инертной и гравитационной масс. Иными словами,

- ускорение системы отсчета и гравитационное поле эквивалентны. (А3)

Здесь стоит вспомнить знаменитый “лифт Эйнштейна” – закрытое помещение без окон, о движении которого наблюдатель, находящийся внутри, может судить, только выполняя некоторые эксперименты. Так, если лифт движется равноускоренно с ускорением $g = 9,82$ м/с², то результаты любых опытов не позволят установить, движется ли этот лифт, либо покоится, – но в однородном поле тяжести (вблизи поверхности Земли). Попытки обнаружить различие величин инертной и гравитационной масс начались опытами Этвеша, впоследствии усложнялись и даже включали в себя опыты на специально запущенном спутнике в рамках программы STEP. В настоящее время можно утверждать, что эти величины равны с точностью не ниже, чем 10^{-12} . Тем не менее, этот вопрос является принципиальным и должен быть решен принципиально, например, путем введения постулата об их точном равенстве.

После введения этого постулата для построения теории с проверяемыми следствиями необходимо поступать так:

- Выберем геометрию для описания рассматриваемого пространства, т.е., в частности, дадим определение метрики. Обратим внимание на то, что в этом пункте присутствует выбор, и у него должны быть основания, что не всегда осознается и обсуждается.
- Введем функционал действия, получим и решим уравнения Эйлера-Лагранжа, соответствующие, с одной стороны, геодезической геометрического пространства, а с другой (в связи с принципом эквивалентности) – уравнению движения частицы в неустранимом гравитационном поле в физическом пространстве.
- Выберем геометрический скаляр, соответствующий функционалу действия. Этот выбор осознается чаще, причем говорят даже об *искусстве* выбора скаляра. Это связано с упоминавшейся проблемой “простоты”.

- Найдем (каким-то образом!) уравнения поля и исследуем частные случаи (предскажем и рассчитаем проверяемые эффекты).

2.4 Пример: построение основ. ОТО Эйнштейна (1916)

Проследим теперь, как был реализован этот алгоритм в классической работе Эйнштейна.

- В связи со стремлением опираться на постулат А1 о постоянстве скорости света во всех системах отсчета, что дает возможность содержательной физической интерпретации геометрических результатов, можно обобщить понятия геометрии Минковского и перейти к геометрии Римана, учитывающей зависимость метрики от координат рассматриваемой точки. Тогда квадрат элемента длины будет равен $ds^2 = g_{ik}(x)dx^i dx^k$. При этом, выбор функции $g(x)$ все еще содержит произвол, и естественное ее обобщение может предусматривать зависимость $g(x, \dot{x})$, т.е. от производной координаты точки по естественному параметру, – т.е. от направления. Это – путь к геометрии Финслера. Но Эйнштейн рассмотрел случай Римановой геометрии.
- Геодезическая имеет вид $\frac{d^2 x^i}{ds^2} + \Gamma_{kl}^i \frac{dx^k}{ds} \frac{dx^l}{ds} = 0$. Следовательно, уравнение движения можно записать $\frac{d^2 x^i}{ds^2} = -\Gamma_{kl}^i \frac{dx^k}{ds} \frac{dx^l}{ds}$, т.е. производные метрического тензора, входящие в символы Кристоффеля, могут представлять собой компоненты напряженности гравитационного поля
- Выбор скаляра:

0.1 С учетом принципа Гамильтона (см. Приложение 2) хотелось бы варьировать энергию, и чтобы при этом она сохранялась. Последнее требование означает, что надо, чтобы обращалась в ноль дивергенция $\frac{\partial T^{ik}}{\partial x^k} = 0$, где T^{ik} – *некоторый* тензор энергии-импульса, определяемый как

$$T_i^k = \sum_l q_{,i}^{(l)} \frac{\partial \Lambda}{\partial q_{,k}^{(l)}} - \delta_i^k \Lambda \quad , \quad (10)$$

, где $S = \int \Lambda(q, \frac{\partial q}{\partial x^i}) dV dt$ – действие, Λ – плотность лагранжиана, q – обобщенные координаты, δ_i^k – символ Кронекера. Из требования равенства нулю дивергенции следует, что сохраняется некоторый вектор (4-импульс) $P_i = \text{const} \int T^{ik} dS_k$ где интегрирование – по гиперповерхности, охватывающей 3-мерное пространство. Из формулы (10) видно, что $T^{00} = \dot{q} \frac{\partial \Lambda}{\partial \dot{q}} - \Lambda$, т.е. ^{00} – плотность энергии, а величины $T^{0\alpha} = T^{\alpha 0}$ – компоненты плотности импульса, $T^{\alpha\beta} = \sigma^{\alpha\beta}$ – компоненты тензора напряжений и связаны с силой, действующей на площадку.

0.2 Например, для пылевидной материи $T^{ij} = \rho_0 u^i u^j$, при этом $T^{00} = \rho c^2$ – плотность энергии, $T^{0\alpha} = \rho c v^\alpha$ – компоненты плотности импульса, $T^{\alpha\beta} = \rho v^\alpha v^\beta$ – поток α -компоненты импульса в направлении β .

0.3 Теперь нужно составить соответствующий геометрический скаляр, т.е. такой, что и дивергенция его – ноль, и число его независимых компонент – такое же, как у T^{ik} , т.е. 10.

0.4 Только из g_{ik} и Γ_{kl}^i , входящих в геодезическую, скаляра не составить (см. замечание выше о тензорном характере используемых величин).

- 0.5 Может быть, подойдет скалярная кривизна $R = g^{ik} R_{ik} = g^{il} g^{km} R_{iklm}$, где $R_{iklm} = \frac{\partial \Gamma_{km}^i}{\partial x^l} - \frac{\partial \Gamma_{kl}^i}{\partial x^m} + \Gamma_{nl}^i \Gamma_{km}^n - \Gamma_{nm}^i \Gamma_{kl}^n$ - тензор кривизны?
- 0.6 Но у R имеется 20 независимых компонент (вместо нужных 10) и нет такой комбинации производных, чтобы потом получить (ср. с задачей в Приложении 3 ϑ) уравнение Пуассона в предельном Ньютоновском случае..
- 0.7 Тогда, может быть, использовать тензор Риччи R_{ik} ? Он симметричен, т.е. у него как раз 10 независимых компонент!
- 0.8 Составим выражение, содержащее R_{ik} , и воспользуемся теоремой Гаусса, чтобы получить такое выражение, что его дивергенция равна нулю, как и дивергенция тензора энергии-импульса: $G^{ik} = R^{ik} - \frac{1}{2} g^{ik} R$. Можно проверить, что $\frac{\partial G^{ik}}{\partial x^k} = 0$, т.к. дивергенция $\frac{\partial g^{ik}}{\partial x^k} = 0$.
- Теперь постулируем “уравнение гравитационного поля” в виде $R^{ik} - \frac{1}{2} g^{ik} R = -\kappa T^{ik}$ (или, что то же, $R^{ik} = -\kappa(T^{ik} - \frac{1}{2} g^{ik} T)$). Это уравнение называется уравнением Эйнштейна. Решая его, найдем компоненты g_{ik} и Γ_{kl}^i , подставим их в уравнение для геодезической и получим уравнения динамики, пригодные для описания движения тел.

2.5 Закон гравитации Ньютона как приближенное решение уравнений ОТО

Построив теорию, следует убедиться, что она включает в себя ранее полученные и работоспособные результаты или их аналоги, которые не хуже старых. В частности, закон тяготения Ньютона в масштабах Солнечной системы вполне применим, за исключением описания весьма тонких эффектов типа смещения перигелия орбиты Меркурия, измеренного Леверье и составляющего доли угловой секунды за столетие. Поэтому надо убедиться, что полученная геометрическая теория дает тот же результат.

Выберем сигнатуру метрики Минковского так, как это делал Эйнштейн, $\gamma_{ik} = \{-1, -1, -1, 1\}$, где временная компонента соответствует g_{44} , и рассмотрим слабое поле с метрикой $g_{ik} = \gamma_{ik} + \varepsilon_{ik}$, $\varepsilon_{ik} \ll 1$, которая стремится к γ_{ik} на бесконечности. Используем следующие приближения:

1. Во-первых, пренебрежем малыми величинами порядков, более высоких, чем величина ε_{ik} . Движение описывается уравнением геодезической

$$\frac{d^2 x^i}{ds^2} = -\Gamma_{kl}^i \frac{dx^k}{ds} \frac{dx^l}{ds}. \quad (11)$$

2. Во-вторых, будем считать, что движение медленное по сравнению со скоростью света, которую положим равной $c = 1$, т.е.

$$\frac{dx^1}{ds}, \frac{dx^2}{ds}, \frac{dx^3}{ds} \ll 1; \frac{dx^4}{ds} \approx 1 \quad (12)$$

Из предположения (1) следует, что $\Gamma_{kl}^i \ll 1$, тогда в уравнение геодезической (52) входят только $k = l = 4$, т.е.

$$\frac{d^2 x^i}{dt^2} = -\Gamma_{44}^i \quad (13)$$

где было использовано $ds = dx_4 = dt$.

Тогда, учитывая, что гравитационное поле квазистатично (12), пренебрежем в (13) производными по времени по сравнению с производными по координатам, т.е.

$$\frac{d^2 x^i}{dt^2} = -\frac{1}{2} \frac{\partial g_{44}}{\partial x^i}; i = 1, 2, 3 \quad (14)$$

Таким образом, видим, что $g_{44}/2$ действительно может быть сопоставлено гравитационному потенциалу, градиент которого пропорционален ускорению (т.е. силе согласно закону динамики Ньютона).

“Уравнение поля” Эйнштейна будет иметь вид

$$\frac{\partial \Gamma_{\mu\nu}^{\alpha}}{\partial x^{\alpha}} + \Gamma_{\mu\beta}^{\alpha} \Gamma_{\nu\alpha}^{\beta} = -\kappa (T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T) \quad (15)$$

В нашем приближении в правой части останется только $T_{44} = \rho = T$, а в левой – только первое слагаемое, которое примет вид $\frac{\partial}{\partial x^1} \Gamma_{\mu\nu}^1 + \frac{\partial}{\partial x^2} \Gamma_{\mu\nu}^2 + \frac{\partial}{\partial x^3} \Gamma_{\mu\nu}^3 - \frac{\partial}{\partial x^4} \Gamma_{\mu\nu}^4$. Полагая $\mu = \nu = 4$ и отбрасывая производные по времени, получим

$$-\frac{1}{2} \left(\frac{\partial^2 g_{44}}{\partial x_1^2} + \frac{\partial^2 g_{44}}{\partial x_2^2} + \frac{\partial^2 g_{44}}{\partial x_3^2} \right) = -\frac{1}{2} \Delta g_{44}.$$

Тогда уравнение (15) дает уравнение Пуассона

$$\Delta g_{44} = \kappa \rho, \quad (16)$$

которое вместе с (14) эквивалентно закону гравитации Ньютона. Зная решение уравнения Пуассона, получаем, таким образом,

$$g_{44} = 1 - \frac{a}{r} \quad (17)$$

где $a = \frac{\kappa M}{4\pi}$, $\kappa = 1,87 \cdot 10^{-27}$; $r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$.

Окончательно, видим, что в низшем приближении ОТО дает требуемый закон гравитации Ньютона. Теперь следует убедиться, что ОТО дает что-то еще, чего не предсказывает старая теория и что можно проверить измерениями.

2.6 Классические тесты ОТО (красное смещение, отклонение луча, прецессия орбиты)

Геометрический подход к динамике позволяет сделать утверждения, некоторые из которых можно проверить.

2.6.1 Геометрия.

Определитель фундаментального тензора $g = -1$. Непосредственной проверкой можно получить

$$\begin{aligned} g_{\rho\sigma} &= -\delta_{\rho\sigma} - \alpha \frac{x^\rho x^\sigma}{r^3}; \rho, \sigma = 1, 2, 3 \\ g_{\rho 4} &= g_{4\rho} = 0; \rho, \sigma = 1, 2, 3 \\ g_{44} &= 1 - \frac{a}{r}; r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} \end{aligned} \quad (18)$$

Тогда для единичного отрезка, расположенного параллельно оси Ox , вдоль которой действует гравитационное поле, выполняется

$$ds^2 = -1; dx^2 = dx^3 = dx^4 = 0 \Rightarrow -1 = g_{11}(dx^1)^2$$

А если отрезок находится непосредственно на оси Ox , то с учетом (18) $g_{11} = -1 - \frac{a}{r}$ и получим $dx = 1 - \frac{a}{2r}$, что соответствует сокращению длины. Если же отрезок расположен в поперечном направлении, то

$$ds^2 = -1; dx^1 = dx^3 = dx^4 = 0; x^1 = r; x^2 = x^3 = 0 \Rightarrow -1 = g_{22}(dx^2)^2 = -(dx^2)^2$$

и сокращения нет.

Таким образом, в гравитационном поле геометрия не является Евклидовой, (но, поскольку $\kappa \ll 1$, проверить это непосредственно довольно трудно)

2.6.2 Течение времени.

Пусть часы, измеряющие единичный интервал времени, расположены вблизи источника гравитационного поля

$$ds = 1; dx^1 = 0; dx^2 = 0; dx^3 = 0.$$

Тогда из условия $1 = g_{44}(dx^4)^2 \Rightarrow$

$$dx^4 = \frac{1}{\sqrt{g_{44}}} = \frac{1}{\sqrt{1 + (g_{44} - 1)}} = 1 - \frac{g_{44} - 1}{2} > 1$$

т.е. часы идут медленнее, следовательно, в спектральных линиях света звезд должно быть гравитационное красное смещение.

Это предсказание подтверждается наблюдениями

2.6.3 Отклонение луча света

Можно показать, что при прохождении луча света рядом с источником гравитационного поля имеет место отклонение луча.

Пусть дано направление распространения луча, т.е. отношения $dx^1 : dx^2 : dx^3$, тогда из $ds^2 = g_{ik}(x)dx^i dx^k$ следует выражение для трехмерной скорости в виде

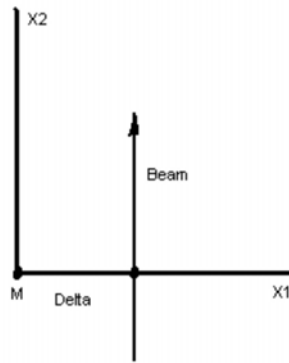
$$\gamma = \sqrt{\left(\frac{dx^1}{dx^4}\right)^2 + \left(\frac{dx^2}{dx^4}\right)^2 + \left(\frac{dx^3}{dx^4}\right)^2}$$

Если n – направление, перпендикулярное направлению распространения луча, то луч света, рассматриваемый в плоскости (γ, n) , обладает кривизной $(-\frac{\partial \gamma}{\partial n})$.

Общее искривление луча

$$B = \int_{-\infty}^{\infty} \frac{\partial \gamma}{\partial x^1} dx^2 = \frac{2\alpha}{\Delta}$$

Для Солнца $B = 1,7$ угловой секунды – это проверено измерениями Эддингтона во время специальной экспедиция в Бразилию 1919 года для наблюдения полного Солнечного затмения



и соответствующих измерений координат звезд. Следует отметить, что Эддингтон, бывший признанным авторитетом в науке того времени, скептически относился к теории Эйнштейна. Однако, выполнив эти измерения признал ее состоятельность и правоту, что является редким случаем.

2.6.4 Прецессия орбиты Меркурия

Расчет орбитального движения этой планеты на основе ОТО привел к отклонению от закона Кеплера, вполне удовлетворительно соответствующему наблюдаемой прецессии орбиты Меркурия.

Таким образом, теория Эйнштейна дала описание известных явлений (гравитация Ньютона), справилась с затруднением старой теории (прецессия орбиты Меркурия) и дала специфические предсказания (отклонение луча и гравитационное красное смещение), которые были обнаружены при наблюдениях.

2.7 Точное решение Шварцшильда (1916)

Решение уравнений ОТО и выражение для метрики, полученные Эйнштейном, носят приближенный характер. Между тем, представляет несомненный интерес и задача аналитического определения вида метрики, входящей в определение линейного элемента

$ds = \sqrt{\sum g_{ik}(x)dx^i dx^k}$, из уравнений поля Эйнштейна $\sum_{\alpha} \frac{\partial \Gamma_{\mu\nu}^{\alpha}}{\partial x^{\alpha}} + \sum_{\alpha, \beta} \Gamma_{\mu\beta}^{\alpha} \Gamma_{\nu\alpha}^{\beta} = 0$ для сферически симметричного случая, соответствующего области около отдельной звезды, и выполнения условия на определитель $g = -1$.

Сформулируем дополнительные условия такой задачи, использованные Шварцшильдом:

1. Все компоненты метрики не должны зависеть от времени x^4 – стационарность.
2. На бесконечности должны обращаться в ноль все компоненты метрики, кроме диагональных, которые соответствуют метрике Минковского (плоское пространство)

Решение задачи:

в сферических координатах метрика, удовлетворяющая указанным условиям, имеет вид

$$ds^2 = (1 - r_S/r)c^2 dt^2 - \frac{dr^2}{1 - r_S/r} - r^2(d\theta^2 + \sin^2 \theta d\varphi^2)$$

где $r_S = \frac{2GM}{c^2}$. Примечательным обстоятельством полученного решения является то, что его сингулярность находится не в нуле, а на некотором расстоянии от него, зависящем от массы гравитирующего объекта и известном как “радиус Шварцшильда”. У этого радиуса имеется прозрачный физический смысл: $r_S = \frac{2GM}{c^2}$ – радиус массивного тела (звезды, планеты), для которого 2-я космическая скорость равна скорости света $v_{II} = \sqrt{\frac{2GM}{r_S}} = c$. То есть материальному телу улететь от планеты с такими параметрами (радиусом и массой) невозможно.

2.8 ОТО и космология

Приложения ОТО к космологии с самого начала оказались эффектными и интригующими, хотя В.А.Фок еще в 30-е годы предупреждал, что на этом пути могут встретиться затруднения. В первой четверти XX века Вселенная еще считалась стационарной, и Эйнштейн пытался устранить возникающие при этом парадоксы. Для построения космологии, т.е. модели Вселенной в целом, можно принять, что скорость движения звезд мала по сравнению со скоростью света. Тогда при формулировке теории надо выбрать между такими условиями:

- а) считать, что на бесконечности должны обращаться в ноль все компоненты метрики, кроме диагональных;
- б) не устанавливать для пространственной бесконечности никаких всегда справедливых граничных условий.

Основной недостаток условия а) состоит в том, что имеющаяся аналогия с газом соответствует тому, что условие нулевой плотности на бесконечности влечет нулевую плотность в центре.

Пусть пространство бесконечно и все же заполнено веществом с некоторой постоянной плотностью ρ_0 . Тогда вместо обычного уравнения Пуассона запишем уравнение

$$\Delta\varphi - \lambda\varphi = 4\pi K\rho,$$

решение которого имеет вид

$$\varphi = -\frac{4\pi K}{\lambda}\rho_0,$$

где λ – некоторая универсальная константа. А гравитационное поле Ньютона создается неравномерностью распределения масс.

Тогда λ войдет и в уравнения поля Эйнштейна, и они примут вид

$$G_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T)$$

В этом случае λ называется космологической константой.

Это уравнение выполнено, если

$$\rho = \rho_0 + \rho_1; \lambda = \frac{\kappa\rho_0}{2}; \lambda = \frac{1}{R^2},$$

мир – цилиндрический (т.е. 3-мерное пространство имеет постоянную кривизну) и существует некоторая “мировая масса”, связанная с ρ_0 . Линейный элемент в такой теории можно выбрать в следующем виде

$$ds^2 = c^2 dt^2 - dr^2 - R^2 \sin^2 \frac{r}{R} [d\theta^2 + \sin^2 \theta d\varphi^2]$$

Де Ситтер (1917) показал, что если принять

$$\rho_0 = 0; \lambda = \frac{3}{R^2},$$

то уравнения также удовлетворены, мир – сферический (с постоянной кривизной 4-мерного пространства), и линейный элемент при этом имеет вид

$$ds^2 = c^2 \cos^2 \frac{r}{R} dt^2 - dr^2 - R^2 \sin^2 \frac{r}{R} [d\theta^2 + \sin^2 \theta d\varphi^2]$$

В 1922 году А.Фридман отказался от ограничения, связанного со стационарностью Вселенной и допустил возможность изменения постоянной пространственной кривизны со временем. При этом миры Эйнштейна и де Ситтера стали частными случаями полученного им общего решения. Однако идея нестационарности Вселенной была чуждой тогдашнему мировоззрению, и Эйнштейн не сразу и весьма неохотно признал, что решение Фридмана было верным, но счел его абстрактной математикой.

Но в 1929 году Э.Хаббл пронаблюдал красное смещение в спектрах удаленных галактик и интерпретировал его как результат эффекта Доплера – все галактики удалялись от Солнечной системы, причем, чем дальше они находились, тем больше было красное смещение, что соответствовало большей скорости удаления. Этот наблюдательный результат перевернул существовавшее мировоззрение, теория Фридмана получила мощное экспериментальное обоснование, и Вселенная была признана расширяющейся. Эйнштейн в этой связи впоследствии упоминал о введении в уравнения космологической постоянной как о самой большой своей ошибке. Поскольку оказалось, что красное смещение наблюдалось в галактиках, расположенных на всех сторонах небосвода, был сделан вывод о существовании в прошлом события Большого Взрыва, которое дало начало нашей Вселенной, до этого существовавшей в состоянии сингулярности. По величине красного смещения, (позволяющей оценить скорость), и оценкам расстояний до галактик был оценен и возраст Вселенной – время, прошедшее после Большого Взрыва. Представление об этом загадочном событии (и еще более загадочном предшествующем состоянии) позволило разработать целый ряд остроумных теорий о стадиях и характере развития Вселенной, некоторые из которых допускали экспериментальную проверку. В частности, Гамов предсказал существование реликтового излучения, и оно было обнаружено.

Впоследствии были получены разнообразные выражения для нестационарных метрик пространства-времени. Так Леметр (1938) для устранения особенности в точке $r = r_S$ решения Шварцшильда выбрал

$$c\tau = ct + \int \frac{f(r)dr}{1 - \frac{r_S}{r}}; R = ct + \int \frac{dr}{(1 - \frac{r_S}{r})f(r)}; f(r) = \sqrt{\frac{r_S}{r}}$$

Тогда нестационарная метрика

$$ds^2 = c^2 d\tau^2 - \frac{dR^2}{[\frac{3}{2r_S}(R - c\tau)]^{2/3}} - [\frac{3}{2}(R - c\tau)]^2 r_S^{2/3} (d\theta^2 + \sin^2 \theta d\varphi^2)$$

не имеет особенности на r_S . Она стала первой из нестационарных метрик, имеющих указанное свойство и известных сейчас как метрики Фридмана-Робертсона-Уолкера. Они являются общепризнанными в современной космологии.

2.9 Наблюдательные проблемы и “темные” понятия

Наряду с успехами теории, предсказания которой подтверждались одними наблюдениями, стали выявляться и ее проблемы, связанные с другими наблюдениями. Так Ф.Цвике еще в 1933 году обнаружил, что наблюдаемая (видимая) масса кластеров галактик слишком мала, чтобы обеспечить гравитационную связанность этих кластеров. Уже тогда он предположил, что какая-то часть материи, входящей в галактики, не видна. Впоследствии эта точка зрения находила все больше подтверждений и сторонников, и в настоящее время убеждение в существовании “темной материи” является распространенным, хотя и не всеобщим.

В десятилетие, следующее за 1998 годом, на основании наблюдении сверхновых типа Ia и соотнесении их результатов с общепринятой теорией расширяющейся Вселенной пришли к выводу, что на современном этапе своего развития Вселенная расширяется ускоренно. Это означает, что на космологическом масштабе имеется неизвестное в настоящее время взаимодействие отталкивания (антигравитация?). При этом оценки величины энергии этого взаимодействия – “темной энергии” – на основе ОТО показывают, что она превосходит наблюдаемую имеющуюся массу вместе с темной материей в 3-4 раза.

При обсуждении ускорения расширения Вселенной и соответствующей ему темной энергии можно ввести много дополнительных подробностей, оставляющих большой простор для теоретических разработок. Так, первое, что вспоминается при описании данной проблемы, это космологическая постоянная, обсуждавшаяся в свое время Эйнштейном. Теперь она должна не удерживать материю Вселенной, а приводить к ее отталкиванию. Такой константе придают своеобразный геометро-физический смысл, а именно – смысл энергии, связанной с самим понятием пространства, что, хотя и не находится в прямом противоречии с идеями геометродинамики, пока еще недостаточно привычно. Далее, вспоминая, что компоненты тензора напряжений, связаны с силой, действующей на площадку, и принимая закон Паскаля $\sigma_{\alpha\beta} df_{\beta} = p df_{\alpha}$, т.е. $\sigma_{\alpha\beta} = p \delta_{\alpha\beta}$ для космологических приложений (для галактик), получим для системы покоя элемента объема

$$T^{ik} = \begin{pmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix},$$

т.е. для произвольной системы

$$T^{ik} = (p + \varepsilon)u^i u^k - p g^{ik}; T_i^k = (p + \varepsilon)u_i u^k - p \delta_i^k.$$

Тогда плотность энергии будет складываться из плотности собственно энергии $\varepsilon = \rho c^2$ и давления

$$T_i^i \equiv \varepsilon_G = \varepsilon + 3p$$

Положительное давление соответствует сжатию, отрицательное давление – расширению. Тогда, вводя параметр

$$w = \frac{p}{\varepsilon}; p = w\varepsilon,$$

можно рассмотреть следующие возможные ситуации и связанные с ними космологические режимы:

1. $w < -1$. “Фантомная энергия”. Мощная “антигравитация”, средняя плотность энергии увеличивается со временем, явление “разрыва”.
2. $w = -1$. Вакуум, обладающий постоянной плотностью энергии (космологическая константа $\lambda = \frac{8\pi G}{c^4}\varepsilon$). Согласно современным измерениям, доля энергии, приходящаяся на вакуум, составляет $\Omega_\lambda = 0.72$ от критического значения для плотности $\rho_c = \frac{3H^2}{8\pi G}$, соответствующей переходу от отрицательной к положительной кривизне (т.е. плоской Евклидовой геометрии).
3. $-1 < w < -1/3$. Для такого интервала значений этого параметра применимо понятие “квинтэссенции” гипотетического динамического скалярного поля. “Антигравитация” при этом оказывается слабее, чем у вакуума.
4. $w = -1/3$. Нет ни гравитации, ни антигравитации.
5. $-1/3 < w < 0$. Слабая гравитация (меньше, чем у пыли).
6. $w = 0$ ($p = 0$). Барионная пыль и “темная материя” $\Omega_b = 0.046$; $\Omega_c = 0.033 \Rightarrow \Omega_m = 0.28$
7. $0 < w < 1/3$. Гравитация (больше, чем у пыли).
8. $w = 1/3$. Гравитация, ультрарелятивистская среда (излучение, фотоны).
9. $w > 1/3$. Гравитация сильнее, чем у излучения.

Имеется обширная литература, посвященная обсуждению соответствующих весьма интересных проблем и их расчету.

Обсудим подробнее некоторые наблюдения, говорящие в пользу существования темной материи. Известен эффект гравитационного линзирования, состоящий в том, что массивные кластеры галактик, расположенные на луче зрения, направленном на более удаленные объекты, отклоняют световые лучи подобно обычной линзе, и возникают симметричные изображения двойников этих удаленных объектов (рис.1). Качественно этот эффект объясняется

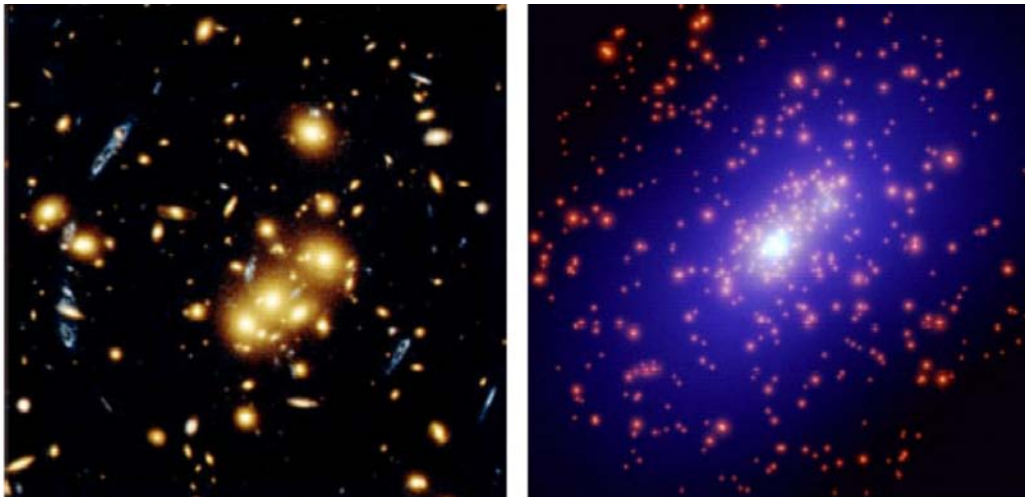


Рис. 1

ОТО, но количественно те оценки масс, которые используются в астрономии, оказываются в несколько раз меньше, чем значения, необходимые для обеспечения наблюдаемого эффекта.

На рис.2 представлены результаты наблюдений космической обсерватории “Чандра” (2006) столкновения кластеров галактик. Слева на оптическом изображении проведены линии, со-

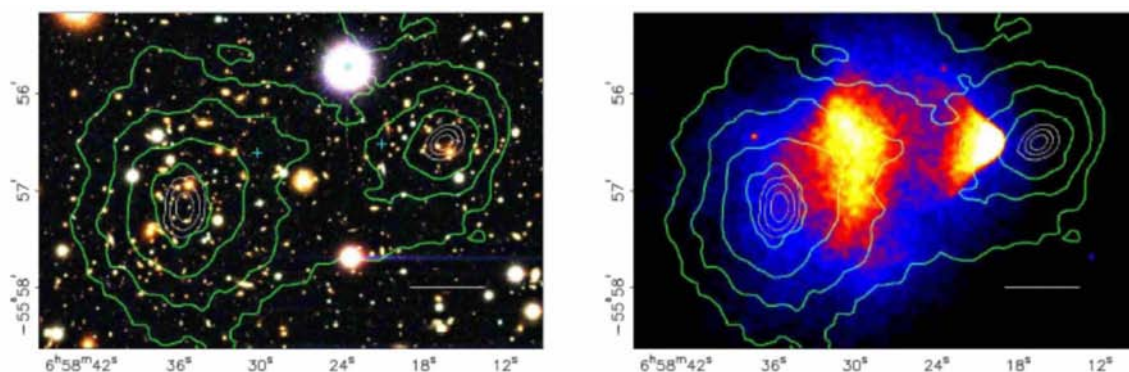


Рис. 2

ответствующие равным гравитационным потенциалам. Видно, что они окружают центры двух галактических кластеров, расположенных поблизости друг от друга. Обратим внимание на масштаб явления – это центры кластеров галактик. На правом рисунке представлено изображение того же участка в рентгеновском диапазоне, что позволяет наблюдать за распределением барионной плазмы в межгалактическом пространстве. Видно, что гравитационные центры плазмы (как полагают, области ударных волн, возникших при прохождении кластеров друг сквозь друга) не совпадают с центрами масс, определенными по результатам линзирования и совпадающими с оптическими центрами кластеров. Расчеты масс, необходимых для наблюдаемого линзирования, дают, что 1-2% массы приходится собственно на видимые звезды галактик, 5-10% массы приходится на межгалактическую барионную плазму, испускающую рентгеновское излучение, но кроме того необходима еще и темная материя, на которую приходятся оставшиеся 88% массы. В стационарном случае центры масс всех трех составляющих совпадают, но на фотографии мы видим, что это не так. Из этого делается вывод о том, что необходимая темная материя, по-видимому, связана с галактиками и является бесстолкновительной.

Приписывание эффектов со столь значительными относительными масштабами ненаблюдаемым “темным” понятиям указывает на глубину научного кризиса, имеющегося в настоящее время. Прежде, чем перейти к обсуждению попыток его преодоления, укажем еще один важный ряд наблюдательных данных.

Наиболее известными, подробными и количественными являются данные, связанные с измерением кривых вращения спиральных галактик – зависимостей орбитальных скоростей (чаще говорят о квадратах скоростей) звезд от их расстояний до центра галактики. На рис.3 представлены соответствующие экспериментальные графики. Известно, что все они описываются эмпирической зависимостью следующего вида

$$v^2 = \frac{\beta^* c^2 N^*}{R} + \frac{\gamma^* c^2 N^* R}{2} + \frac{\gamma_0 c^2 R}{2} \quad (19)$$

где R – расстояние до центра галактики, c – скорость света, N^* – число звезд в галактике (что придает соответствующему члену в выражении для квадрата скорости пропорциональность массе), β^* , γ^* , γ_0 – параметры. В формуле (19) наряду с первым слагаемым, соответствующим теории Ньютона-Эйнштейна-Шварцшильда, имеются и два других, прямо пропорциональных расстоянию до центра галактики, что категорически не соответствует имеющимся теоретическим представлениям.

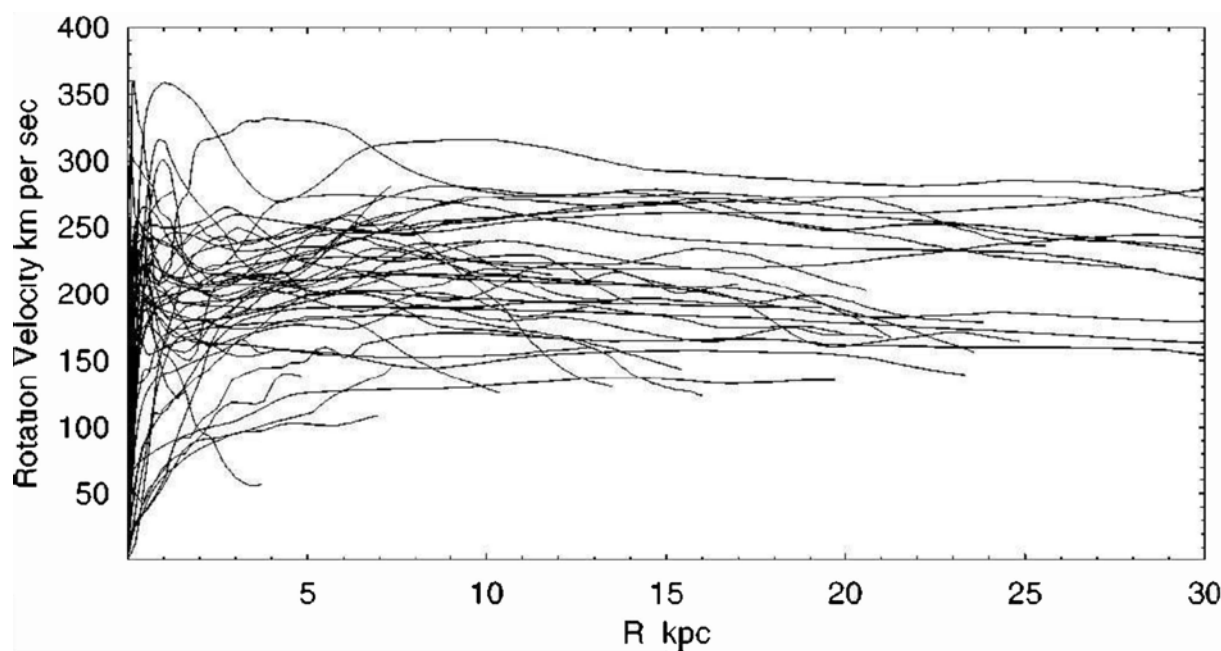


Рис. 3

Для того, чтобы обеспечить получаемый вид кривых вращения, предполагают, что существует сферически симметричное облако темной материи, окружающее галактику (гало), обладающее следующими свойствами:

1. частицы облака не взаимодействуют с электромагнитным излучением, не поглощают его и не испускают. Предполагают, что ими могут являться бозоны Хиггса – носители свойства массивности. Имеется надежда на обнаружение этих частиц в экспериментах на Большом адронном коллайдере в ЦЕРНе;
2. масса облака должна быть в 3-4 раза больше, чем наблюдаемая барионная масса галактики (звезды + межзвездный газ);
3. размер облака должен превышать линейный размер галактики примерно в 4-5 раз;
4. облако имеет сферическую симметрию. Это приводит к проблеме различия наблюдаемого гравитационного потенциала в плоскости галактики (для обеспечения наблюдаемых кривых вращения) и перпендикулярно ей (3D-проблема). Так, наблюдаемые глобулярные кластеры – сфероидальные скопления звезд, принадлежащих галактике, но лежащих вне ее плоскости, двигаются в соответствии с обычной теорией. С ними, однако, связана другая проблема: из статистических соображений они должны большую часть времени находиться на периферии своих орбит, однако там их оказывается значительно меньше, чем в областях более близких к центру галактики;
5. параметры необходимого сферического распределения темной материи подбираются по наблюдательным данным о кривых вращения, поэтому рассчитать сами эти кривые по данным только о светимости галактики невозможно. Это лишает обсуждаемую гипотезу внутренней самосогласованности.

Облако темной материи с перечисленными свойствами должно существовать для того, чтобы классическая ОТО была справедлива на галактическом масштабе и не требовала модифи-

кации. Похожая ситуация уже встречалась в истории науки, когда в движении Урана были обнаружены незначительные отклонения от предсказаний теории, основанной на законе гравитации Ньютона. Тогда высказывалось мнение, что так далеко от Солнца теория Ньютона не работает. Но Леверье и Адамс, не желая так просто отказываться от взглядов Ньютона, предположили существование неизвестной планеты, и вскоре Галле, следуя указаниям Леверье, действительно обнаружил Нептун. Видно, однако, что относительные масштабы поправок, требуемых тогда и теперь, несоизмеримы (см. рис.4). При этом, если бозоны Хиггса

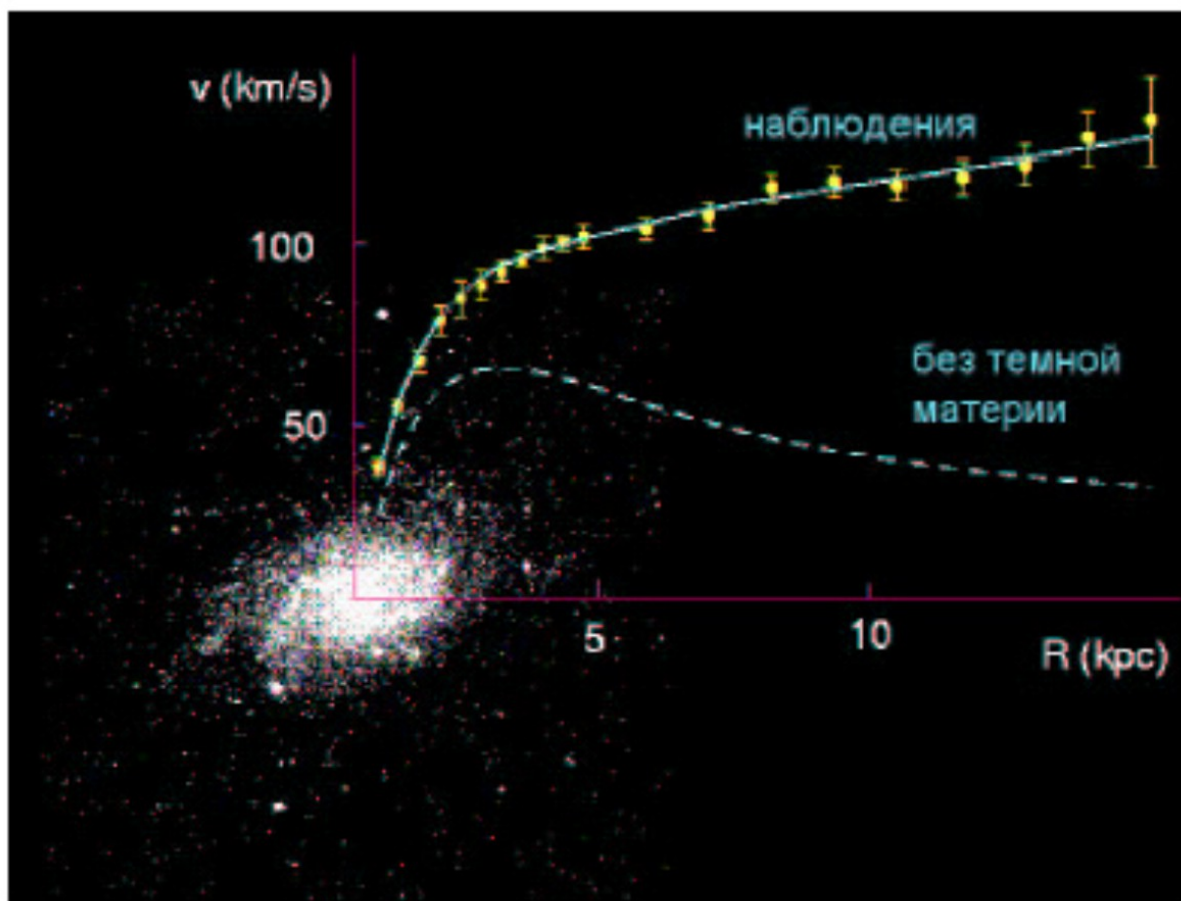


Рис. 4

будут обнаружены, сам этот факт не будет достаточным обоснованием для утверждения гипотезы темной материи как космологической причины наблюдаемых эффектов.

Упомянем, наконец, менее вызывающий, но оттого не менее интересный установленный экспериментальный факт, состоящий в том, что светимость галактики пропорциональна четвертой степени орбитальной скорости звезд, находящихся на ее периферии

$$L_{lum} \sim v_{orb}^4 \quad (20)$$

Это соотношение известно как закон Талли-Фишера, и в рамках ОТО оно не имеет удовлетворительного объяснения в настоящее время.

2.10 Попытки модификации ОТО – классический (геометрический) подход

Перечислим некоторые подходы, имеющие целью модифицировать основы ОТО с тем, чтобы удовлетворить данным наблюдений на космологических масштабах, не ставя задачу дать исчерпывающий обзор достигнутых результатов.

I. Наиболее прямолинейным подходом является последовательное усложнение известного квадратичного выражения Эйнштейна-Гильберта для действия

$$S_{EH} = -\frac{c^3}{16\pi G} \int d^4x (-g)^{1/2} R_\alpha^\alpha \quad (21)$$

за счет введения чисто метрических членов следующих порядков. Например, Де Витт ввел следующие (квадратичные) поправки

$$S_{W_1} = -\frac{c^3}{16\pi G} \int d^4x (-g)^{1/2} (R_\alpha^\alpha)^2 \quad (22)$$

или

$$S_{W_2} = -\frac{c^3}{16\pi G} \int d^4x (-g)^{1/2} R_{\alpha\beta}^{\alpha\beta} \quad (23)$$

которые должны давать пренебрежимо малый вклад в решение Шварцшильда. Развитие этого подхода привело к так называемым $f(R)$ -теориям, где скаляр представлял собой разложение по соответствующим степеням.

II. Следующим естественным ходом может быть введение дополнительного аддитивного скаляра, соответствующего макроскопическому гравитационному полю. Например, Бранс и Дике предложили следующее выражение для действия

$$S_{BD} = -\frac{c^3}{16\pi G} \int d^4x (-g)^{1/2} (SR_\alpha^\alpha - w \frac{S_{;\mu} S^{;\mu}}{S}) \quad (24)$$

где w – константа.

III. Третьим направлением является увеличение числа измерений пространства-времени с последующим выходом измерений высших порядков на Планковский масштаб длин. Соответствующие работы, начавшись с работы Калуцы, привели к развитым в математическом смысле современным теориям струн, а затем и бран. Следует, однако, отметить, что при всей эстетической привлекательности этих теорий, заметных физических результатов на их основе к настоящему времени добиться не удалось.

Перечислим теперь подходы, подразумевающие не столько уточнение деталей существующих структур, предположительно позволяющие получить решение, более соответствующее наблюдениям, сколько подходы, предусматривающие пересмотр самих структур, предположительно позволяющие сделать то же.

IV. Ревизии подверглись и классические основы. Анализ кривых вращения навел Мильгрема на мысль, что поправки следует искать глубже. Им был предложен феноменологический подход, известный как МОНД – Модифицированная Ньютоновская Динамика, в котором изменению подвергается либо классическое уравнение динамики Ньютона

$$\mu\left(\frac{a}{a_0}\right) \vec{a} = \vec{f}, \quad (25)$$

либо классическое уравнение гравитации Ньютона

$$\vec{a} = \nu\left(\frac{f}{a_0}\right) \vec{f} \quad (26)$$

и вводится новая мировая константа, имеющая размерность ускорения. Предлагается так подобрать функции $\mu(x)$ или $\nu(x)$ и величину a_0 , чтобы в масштабах Солнечной системы получать классический результат, а в масштабе галактики – результат, соответствующий эмпирической формуле для кривых вращения. Введенные Мильгремом наиболее известные выражения для функций МОНД имеют вид:

$$\mu(x) = \frac{1}{1+x}; \mu(x) = \frac{x}{\sqrt{1+x^2}} \quad (27)$$

Этот подход, естественно, позволяет вполне удовлетворительно описать наблюдаемые кривые вращения галактик, однако произвол в выборе соответствующих функций не позволяет приписать физический (или геометрический, как в ОТО) смысл предлагаемой модификации.

V. Г. Вейль еще в 1918 году отступил от использованной Эйнштейном геометрии Римана в целях объединения гравитации и электромагнетизма с помощью введения новой метрики. Он предложил использовать преобразования следующего вида

$$g_{\mu\nu}(x) \rightarrow e^{2\alpha(x)} g_{\mu\nu}(x) \quad (28)$$

$$A_\mu(x) \rightarrow A_\mu(x) - e\partial_\mu\alpha(x) \quad (29)$$

в которых гравитацию и электромагнетизм объединяет общая функция $\alpha(x)$. Это привело к созданию новой – Вейлевой – геометрии. Получающиеся далее уравнения не сводятся к уравнениям Эйнштейна, но содержат решения Швардшильдковского типа для масштабов Солнечной системы. Преобразование (28) Вейль назвал калибровочным, т.е. зависящим от масштаба, однако в дальнейшем этот термин стал использоваться в других областях физики и преимущественно для случая, когда показатель экспоненты является мнимым. В применении к гравитации такие преобразования обычно называют “конформными”.

VI. Дальнейшее развитие подобных идей привело к теориям конформной гравитации, в которых метрика лишь наделяется дополнительной симметрией, соответствующей (28), электромагнитные переменные не вовлекаются, а, значит, геометрия остается Римановой. Формально этот подход оказывается аналогичным подходу I, но уже со специфическим выбором коэффициентов при (22) и (23). Уравнения Эйнштейна, получающиеся у Манхейма и Казанаса, имеют вид

$$4\alpha_g W^{\mu\nu} = 4\alpha_g (2C^{\mu\lambda\nu\kappa}_{;\lambda;\kappa} - C^{\mu\lambda\nu\kappa} R_{\lambda\kappa}) = T^{\mu\nu} \quad (30)$$

где α_g – безразмерная константа, $C^{\mu\lambda\nu\kappa}$ – так называемый тензор Вейля, не изменяющийся при преобразованиях (28). После замены

$$\begin{aligned} W^{\mu\nu}(x) &\rightarrow e^{-6\alpha(x)} W^{\mu\nu}(x) \\ T^{\mu\nu}(x) &\rightarrow e^{-6\alpha(x)} T^{\mu\nu}(x), \end{aligned}$$

преобразования координат с использованием некоторой функции $B(r)$ и введения функции источника в виде $f(r)$ уравнение (30) в стационарном случае приводится к уравнению Пуассона, но не второго порядка, как в обычном случае, а четвертого

$$\nabla^4 B(r) = f(r) \quad (31)$$

При наличии сферической симметрии уравнение (31) имеет точное решение. При этом оно не только содержит слагаемое, соответствующее решению Ньютона-Шварцшильда, но и слагаемые того же вида, что и эмпирическое уравнение (19) для кривых вращения

$$\begin{aligned} B(r > R) = -g_{00} &= 1 - \frac{2\beta}{r} + \gamma r \\ 2\beta &= \frac{1}{6} \int_0^R dr' r'^4 f(r'); \gamma = \frac{1}{2} \int_0^R dr' r'^2 f(r') \end{aligned} \quad (32)$$

Описанный подход не требует введения дополнительной (темной) материи, т.е. дополнительного скалярного поля. Вместо этого он использует другой выбор скалярной функции при формулировке вариационного принципа. Это сохраняет геометрию пространства-времени, но приводит к уравнению Эйнштейна вида (30), которое, в частности, в пустом пространстве не имеет структуры волнового уравнения. Таким образом, нет оснований говорить о каких-либо гравитационных волнах. Но тогда возникают проблемы с интерпретацией эффекта Хюлса-Тэйлора, наблюдавших изменение периода обращения двойного пульсара с течением времени, которое оказалось согласующимся с предсказанием ОТО, основанном на испускании гравитационных волн, с точностью до долей процента.

2.11 Анизотропная геометродинамика – наводящие соображения и первые шаги

Рассмотрим другую возможность геометрического подхода к решению возникших проблем. Те преобразования простейшего скаляра или выбор нового, изменение физического закона или выбор другой геометрии, которые были упомянуты выше, носили произвольный и не всегда обоснованный характер. Между тем, можно предложить путь обобщения, который выглядит довольно естественным. Речь идет об отказе от постулируемой изотропии пространства. Наводящим соображением служит плоский характер многих галактик, а естественность состоит в том, что, полагая метрику анизотропной, мы будем неизбежно получать и обобщенный “простейший скаляр”, поскольку метрический тензор участвует в его построении. При этом следует не пытаться угадать нужный вид метрики, а задать его в общем виде и попытаться уточнить, когда дело дойдет до конкретных результатов, узнав тем самым характеристики анизотропии физического пространства.

Анизотропия означает, что метрика в каждой точке будет зависеть от некоторого направления. Является ли возможная анизотропия глобальной, локальной или смешанной, сразу сказать невозможно. Поэтому прежде, чем модифицировать теорию, представляется естественным понять, какого сорта проблемы мы намерены решить. Если начинать с тех, которые были описаны выше, и не ставить радикальный отход от классической геометродинамики в качестве самостоятельной задачи, то следует осознать, как теперь следует воспринимать основные постулаты ОТО и, в частности, принцип эквивалентности. Он утверждает, что мы не можем различить эффекты, связанные с ускорением системы отсчета, и эффекты, связанные с действием гравитационного поля. Но если ускорение системы отсчета не является прямолинейным, то инерциальные силы, действующие на тело, зависят от скорости его движения (например, сила Кориолиса). Тогда в силу принятого принципа эквивалентности то, что принято называть гравитационным воздействием, также должно зависеть от скорости движения тел. Иными словами, измеряя действие гравитационных сил, мы никогда не знаем, в какой мере в наблюдениях учитывается кинематика. С математической точки зрения это будет означать, что пространство является локально анизотропным.

Наглядно можно представить себе это как переход от прямолинейно ускоряющегося “лифта Эйнштейна” к некоторой комнате, находящейся на вращающейся “карусели” так, что радиус, проведенный от центра вращения к ободу, перпендикулярен полу и потолку. Падающий шарик будет отклоняться от этого радиуса, другие движения будут также зависеть от начальной скорости тела. И, если о движении карусели ничего не известно, то наблюдаемая “гравитация” будет зависеть от скоростей как самого тела, контролируемой наблюдателем, так и от скорости вращения, о которой он может и не подозревать.

В отличие от классической ОТО и в свете возникших проблем, может оказаться более естественным начать построение теории для тех масштабов, которые характеризуют большую часть Вселенной, и лишь затем убедиться, что и в частном случае, например, для планетной системы, теория остается работоспособной. Тогда то, что сейчас является необъяснимой проблемой, станет своего рода исходной посылкой, которая в пределе не должна противоречить представлениям Ньютона-Эйнштейна-Шварцшильда о гравитации для точечной массы.

Учет зависимости метрики не только от координаты, но и от направления (от “скорости”) в каждой точке, между прочим, приводит и к тому, что касательное пространство к основному многообразию становится восьмимерным *фазовым* пространством, имеющим в физике наглядный смысл.

Опишем вкратце первые результаты, возникающие на этом пути. Рассмотрим анизотропное пространство с деформированной метрикой простейшего вида

$$g_{ik}(x, y) = \gamma_{ik} + \varepsilon_{ik}(x, y)$$

$\gamma = \text{diag}\{1, -1, -1, -1\}$ где метрика Минковского, не зависящая от y , а $\varepsilon_{ik}(x, y)$ – малое анизотропное возмущение. Такая метрика является обобщенной Лагранжевой метрикой, и -переменные должны играть точно такую же роль, как и -переменные, поэтому имеет место соотношение $y^i = \frac{1}{H} \frac{dx^i}{dt} = \frac{1}{H} v^i$, где H – некоторая константа с размерностью (секунда)⁻¹. Будем считать $\varepsilon(,)$ достаточно малым для того, чтобы использовать линейное приближение и сделаем следующие упрощающие предположения, полностью соответствующие исходным предположениям Эйнштейна, рассмотренным в разделе 5:

1. Скорости рассматриваемых объектов значительно меньше, чем фундаментальная скорость. Это означает, что компонентами ², ³ и ⁴ можно пренебречь по сравнению с компонентой ¹, которая равна единице с точностью до второго порядка;
2. Поскольку скорости малы, производной метрики по времени $\frac{\partial \varepsilon_{hj}}{\partial x^1}$ можно пренебречь по сравнению с производными по пространственным координатам $\frac{\partial \varepsilon_{hj}}{\partial x^\alpha}$; $\alpha = 2, 3, 4$;
3. Последнее будем считать верным и для соответствующих ускорений, т.е. на u -подпространстве фазового пространства производной по ¹ можно пренебречь по сравнению с производными по ², ³ и ⁴.

2.11.1 Геодезическая

Поскольку метрика теперь зависит от y , уравнения Эйлера-Лагранжа получаются при варьировании следующего (Финслеровского) выражения для квадрата расстояния

$$F^2 = (\gamma_{ik} + \varepsilon_{ik}(x, y)) y^i y^k$$

В этом случае уравнение получающейся обобщенной геодезической принимает вид:

$$\frac{dy^i}{ds} + (\Gamma_{lk}^i + \frac{1}{2}\gamma^{it}\frac{\partial^2\varepsilon_{kl}}{\partial x^j\partial y^t}y^j)y^ky^l = 0$$

Здесь $\Gamma_{jk}^i = \frac{1}{2}\gamma^{ih}\{\frac{\partial\varepsilon_{hj}}{\partial x^k} + \frac{\partial\varepsilon_{hk}}{\partial x^j} - \frac{\partial\varepsilon_{jk}}{\partial x^h}\}$ представляет собой символ Кристоффеля, но зависящий и от направления, т.е. от y .

Полученная геодезическая представляет собой геометрические уравнения, которые следует использовать для описания динамики частицы в гравитационном поле, соответствующем анизотропной метрике. Как и в работе Эйнштейна, упомянутой в разделе 5, используем сделанные предположения для того, чтобы сохранить только члены с $k = l = 1$. Тогда из всех ε_{ik} , имеющих в уравнении, останется только ε_{11} , как и в классической работе, при этом $k = 1 = 1$.

Введем новые обозначения

$$\frac{1}{2}\frac{\partial\varepsilon_{11}}{\partial y^t} \equiv A_t$$

соответствующие компонентам тензора Картана. Тогда получим

$$\frac{dy^i}{ds} + \Gamma_{11}^i + \gamma^{it}\frac{\partial A_t}{\partial x^j}y^j = 0$$

Добавим и вычтем одну и ту же величину $\gamma^{it}\frac{\partial A_j}{\partial x^t}y^j$ и получим

$$\frac{dy^i}{ds} + \Gamma_{11}^i + \gamma^{it}[(\frac{\partial A_t}{\partial x^j} - \frac{\partial A_j}{\partial x^t}) + \frac{\partial A_j}{\partial x^t}]y^j = 0$$

где $(\frac{\partial A_t}{\partial x^j} - \frac{\partial A_j}{\partial x^t})$ могут рассматриваться как компоненты антисимметричного тензора F_{jt} точно такого же, как и в электродинамике. Выражения $(\frac{\partial A_t}{\partial x^j} - \frac{\partial A_j}{\partial x^t})$ представляют собой компоненты ротора вектора A . Тогда для геодезической получим

$$\frac{dy^i}{ds} + \Gamma_{11}^i - \gamma^{it}F_{tj}y^j + \gamma^{it}\frac{\partial A_j}{\partial x^t}y^j = 0$$

2.11.2 Снова “тождества Максвелла”

Формально введенный тензор F_{ij} обладает полезными чисто геометрическими свойствами. А именно, он удовлетворяет тому же самому тождеству,

$$\frac{\partial F_{ij}}{\partial x^k} + \frac{\partial F_{jk}}{\partial x^i} + \frac{\partial F_{ki}}{\partial x^j} = 0$$

которое было рассмотрено в разделе 2.

Повторяя ту же схему рассуждений, введем формальное обозначение

$$F_{12} = E_y^{(g)}; F_{13} = E_z^{(g)}; F_{14} = E_x^{(g)}; F_{23} = -B_z^{(g)}; F_{24} = B_y^{(g)}; F_{34} = -B_x^{(g)}$$

и немедленно получим ту же пару “тождеств Максвелла”, которые в электродинамике известны как уравнения Максвелла

$$\begin{aligned}\frac{\partial \vec{B}^{(g)}}{\partial t} + \text{rot} \vec{E}^{(g)} &= 0 \\ \text{div} \vec{B}^{(g)} &= 0\end{aligned}$$

Следуя геометрическому рецепту, как и в разделе 2, перейдем к контравариантному тензору $F^{ij} = \gamma^{ik}\gamma^{jm}F_{mk}$ (для поднятия значков можно использовать γ^{it} поскольку A_t уже содержит ε), и введем новый 4-вектор I^i следующим образом

$$I^i = \frac{\partial F^{ij}}{\partial x^j}$$

Тогда аналогично может быть получена и вторая пара “тождеств Максвелла”

$$\begin{aligned}\text{rot} \vec{B} - \frac{\partial \vec{E}}{\partial t} &= \vec{j} \\ \text{div} \vec{E} &= \rho\end{aligned}$$

где мы формально обозначили

$$I^1 \equiv j_x; I^2 \equiv j_y; I^3 \equiv j_z; I^4 \equiv \rho$$

Физическая интерпретация полученных геометрических результатов теперь не может включать никаких полей, кроме гравитационных. Смысл функций в полученных уравнениях не такой, как в уравнениях Максвелла в электромагнетизме. Вектора $^{(g)}$ и $^{(g)} = \text{rot}_{(x)} A$ описывают пространство-время с анизотропной метрикой и, таким образом, характеризуют гравитацию, зависящую от “скорости”. Если интерпретировать как вектор-потенциал гравитационного поля, $\rho^{(m)}$ – как плотность массы источника гравитации, и $j^{(m)} = \rho^{(m)}v$ – как плотность потока массы, или массовый ток, соответствующий собственному движению источника и его частей, мы получим впечатляющую аналогию с электромагнетизмом, и весь соответствующий ему формализм может быть использован при расчетах.

Ранее в теории встречалось исследование влияния вращения (точечного) источника на гравитационные эффекты. Вероятно, наиболее известным является эффект Лензе-Тирринга, в котором рассчитывается прецессия гироскопа (например, планеты) в поле вращающейся массы (например, звезды). Стоит обратить внимание на то, что исходной целью авторов было исследование принципа Маха о происхождении инерции. Известны и другие гравитоманнитные эффекты, рассчитанные в рамках релятивистского подхода: разница показаний часов, совершивших оборот по и против вращения центрального тела в его экваториальной плоскости (используется метрика Керра); эффект Саньяка; вращающаяся черная дыра и другие, включая расщепление спектральных линий, аналогичное эффекту Штерна-Герлаха. Во времена их первых расчетов все эти эффекты были слишком малыми для экспериментальной проверки, но с тех пор техника эксперимента существенно продвинулась. Существование гравитоманнитных явлений подтверждается измерениями так называемой прецессии Де Ситтера для системы Земля-Луна, совпадающими с расчетом с точностью до 1%, и для гироскопа, установленного на космическом аппарате Gravity Probe B, где достигнутая точность была еще выше. Однако в рамках гравитоманнитного подхода никогда не рассматривалась

метрика анизотропного пространства в общем виде, а масштаб гравитомагнитных явлений, обсуждавшихся в литературе, никогда не превышал масштаб планетной системы, оставаясь малой поправкой к решениям Ньютона-Шварцшильда для сферически симметричного точечного источника.

2.11.3 Гравитационная сила и ее компоненты

Возвращаясь к уравнению геодезической с учетом тождеств Максвелла, получим

$$\frac{dy^i}{ds} = -\Gamma_{11}^i + \gamma^{it} F_{tj} y^j - \gamma^{it} \frac{\partial A_j}{\partial x^t} y^j$$

и увидим, что гравитационное ускорение теперь представляет собой сумму трех членов. Первый из них связан с ε_{11} и приводит к появлению классического выражения $-\frac{1}{2}\nabla_{(x)}\varepsilon_{11}$,

приводящего к гравитации Ньютона $F_N^{(g)}$, появление двух остальных связано с анизотропией метрики. Второй член представляет собой гравитационный аналог силы Лоренца $F_L^{(g)}$, пропорциональный

$$F_L^{(g)} \sim (E^{(g)} + [y, B^{(g)}])$$

где

$$E^{(g)} = -\frac{1}{2}\nabla_{(x)}\frac{\partial\varepsilon_{11}}{\partial y^1}; B^{(g)} \equiv \text{rot}_{(x)}A; A = \frac{1}{2}\left(\frac{\partial\varepsilon_{11}}{\partial y^2}, \frac{\partial\varepsilon_{11}}{\partial y^3}, \frac{\partial\varepsilon_{11}}{\partial y^4}\right)$$

С учетом третьего предположения слагаемым $^{(g)}$ в уравнении можно пренебречь. Третье слагаемое в геодезической можно сначала преобразовать с учетом второго предположения

$$-\gamma^{\alpha t} \frac{\partial A_j}{\partial x^t} y^j = -(-\delta^{\alpha t}) \left[\frac{\partial}{\partial x^t} (A_j y^j) - \frac{\partial y^j}{\partial x^t} \right] = (\nabla_{(x)}(A_j y^j))^{\alpha} = (\nabla_{(x)}(A, y))^{\alpha}$$

где $\frac{\partial y^j}{\partial x^t}$ исчезает, поскольку y и x — независимые переменные. Вспоминая, что правая часть уравнения геодезической вначале была умножена на ^{11} , и $^1 = 1$ единица длины,

учтем теперь влияние размерных множителей (y и x) и сделанного для удобства выбора констант ($c = H = 1$) непосредственно

$$H \frac{dy}{c^2 dt} = \frac{1}{2} \left\{ -\nabla_{(x)}\varepsilon_{11} + [y, \text{rot}_{(x)}\frac{\partial\varepsilon_{11}}{\partial y}] + \nabla_{(x)}\left(\frac{\partial\varepsilon_{11}}{\partial y}, y\right) \right\} \cdot \left(\frac{H}{c}y^1\right)^2$$

Поскольку $y = \frac{1}{H}v$; $v^1 = c$, выражение для гравитационной силы, действующей на частицу массы m , имеет вид

$$m \frac{dv}{dt} = F^{(g)} = \frac{mc^2}{2} \left\{ -\nabla_{(x)}\varepsilon_{11} + [v, \text{rot}_{(x)}\frac{\partial\varepsilon_{11}}{\partial v}] + \nabla_{(x)}\left(\frac{\partial\varepsilon_{11}}{\partial v}, v\right) \right\}$$

Первый член в скобках соответствует выражению для обычной гравитационной силы $F_N^{(g)}$, действующей на частицу с массой m со стороны источника с массой M и направленной по радиусу к центру распределения масс источника. Решение уравнения Пуассона для стационарного гравитационного поля дает $\varepsilon_{11} \sim \frac{1}{r}$ где r — расстояние от частицы до центра, и

выбор $\varepsilon_{11} = \frac{2GM}{c^2} \frac{1}{r}$ для точечного источника на достаточных расстояниях приводит к закону Ньютона

$$F_N^{(g)} = G \frac{Mm}{r^2}.$$

Величина $r_S = \frac{2GM}{c^2}$ соответствует радиусу Шварцшильда, упоминавшемуся в разделе 7.

Третий член в скобках является новым. Полное ускорение частицы теперь зависит не только от распределения масс, но и от их собственного движения и от скорости частицы. Компонента гравитационной силы $F_3^{(g)}$ соответствует воздействию, оказываемому на движущуюся частицу расширением или сжатием (распределенного) источника. Расширение гравитирующего источника приводит к дополнительной силе притяжения частицы, двигающейся по радиусу к центру.

Выражение для второй компоненты гравитационной силы $F_L^{(g)}$ требует рассмотреть взаимодействие между движущейся частицей и движущимся распределением масс. И если система отсчета связана с последним, и его ускорение не прямолинейно, а, например, соответствует движению по окружности, принцип эквивалентности неизбежно требует появления компоненты гравитации типа силы Кориолиса. В простейшем случае источник может представлять собой массу, близкую к точечной, вращающуюся с угловой скоростью

$$\Omega = \frac{c^2}{4H} \text{rot}_{(x)} \frac{\partial \varepsilon_{11}}{\partial y}$$

Тогда второе слагаемое принимает точный вид “грави-Кориолисовой силы”

$$F_L^{(g)} = 2m[v, \Omega] \equiv F_C^{(g)},$$

создаваемой вращающейся массой, и непосредственно демонстрирует отмеченный выше общий характер связи между инерцией и гравитацией. На плоскости, проходящей через перпендикулярно Ω , можно получить траектории частиц, двигающихся в различных направлениях с различными скоростями. Таким образом, можно проследить различные действия силы $F_L^{(g)}$: притяжение, отталкивание и тангенциальное воздействие, зависящие от угла между v и $V = [R_{eff}\Omega]$, где R_{eff} – радиус-вектор в направлении частицы, длина которого есть эффективный радиус вращающегося источника.

Из этих определений следует, что компоненты скорости, перпендикулярные плоскости, не испытывают действия силы $F_L^{(g)}$, что снимает 3D-проблему различия наблюдаемого гравитационного потенциала в разных направлениях упомянутую в разделе 9. Очевидно, что аналогия с электромагнетизмом может быть весьма полезна при расчетах.

Важно отметить, что в соответствии с последней формулой в анизотропной геометродинамике гравитационное взаимодействие уже не является только притяжением, как раньше, а вклад той или иной компоненты в наблюдаемые явления зависит от параметров системы.

2.11.4 Кривые вращения и закон Талли-Фишера

Рассмотрим сначала упрощенный пример. Пусть источник гравитации представляет собой двойную звезду с характерным расстоянием r_0 , массой и периодом. Пусть планета m движется вокруг звезд по орбите с характерным радиусом r , $r > r_0$, которая принадлежит плоскости движения звезд. Можно сказать, что движение массивных звезд представляет собой

круговой массовый ток $J^{(m)}(r_0)$ такой, что соответствующая величина $B^{(g)}(r)$ может быть рассчитана по закону Био-Савара. Компонента $B^{(g)}(r)$, перпендикулярная плоскости, пропорциональна $B_z^{(g)}(r) \sim J^{(m)}(r)/r$. (Заметим, что она направлена “вверх” и “вниз” снаружи и внутри контура, достигая максимума в его центре). Центробежная сила $m \frac{v_{orb}^2}{r}$, действующая на обращающуюся планету, равна сумме обычного Ньютоновского притяжения $F_N^{(g)} = \frac{mC_1}{r^2}$, и “грави-Лоренцевой” силы $F_L^{(g)} = \frac{mv_{orb}C_2(r_0)}{r}$, где $C_1, C_2(r_0)$ - константы. Таким образом, квадрат орбитальной скорости планеты v_{orb}^2 , будет равен

$$v_{orb}^2 = \frac{C_1}{r} + v_{orb}C_2(r_0)$$

Нетрудно видеть, что у этого уравнения два корня, и при $r \rightarrow \infty$ первый из них дает Ньютоновский вклад $v_{orb} = 0$, а второй соответствует

$$v_{orb}|_{r \rightarrow \infty} = C_2(r_0),$$

что приводит к плоской кривой вращения.

Для более сложной системы, например, для спиральной галактики, следует обобщить понятие r_0 и найти эффективный радиус R_e . Это можно сделать, например, так

$$I_{eff} \sum I_n = MR_e^2 \Rightarrow R_e^2 = \sqrt{\frac{I_{eff}}{M}}$$

где I_{eff} – момент инерции системы с общей массой. Таким образом, рассматривая теория дает наблюдаемые плоские кривые вращения, причем R_e характеризует конкретную галактику, а эффективная угловая скорость Ω_e может быть найдена из условия $I_{eff}\Omega_e = L_{eff} = \sum L_n$, где L_n – моменты импульсов тел составляющих систему. Таким образом, получаем

$$\Omega_e = \frac{L_{eff}}{I_{eff}}$$

Внутри окружности радиуса R_e компонента $B_z^{(g)}(r)$ меняет знак и $F_L^{(g)}$ становится отталкивающей силой, действующей от центра распределения вращающихся масс. Это означает, что областью устойчивости является некоторое характерное расстояние, которое не является микроскопическим для масштаба рассмотрения, как радиус Шварцшильда. Поэтому можно рассмотреть в этой связи и проблему глобулярных кластеров, упомянутую в разделе 9.

Оценим величину $C_2(R_e) \sim J^{(m)}(R_e)$. Массовый ток $J^{(m)}(R_e) \sim \frac{M}{T}$, где масса M пропорциональна площади спиральной галактики R_e^2 , а период $\sim R_e^{3/2}$ в соответствии с законом Кеплера. Это дает $J^{(m)}(R_e) \sim \sqrt{R_e}$. Поскольку светимость L_{lum} также пропорциональна площади галактики, получим $R_e \sim \sqrt{L_{lum}}$. Поэтому второй корень выражения для квадрата орбитальной скорости дает

$$v_{orb}|_{r \rightarrow \infty} \sim L_{lum}^{1/4},$$

что соответствует закону Талли-Фишера (20).

Для того, чтобы найти дополнительное ускорение $a_L = 2[v, \Omega]$, где Ω определена выше, удобно воспользоваться электромагнитной аналогией и выбрать

$$rot_{(x)} \frac{\partial \varepsilon_{11}}{\partial y} = B^{(g)}; B_z^{(g)}(R_e) \sim \frac{J^{(m)}(R_e)}{r}$$

для данного распределения движущихся масс. Если выбрать приближение в виде кругового массового тока $J^{(m)}$ с радиусом R_e и скоростью движения масс V , то

$$\Omega \sim \frac{G}{c^2 r} \frac{MV}{R_e}$$

и дополнительное ускорение тела пропорционально

$$a_L \sim v \frac{GMV}{c^2 r R_e}$$

Отношение этого ускорения к Ньютоновскому ускорению $a_N = \frac{GM}{r^2}$ равно

$$\frac{a_L}{a_N} \sim \frac{vV}{c^2} \frac{r}{R_e} = \frac{vr}{c^2} \Omega$$

и можно определить области, где специфика анизотропии пространства становится существенной. Для заданной частицы, двигающейся со скоростью v на расстоянии r от центра, это отношение дает

$$\frac{a_L}{a_N} \sim \frac{vr}{c^2} \frac{L_{eff}}{I_{eff}},$$

где L_{eff} и I_{eff} характеризуют распределенный источник гравитации.

2.12 Интерпретация известных наблюдений, космологические следствия и специфический эффект

Как следует из полученных выражений, предложенная анизотропная модификация теории естественным образом обеспечивает плоский характер кривых вращения и выполнение закона Талли-Фишера. Проблема, связанная с различием гравитационного взаимодействия в плоскости галактики и вне ее также теряет остроту, поскольку движение, перпендикулярное галактической плоскости, не испытывает воздействия дополнительных сил, зависящих от скоростей, и таким образом, соответствующие “законы гравитации” действительно различаются. Проблема глобулярных кластеров может оказаться связанной с существованием характерного радиуса устойчивости, в окрестности которого сила меняет знак.

На данном этапе развития модели анизотропной геометродинамики количественные расчеты всех этих эффектов еще не получены. Но нетрудно видеть, что предложенная модификация теории гравитации приводит к дополнительному взаимодействию, обусловленному собственными движениями как галактик в кластере, так и в самих галактиках. Это может позволить обойтись и без введения темной материи для объяснения астрофизических эффектов, хотя и не будет противоречить самой возможности существования таких объектов, как бозоны Хиггса.

Как уже упоминалось, идея расширяющейся Вселенной утвердилась в связи с измеренным Хабблом красным смещением спектральных линий удаленных галактик, которое принято объяснять с помощью эффекта Доплера, и в настоящее время общепринятыми являются космологические модели расширяющейся Вселенной с метриками типа Фридмана-Робертсона-Уолкера, соответствующие классической геометродинамике. Именно с этих позиций трактуется и обнаруженное “ускорение расширения”, потребовавшее учета в теории

λ -члена, соответствующего взаимодействию отталкивания. Но теперь можно предположить, что красное смещение является *гравитационным* красным смещением, обусловленным тангенциальными движениями далеко удаленных объектов. Эта идея находит подтверждение в наблюдениях тангенциального движения квазаров, которое, как показывают наблюдения, происходит при удивительно высоких скоростях (рис.5), и точка зрения Маха о влиянии “удаленных звезд” на инерцию/гравитацию приобретает новый смысл. Таким образом, ока-

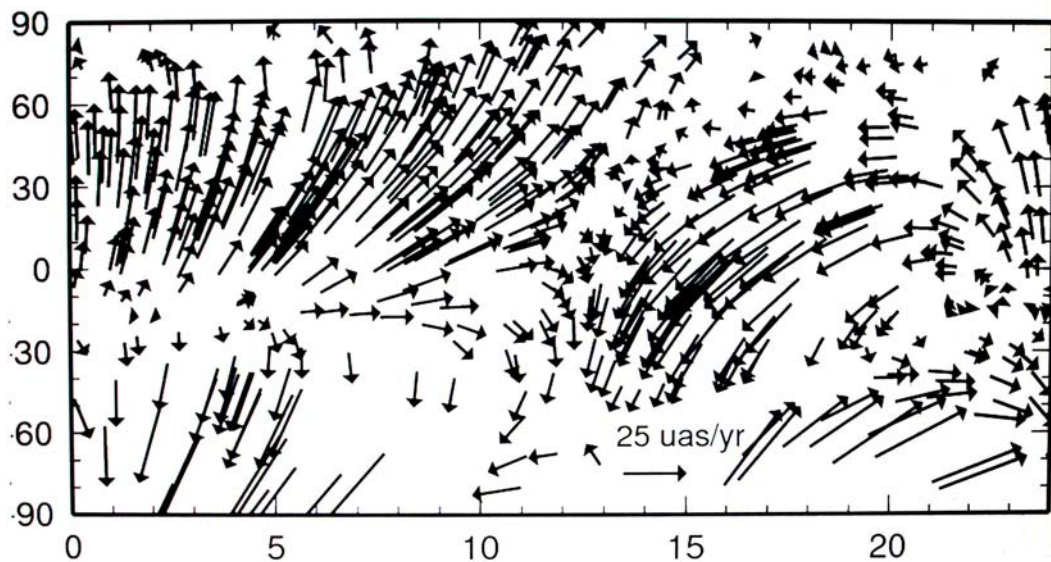


Рис. 5

зывается, что в рамках анизотропной геометродинамики наблюдаемые эффекты могут найти объяснение и без предположения о радиальном расширении Вселенной, требующем введения темной энергии, или наряду с ним, но при учете тангенциального движения объектов, вызывающего анизотропию метрики. Это означает, что в астрофизике и космологии повышается роль гидродинамики, поскольку любые скопления масс, такие, как галактики, кластеры галактик и сама Вселенная, можно интерпретировать как турбулентные вихри с присущими им свойствами взаимодействия, проявляющимися на соответствующих масштабах.

В каких случаях и в какой степени можно пренебречь эффектами крупномасштабных движений, следует каждый раз решать отдельно, ориентируясь на последнее из полученных соотношений, которое устанавливает количественный критерий для учета анизотропии. Заметим также, что при сопоставлении теории с измерениями на больших масштабах необходимо модифицировать как уравнение Эйнштейна, так и электромагнитное волновое уравнение с учетом анизотропной метрики.

Уже эти первые результаты носят слишком фундаментальный характер для того, чтобы можно было их безоговорочно принять. Как указывалось во Введении, любая новая теория должна не только включать в себя все результаты старой и решать некоторые ее проблемы, но и предсказывать специфический проверяемый эффект. Ясно, что эффекты типа классических тестов ОТО здесь не годятся, поскольку масштаб рассмотрения существенно иной. Следует предложить и рассчитать эксперимент, как минимум, в масштабах галактики. Удачным обстоятельством является то, что наша галактика – Млечный путь – является плоской спиральной галактикой, и эффекты анизотропии, если она имеется, должны проявляться и

в ней. Поэтому экспериментом, который может позволить исследовать геометрические свойства физического пространства, являются наблюдения гравитационно-волнового эффекта нового типа, так называемого оптико-метрического параметрического резонанса (ОМПР).

Суть его состоит в том, что при выполнении набора определенных условий на астрофизическую систему гравитационные волны (ГВ), испускаемые периодическим источником, например, тесной двойной звездной системой или пульсаром, начинают взаимодействовать с монохроматическим электромагнитным излучением космического мазера резонансным образом. В отличие от всех (18-ти!) известных эффектов, пригодных для детектирования ГВ, этот эффект в силу своего резонансного характера имеет не первый, но нулевой порядок по чрезвычайно малой амплитуде ГВ и может наблюдаться с помощью уже имеющейся аппаратуры (см. Приложение 4). При построении соответствующей модели, подлежащей расчету, используется выражение для метрики пространства, поэтому такой эксперимент подходит для заявленной цели – исследования геометрии пространства в масштабах галактики. Выполненный расчет указывает, что при различных ориентациях системы (источник ГВ)-(мазер)-(Земля) относительно плоскости галактики условия ОМПР будут различаться, что позволит судить о наличии анизотропии с помощью измерений и, возможно, оценить ее параметры. В настоящее время соответствующие исследования ведутся в Пущинской РАО РАН.

Заклучая изложение, подчеркнем следующее. Проблемы, стоящие в настоящее время перед ОТО, представляются слишком серьезными, чтобы можно было справиться с ними, вводя поправки, не выходящие за пределы ее основ, и наращивая сложность вычислений. Следует обратиться к глубоким физическим и математическим идеям, лежащим в ее основе, и проанализировать именно их применимость к описанию физической реальности с учетом данных наблюдений. Рассмотренная здесь возможная локальная анизотропия пространства является лишь одним из возможных подходов, с помощью которого наука может выйти на следующий виток своего развития.

Приложение 1

Метрический тензор и метризация

Исследуем пару, состоящую из единичных векторов линейного пространства E_1 , связанного с отображением, т.е. из ортов канонического базиса. Изменяя систему координат, получим

$$(\vec{e}_i, \vec{e}_k) = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{k'}}{\partial x^k} (\vec{e}_{i'}, \vec{e}_{k'})$$

Элементов, стоящих справа в идее множителя, получится n^2 штук, и говорят, что рассматриваемая пара преобразуется как дважды ковариантный тензор, который обозначим

$$(\vec{e}_i, \vec{e}_k) = g_{ik}$$

Этот тензор немедленно появится, когда мы рассмотрим пару векторов

$$(V^i \vec{e}_i, W^k \vec{e}_k) = V^i W^k g_{ik}$$

Здесь каждый элемент введенного тензора, сконструированного из ортов канонического базиса, умножен на соответствующее число. Существует и его аналог, связанный с сопряженным базисом. Действительно,

$$(\cdot, \cdot) = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^k}{\partial x^{k'}} (\cdot, \cdot)$$

и говорят, что такая пара представляет собой контравариантный тензор второго ранга, т.е.

$$(\cdot, \cdot) = g^{ik}$$

Коэффициенты, возникающие при преобразовании координат, представляют собой компоненты матрицы Якоби (якобиан). Поэтому можно найти, как связаны между собой эти два тензора. Фактически, можно сказать, что тензор – один и тот же, а рассматриваются то его ковариантные компоненты, то контравариантные, относятся же они к одному и тому же объекту.

Исследуем свойства тензора g_{ik} . Представим один из уже упомянутых векторов в виде разложения по сопряженному базису

$$\vec{V} = V_i$$

Тогда ту же пару можно записать так

$$(V_i, W^k \vec{e}_k) = V_i W^k \vec{e}_k = V_i W^k \delta_k^i = V_k W^k$$

Откуда

$$g_{ik} V^i = V_k$$

Таким образом, тензор g_{ik} представляет собой оператор перехода от одной из контравариантных компонент вектора к его ковариантной компоненте с другим значком (эффективно это выглядит как опускание значка) путем суммирования всех контравариантных компонент с соответствующими коэффициентами. Такая возможность является фундаментальной для тензоров, и сам тензор g_{ik} называется фундаментальным.

Аналогично

$$(V_i, W_k) = V_i W_k = V_i W_k g^{ik}$$

и

$$(V_i, W^k \vec{e}_k) = V_i W^k \vec{e}_k = V_i W^k \delta_k^i = V_i W^i$$

Откуда

$$g^{ik} W_k = W^i$$

т.е. дважды контравариантный тензор g^{ik} выполняет обратный переход от ковариантной компоненты вектора к его контравариантной компоненте с другим значком. Прямое вычисление дает

$$g_{ik} g^{kj} = \vec{e}_i \vec{e}_k = \vec{e}_i = \delta_i^j$$

Компоненты дважды контравариантного тензора g^{ik} могут быть найдены как алгебраические дополнения к матричным элементам дважды ковариантного тензора

$$g^{ik} = \frac{1}{g} \frac{\partial g}{\partial g_{ik}}$$

где g – определитель матрицы.

Введем еще одну характеристику геометрических объектов. Сопоставим каждому тензору число, не зависящее от выбора координат, которое будем называть нормой. Для тензора

нулевого ранга – скаляра – это будет сам скаляр. Для тензора первого ранга – вектора – рассмотрим сначала выражение

$$||\vec{V}||^2 \equiv V^i V^k g_{ik} \quad (33)$$

Непосредственной проверкой нетрудно убедиться, что это число, представляющее собой сумму произведений контравариантных компонент вектора на соответствующие компоненты фундаментального тензора, не зависит от выбора координат. Это же число может быть получено и иначе

$$||\vec{V}||^2 \equiv V^i V^k g_{ik} = V_i V_k g^{ik} = V_i V^i \quad (34)$$

В отличие от выражения (33) здесь могут использоваться ковариантные компоненты вектора пространства, связанные только с сопряженным базисом, и контравариантный фундаментальный тензор, элементами которого являются комбинации ортов сопряженного базиса. Это дает возможность говорить о свойствах пространства и его векторов, не требуя введения какого-либо отображения непосредственно. Традиционно в качестве нормы выбирается корень квадратный из введенного выражения. Также эта величина называется длиной или модулем вектора.

Введем бинарную операцию. Как и ранее, рассмотрим два вектора, у которых определены нормы (длины)

$$\begin{aligned} ||\vec{V}|| &= \sqrt{V^i V^k g_{ik}} \\ ||\vec{W}|| &= \sqrt{W^i W^k g_{ik}} \end{aligned}$$

Введем скалярное 2-произведение

$$\vec{V} \cdot \vec{W} = g_{ik} V^i W^k \equiv ||V|| \cdot ||W|| \cos(\vec{V}, \vec{W}), \quad (35)$$

с помощью которого каждым двум векторам сопоставляется, представляющий собой обратную тригонометрическую функцию от выражения

$$\cos(\vec{V}, \vec{W}) = \frac{\vec{V} \cdot \vec{W}}{||V|| \cdot ||W||} = \frac{g_{ik} V^i W^k}{\sqrt{g_{lm} V^l V^m} \sqrt{g_{jn} W^j W^n}} \quad (36)$$

Таким образом, фундаментальный тензор g_{ik} позволяет ввести привычные понятия длины вектора и угла между векторами на основе бинарной операции (скалярного 2-произведения), и понятия нормы. По этой причине в данном случае этот тензор называют также метрическим тензором, т.е. таким объектом, который позволяет ввести меру для векторов. Одновременно он, а точнее, его контравариантный аналог характеризует свойства пространства, в котором заданы вектора.

Замечание: метрический тензор может и не совпадать с фундаментальным тензором.

Рассмотрим операции более высоких порядков. Так, вводя норму иначе, получим

$$\begin{aligned} ||\vec{V}||^3 &\equiv V^i V^j V^k \vec{e}_i \vec{e}_j \vec{e}_k = V^i V^j V^k g_{ijk} \\ ||\vec{V}|| &= \sqrt[3]{V^i V^j V^k g_{ijk}} \end{aligned} \quad (37)$$

где g_{ijk} – некоторый трижды ковариантный (метрический) тензор, составленный из ортов канонического базиса. Такое выражение также является числом, не зависящим от координат.

Если ассоциативность не предполагается, то скалярное 3-произведение и 3-угол определяются следующим образом

$$\begin{aligned}\vec{V} \cdot \vec{U} \cdot \vec{W} &= g_{ijk} V^i U^j W^k \equiv ||V|| \cdot ||U|| \cdot ||W|| \cos(\vec{V}, \vec{U}, \vec{W}) \\ \cos(\vec{V}, \vec{U}, \vec{W}) &= \frac{g_{ijk} V^i U^j W^k}{\sqrt[3]{V^i V^j V^k g_{ijk}} \sqrt[3]{U^i U^j U^k g_{ijk}} \sqrt[3]{W^i W^j W^k g_{ijk}}}\end{aligned}\quad (38)$$

Если же ассоциативность имеет место, то, представив вектора $\vec{V}, \vec{U}, \vec{W}$ в виде их разложений по каноническому базису, как в формуле (37), нетрудно убедиться, что $\cos(\vec{V}, \vec{U}, \vec{W})$ будет представлять собой вектор. Этот объект не будет изменяться при изменении координат, но действия с ним будут отличаться от действий с обычными тригонометрическими функциями.

Продолжая увеличивать число компонент, участвующих в определении нормы, получим

$$\begin{aligned}||\vec{V}||^4 &\equiv V^i V^j V^k V^l \vec{e}_i \vec{e}_j \vec{e}_k \vec{e}_l = V^i V^j V^k V^l g_{ijkl} \\ ||\vec{V}|| &= \sqrt[4]{V^i V^j V^k V^l g_{ijkl}}\end{aligned}\quad (39)$$

где g_{ijkl} – ковариантные компоненты некоторого (метрического) тензора четвертого ранга. Если для операций с единичными векторами канонического базиса имеет место ассоциативность, то этот метрический тензор может быть выражен через фундаментальный тензор

$$\begin{aligned}||\vec{V}||^4 &\equiv V^i V^j V^k V^l \vec{e}_i \vec{e}_j \vec{e}_k \vec{e}_l = V^i V^j g_{ij} V^k V^l g_{kl} = V^i V_i V^k V_k \\ ||\vec{V}|| &= \sqrt[4]{V^i V^j V^k V^l g_{ij} g_{kl}} = \sqrt[4]{V^i V_i V^k V_k}\end{aligned}\quad (40)$$

и норма (40) совпадает с обычной нормой (33) – той, которая определяется с помощью бинарной операции. При этом и углы сохраняют свой привычный смысл. Если ассоциативность не предполагается, то скалярное 4-произведение и 4-угол будут иметь вид

$$\begin{aligned}\vec{V} \cdot \vec{U} \cdot \vec{W} \cdot \vec{Z} &= g_{ijkl} V^i U^j W^k Z^l \equiv ||V|| \cdot ||U|| \cdot ||W|| \cdot ||Z|| \cos(\vec{V}, \vec{U}, \vec{W}, \vec{Z}) \\ \cos(\vec{V}, \vec{U}, \vec{W}, \vec{Z}) &= \frac{g_{ijkl} V^i U^j W^k Z^l}{\sqrt[4]{V^i V^j V^k V^l g_{ijkl}} \sqrt[4]{U^i U^j U^k U^l g_{ijkl}} \sqrt[4]{W^i W^j W^k W^l g_{ijkl}} \sqrt[4]{Z^i Z^j Z^k Z^l g_{ijkl}}}\end{aligned}\quad (41)$$

Здесь $\cos(\vec{V}, \vec{U}, \vec{W}, \vec{Z})$ представляет собой ковариантные компоненты тензора 4-го ранга, структура которого совпадает со структурой метрического тензора. В выражениях для углов (38) и (41) суммирование по повторяющимся индексам проводится только внутри каждого из входящих в них сомножителей по отдельности.

Приложение 2

Принцип наименьшего действия

Постулаты, определения, законы

Введем понятие сигнала – информационной единицы, выделяемой из окружающего мира и способной обладать различными свойствами. Физическая природа сигнала также может быть различной, и каждый раз должна быть оговорена отдельно. Например, она может быть механической, акустической, электромагнитной. . . Рассмотрим два события: испускание сигнала, природу которого пока не оговариваем, в точке 1 и прием его в точке 2 в соответствующие моменты времени. Таким образом, можно сказать, что между событиями испускания-приема имеется причинно-следственная связь (физика) или отношение порядка (математика). Примем два исходных постулата. Один из них известен как первый закон (динамики) Ньютона

- Существуют такие системы отсчета, называемые инерциальными (ИСО), в которых тела сохраняют состояния покоя или равномерного прямолинейного движения пока и поскольку сумма всех сил, действующих на них, равна нулю.

Второй постулат представляет собой так называемый принцип относительности Галилея, состоящий в том, что

- Все законы физики (механики) одинаковы во всех инерциальных системах отсчета.

Особенностями этих постулатов является то, что первый из них в строгом (физическом) смысле никогда не верен (невозможно исключить влияние тела отсчета, с которым связана система отсчета), а второй – невозможно проверить (как в силу своего исчерпывающего характера, так и в силу указанного отсутствия инерциальных систем).

Введем две идентично устроенные ИСО, и будем пока считать оси координат ортогональными. Располагая эталоном, наблюдатель мог бы попытаться непосредственно измерить расстояние между точками 1 и 2, находясь в любой из ИСО, (и для дальнейших рассуждений уже потребовался бы постулат о неизменности свойств самого эталона при переходе из системы в систему). Однако, как правило, эталон используют для градуировки осей систем координат, (что, вообще говоря, требует такого же или аналогичных постулатов), после чего вводят определение меры удаленности двух точек друг от друга, т.е. расстояния R , которое может быть вычислено по некоторым определенным правилам, которые, вообще говоря, могут быть различными.

При определении расстояния между точками 1 и 2 в качестве экспериментальной основы измерений координат используем свет. Тогда необходимо установить соответствие между абстрактными (геометрическими) прямыми, используемыми при расчете, и траекториями распространения света, используемыми при измерениях. Исторически это соответствие описывалось с помощью одного из двух постулатов:

- свет распространяется по кратчайшему пути (т.е. в мире, геометрия которого евклидова, световые лучи есть прямые линии – геометрическая оптика),
- свет распространяется по пути, требующим кратчайшего времени (принцип Ферма). Для однородного евклидова пространства этот постулат сводится к предыдущему, а для неоднородного – например, состоящего из двух полупространств, – позволяет получить законы отражения и преломления без ссылок на природу света и характер его взаимодействия с веществом, заполняющим пространство.

Могут существовать поля, влияющие на движение исследуемого сигнала (в том числе, в частности, и светового), которое может быть и неинерциальным. Будем считать, что, судя по измерениям, физический сигнал движется из точки 1 в точку 2 с переменной (не меняющей знака, конечной) скоростью по некоторой траектории, длина которой может отличаться от расстояния между точками. Длину l этой траектории можно найти, если известна (измерена, вычислена) зависимость скорости сигнала от естественного параметра – времени – в данной ИСО

$$l = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt \quad (42)$$

Это обстоятельство позволяет при необходимости переходить от одного естественного параметра – времени, к другому, не менее естественному, – пройденному пути.

Кроме того, надо уточнить, с помощью какого эффекта измеряется скорость и положение сигнала. По-видимому, на современном этапе развития науки естественно и здесь опираться на распространение света.

Для абстрактного (математического) описания движения “сигнала” во внешних полях потребуются так называемые “законы физики” – абстрактные математические структуры (уравнения), вид которых был получен или выбран на основании интуитивных соображений. Последние могут носить как дедуктивный характер, так и индуктивный, основанный на логическом анализе наблюдений природы. Для того чтобы результаты, получаемые с помощью математических уравнений, можно было использовать, потребуется постулат о выборе той или иной геометрии для описания физического (измеряемого) мира и последующая проверка соответствия модели и мира. Этот выбор геометрии естественно связать с основным физическим эффектом, используемым при измерениях, т.е. геометрию удобно выбрать такой, что “закон физики”, описывающий эффект, применяемый для измерений, наиболее просто использовать в максимально возможном числе случаев.

Так, если использовать измерения, основанные на световых сигналах, и считать свет электромагнитной волной, то геометрию следует выбрать такой, чтобы уравнения Максвелла сохраняли свой вид (“наиболее просто использовать”) во всех ИСО (“в максимально возможном числе случаев”). Фактически, это и было сделано Лоренцем, а затем Эйнштейном при построении специальной теории относительности (СТО). Соответствующая геометрия оказалась римановой и включала одновременно и пространство, и время, что требовалось преобразованиями Лоренца, обеспечивающими инвариантность уравнений Максвелла при переходе из одной ИСО в другую. Естественным параметром решений уравнений движения и траекторий стало уже не время или соответствующий ему пройденный путь (как в евклидовом пространстве), а интервал. Время же стало лишь одной из координат четырехмерного пространства-времени, не слишком отличающейся от трех остальных.

Рассмотрим более подробно движение физического сигнала в силовом поле. Наиболее последовательным представляется дедуктивный аксиоматический подход Гамильтона-Лагранжа, хотя следует отметить, что эти последние его качества вторичны и имеют в основе индуктивные эмпирические законы. Поэтому сделаем несколько предварительных замечаний.

Пусть теперь физический сигнал представляет собой тело, характеризующееся двумя измеряемыми числовыми коэффициентами, называемыми массой m и *электрическим зарядом* e . Масса характеризует инертность тела, т.е. его “стремление” сохранять состояние покоя или равномерного прямолинейного движения, заряд – способность участвовать в электромагнитном взаимодействии. Тело перемещается из точки 1 в точку 2 и находится во внешнем электромагнитном поле. О наличии поля можно узнать, наблюдая движение различных заряженных (наблюдателем) и незаряженных тел в рассматриваемой области пространства. Электромагнитное поле характеризуется т.н. *скалярным потенциалом* $\varphi(q, t)$ и *векторным потенциалом* $\vec{A}(q, t)$, где q – радиус-векторы точек пространства. Измерения, выполненные с учетом системы единиц, показывают, что движение заряженного тела в электромагнитном поле согласуется с некоторыми уравнениями, известными как “законы физики”.

Так, уравнение

$$\frac{dp}{dt} = Fp \equiv mv \quad (43)$$

называется основным уравнением динамики или законом Ньютона. Здесь – вектор импульса или количества движения, F – сила, мера воздействия на тело со стороны внешних тел или полей, приводящего к появлению ускорения. Фактически, выражение (43) есть наиболее конструктивное определение понятия силы. Нетрудно увидеть его место в ряду следующих исторически встречавшихся определений: сила есть мера воздействия на тело со стороны внешних тел, приводящего к появлению или деформации тела (например, закон упругости Гука $F = -kx$); сила есть мера воздействия на тело со стороны внешних тел, приводящего к появлению у тела *скорости*, – в том числе и постоянной (так ее определял не учитывавший трения Галилей на основании опытов с тележкой). Ньютон же обнаружил, что, прикладывая различно сжатую пружину к одному и тому же телу, получим $\frac{a_1}{a_2} = \frac{F_1}{F_2}$, где – ускорение, а, прикладывая одинаково сжатую пружину к различным телам, получим $\frac{a_1}{a_2} = \frac{m_2}{m_1}$, где m пропорционально количеству вещества. Уравнение (43) служит одновременно обобщением этих экспериментальных фактов, определением силы и основой для выбора системы единиц измерения. Словом называют также и воздействия различной природы (сила упругости, сила тяжести). Определяя зависимости природных сил от различных параметров и подставляя эти зависимости в уравнение (43), можно, пользуясь математическими методами, найти его решение. С точки зрения физики это означает возможность предсказать положение тела в различные моменты времени, найти траекторию его движения, сравнить полученные результаты с экспериментальными измерениями и при необходимости уточнить использованные зависимости для сил. В этом состоит научный метод Ньютона, используемый до сих пор.

В рассматриваемом случае сила F действует на заряженную частицу со стороны электромагнитного поля. При этом она удовлетворительно описывается выражением, известным как формула силы Лоренца

$$F = eE + \frac{e}{c}[vH] \quad (44)$$

где – напряженность электрического поля, – напряженность магнитного поля, – скорость света, возвращаясь к упомянутым ранее характеристикам поля, получим

$$E \equiv -\frac{1}{c} \frac{\partial A}{\partial t} - \text{grad} \varphi; H \equiv \text{rot} A \quad (45)$$

Воспользовавшись соотношением

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + (v \cdot \nabla)A$$

и подставляя (45) в (44), а затем в (43), получим уравнение движения сигнала в виде

$$\frac{d}{dt}(p + \frac{e}{c}A) = \frac{e}{c}(v \cdot \nabla)A + \frac{e}{c}[v \times \text{rot} A] - e \text{grad} \varphi \quad (46)$$

Сравнивая вид получившегося уравнения с видом основного уравнения динамики (43), можно (например, из эстетических соображений) ввести новое определение импульса для данной конкретной ситуации $P = p + \frac{e}{c}A$. Но можно этого и не делать. Уравнение (46) является примером “закона физики”, упоминавшегося в начале этого раздела.

Не будем теперь конкретизировать природу внешнего поля, но допустим лишь, что зависимость силы, действующей на тело, от параметров такова, что при описании ей может быть

сопоставлен градиент некоторой функции, называемой потенциалом поля, т.е.

$$F = -\frac{\partial U}{\partial q} \quad (47)$$

и уравнение движения тела, согласующееся с измерениями, по-прежнему имеет вид

$$\frac{dp}{dt} = -\frac{\partial U}{\partial q} \quad (48)$$

Тогда с математической точки зрения у этого дифференциального уравнения существует так называемый *интеграл движения*, имеющий простой вид

$$E = \frac{mv^2}{2} + U \equiv T + U \quad (49)$$

В физике он носит название полной механической энергии (излучение здесь пока не рассматриваем) и состоит из двух частей, каждой из которых обычно придают физический смысл. Первое слагаемое называется кинетической энергией и характеризует только свойство самого тела. Второе слагаемое называется потенциальной энергией и характеризует взаимодействие тела и физического поля, в котором тело движется. Закон сохранения полной механической энергии (физика), таким образом, эквивалентен наличию интеграла движения у дифференциального уравнения динамики (математика). Иными словами, если считать, что математика применима для описания физического мира, то измеряемая величина “полная механическая энергия” не может не сохраняться. Если же при использовании выбранной математической модели измеряемая энергия сохраняться не будет, то это укажет на наличие физических полей, не учтенных в модели. У уравнений динамики существуют и другие интегралы движения, также находящие применение в физике.

Проверяемая гипотеза о сохранении энергии, которая в данном случае просто следует из свойств математических структур, наводит на мысль сформулировать постулат, аналогичный принципу Ферма, для описания произвольных физических процессов. Помимо того, что он сообщит законченность теории, он позволит также рассмотреть пригодность той или иной геометрии для описания реального мира, используя не только световые, но и прочие явления.

Принцип наименьшего действия и канонические уравнения

Рассмотрим векторное пространство $q = \{q^i\}$, в котором задана скалярная функция $L(q, \dot{q}, t)$, зависящая от параметра t , и связанный с ней функционал, имеющий вид

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \quad (50)$$

Вариационное исчисление позволяет найти уравнения кривых, доставляющих этому функционалу экстремум. Заметим при этом, что свойство математической кривой быть экстремалью функционала не зависит от выбора системы координат (декартовой, полярной и т.д.). Но что еще более важно, – скаляр, стоящий под интегралом, есть объект, не зависящий от изменения выбранной системы координат.

Элементам этого (абстрактного) векторного пространства можно сопоставить некоторые обобщенные координаты, характеризующие движение системы, и соответствующие обобщенные скорости. Вначале, однако, для простоты сопоставим им радиус-вектора физических объектов, измеряемые в ИСО. При этом параметру t можно также сопоставить измеряемую физическую величину – время. Будем считать, что каждую физическую систему можно охарактеризовать некоторой определенной скалярной функцией $L(q, \dot{q}, t)$. Будем, как и ранее, следить за движением физического сигнала, который пока считаем точечной массой, в некоторой ИСО. Пусть в моменты времени t_1 и t_2 измерения показывают, что сигнал занимает определенные положения, характеризующиеся двумя наборами координат $q^{(33)}$ и $q^{(34)}$. Функционал (50) будем называть *действием*.

Постулируем принцип наименьшего действия (ПНД):

- любая физическая система движется таким образом, что действие, вычисленное для нее, имеет минимально возможное значение.

Из этого принципа можно с помощью варьирования функционала

$$\delta S = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt = 0 \quad (51)$$

получить так называемое уравнение Эйлера-Лагранжа (или уравнение Лагранжа) для функции $L(q, \dot{q}, t)$,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \quad (52)$$

которое фактически должно приводить к соответствующим экспериментально проверяемым уравнениям движения, иначе ПНД останется формальным математическим утверждением. Функция $L(q, \dot{q}, t)$, удовлетворяющая уравнению Лагранжа (52), называется функцией Лагранжа или лагранжианом. Таким образом, придерживаясь данного формализма, следует позаботиться о таком выборе функции $L(q, \dot{q}, t)$, что принятый постулат будет недвусмысленно согласован с известными законами физики. Если это удастся сделать, то для теоретического исследования физических процессов можно будет использовать мощный аппарат вариационного исчисления. Сделаем важное

Замечание 1:

- В чем бы ни состояли требования к функции Лагранжа, и каким бы ни был ее вид, в ее выборе неизбежно имеется произвол. Функциями Лагранжа всегда окажется целый класс функций, отличающихся друг от друга на полную производную по времени от произвольной функции координат и времени $f(q, t)$. Выберем

$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q, t)$ (78) Подставляя (78) в (51), видим, что при варьировании дополнительный член пропадает. Это означает, что уравнения движения, получаемые с помощью новой функции Лагранжа и пригодные для описания измеряемого движения сигнала, останутся теми же, что и раньше. Иными словами, аддитивная составляющая *функции Лагранжа*, которая может быть представлена в виде полной производной по времени, не будет соответствовать никакой наблюдаемой или измеряемой величине. Естественно, что к аддитивной составляющей в виде константы это тоже относится.

Сложившаяся (вероятно, благодаря известному курсу физики) традиция предлагает в качестве эффективного базисного допущения при дедуктивном выборе $L(q, \dot{q}, t)$ учесть однородность пространства и времени (независимость L от q и t) и изотропность пространства (независимость L от v). Эта традиция как таковая удовлетворительной не является ни с точки зрения рассматриваемых здесь вопросов, ни с учетом существования движений, описываемых не дифференцируемыми функциями, ни в связи с приписыванием физическим пространству и времени субстанциональных свойств. Апеллировать к возможности подобного дедуктивного выбора тем более не стоит, если обратить внимание на (столь же традиционный) способ выбора функции Лагранжа в случаях, не ограниченных механикой. Гораздо более важно учесть следующее обстоятельство, которое зачастую обсуждается в курсах физики недостаточно.

Исследуя экспериментально измеряемые величины и сопоставляя им математические объекты, например вектора, следует внимательно следить за адекватностью таких сопоставлений, имея в виду последующие операции с этими объектами. В частности, измеряемые компоненты вектора скорости можно сопоставлять только какому-то одному классу геометрических объектов, например, контравариантным компонентам вектора. Но, как известно из дифференциальной геометрии, для получения скаляра из (контравариантного) вектора понадобится так называемая 1-форма (ковариантный вектор). Ее компоненты могут быть получены из компонент имеющегося (контравариантного) вектора с помощью фундаментального (метрического) тензора.

Таким образом, при построении скаляра (функции Лагранжа) с помощью объектов, связанных с измеряемыми величинами, в получаемое выражение должен с необходимостью входить метрический тензор в явном виде. Он будет характеризовать геометрию (математического) пространства, которое используется для моделирования физической реальности. Этот момент является принципиально важным для построения фундаментальной теории.

Вернемся пока к построению функции Лагранжа. Постулируя, что для “свободной частицы” движение сигнала является равномерным и прямолинейным, и, убеждаясь, что это не противоречит соответствующим наблюдениям, запишем уравнение такого движения

$$q = q_1 + vt \quad (53)$$

Выберем функцию Лагранжа в виде

$$L = bv^2 = b\dot{q}^2; b = \text{const} \quad (54)$$

где квадраты вектора скорости представляют собой скалярные произведения этого вектора самого на себя. При этом квадраты соответствующих норм вектора равны

$$v^2 = (v, v) = v_i v^i = \|v\|^2; \dot{q}^2 = (\dot{q}, \dot{q}) = \dot{q}_i \dot{q}^i = \|\dot{q}\|^2 \quad (55)$$

Можно было бы записать функцию Лагранжа и так

$$L = bv^2 = (v^i)(bv_i). \quad (56)$$

Подстановка (54) в (52) даст то уравнение движения, которое требуется:

$$\ddot{q} = 0 \Rightarrow q = q_1 + vt \quad (57)$$

Привнесем физический смысл в рассматриваемые формальные математические структуры. С этой целью

- Выберем константу b из математического уравнения (54) в виде, имеющем физический смысл и соответствующую размерность: $b \equiv \frac{m}{2}$.

Тогда функция Лагранжа свободной частицы (движущейся равномерно и прямолинейно в трехмерном евклидовом пространстве) будет иметь вид

$$L = \frac{mv^2}{2} = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = T \quad (58)$$

что совпадает с выражением для кинетической энергии. Если же частица находится в поле сил, как в случае, соответствующем формуле (47), и движется неравномерно, то, выбирая функцию Лагранжа в виде

$$L = T - U \quad (59)$$

и подставляя это выражение в формулу (52), немедленно получим

$$\frac{dp}{dt} = -\frac{\partial U}{\partial q} \quad (60)$$

Сравнивая уравнения (60), (52) и (43), можно с помощью функции Лагранжа системы дать определения

- обобщенного импульса $p = \frac{\partial L}{\partial \dot{q}}$ (90)
- обобщенной силы $F = \frac{\partial L}{\partial q}$ (29)

Оказалось, что “подправить внешний вид” уравнения (46) можно и с точки зрения требований формализма. Полученное в рамках рассматриваемого вариационного формализма уравнение (60) – то же самое проверяемое дифференциальное уравнение движения, что и (48), с интегралом движения $E = T + U$.

Именно такую формулировку ПНД дал Гамильтон:

- движения механической системы совпадают с экстремалиями функционала (50), где L есть разность кинетической и потенциальной энергий системы.

Уравнение Лагранжа – дифференциальное уравнение второго порядка для координаты. В этой связи оно эквивалентно системе уравнений первого порядка, в которой вторая независимая переменная – компонента скорости. Но Гамильтон предложил перейти от одного набора независимых переменных к другому – использовать преобразования Лежандра, состоящие в переходе от функций на данном линейном пространстве к функциям на сопряженном пространстве, и ввести определения обобщенных импульсов и обобщенных координат. В этом случае от (некоторой) функции Лагранжа следует перейти к (соответствующей) функции Гамильтона по формуле

$$H(p, q, t) = p\dot{q} - L(q, \dot{q}, t) \quad (61)$$

которая и представляет собой преобразование Лежандра от функции Лагранжа, рассматриваемой как функция $\dot{q}$. Заметим, что преобразованием Лежандра от функции Гамильтона

будет исходная функция Лагранжа. Тогда вместо системы уравнений Лагранжа, следует использовать уравнения Гамильтона или так называемые канонические уравнения

$$\dot{p} = -\frac{\partial H}{\partial q} \quad (62)$$

$$\dot{q} = \frac{\partial H}{\partial p} \quad (63)$$

Если $q(t)$ удовлетворяет уравнениям Лагранжа, то $(p(t), q(t))$ удовлетворяют уравнениям Гамильтона и наоборот. Нетрудно убедиться, что функция Гамильтона (или гамильтониан) есть полная энергия

$$H = T + U \quad (64)$$

В терминах гамильтониана физический закон сохранения энергии следует из математического формализма особенно просто. Поскольку

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \quad (65)$$

то, если функция Гамильтона явно не зависит от времени, величина полной энергии является постоянной. Другие физические законы сохранения также коррелируют с математическими особенностями. Так, если координата q_i не входит в гамильтониан явно (при этом она называется циклической), то соответствующий ей импульс p_i является интегралом движения (“закон сохранения импульса”).

Существует развитая теория канонических уравнений, играющая значительную роль в теоретической и математической физике. Можно сказать, что, построив функцию Гамильтона для описания некоторой физической системы, сопоставимую с наблюдаемыми и измеряемыми процессами, мы полностью решаем задачу предсказания поведения системы в различных условиях.

Введенная в начале данного раздела величина действия $S = \int_{t_1}^{t_2} L dt$ тесно связана с каноническими уравнениями. Если фиксировать только начальную точку движения, но по-прежнему требовать, чтобы оно происходило по физической траектории, т.е. удовлетворяло уравнениям Лагранжа, и учитывать определение обобщенного импульса $p = \frac{\partial L}{\partial \dot{q}}$, то получим соотношение

$$p_i = \frac{\partial S}{\partial q^i} \quad (66)$$

где i нумерует (обобщенную) координату или степень свободы.

При движении системы вдоль наблюдаемой траектории полная производная по времени равна

$$\frac{dS}{dt} = L \quad (67)$$

по определению. С другой стороны

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \sum_i \frac{\partial S}{\partial q^i} \dot{q}^i = \frac{\partial S}{\partial t} + \sum_i p_i \dot{q}^i \quad (68)$$

Поэтому

$$\frac{\partial S}{\partial t} = L - \sum_i p_i \dot{q}^i \quad (69)$$

Откуда

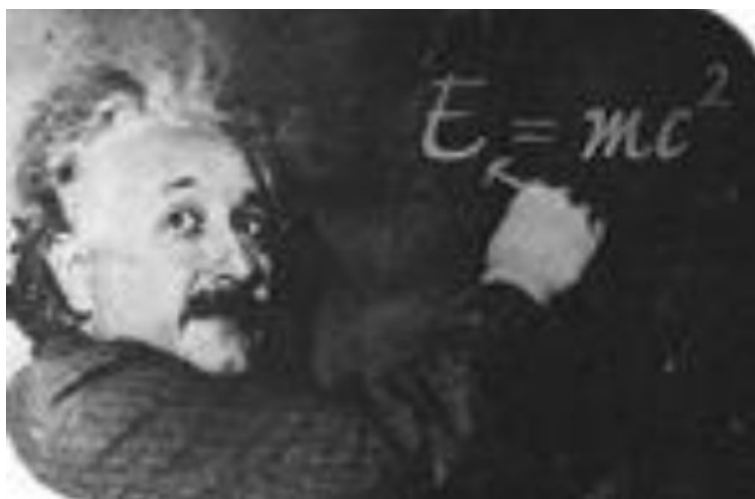
$$H = -\frac{\partial S}{\partial t} \quad (70)$$

Формулы (66) и (70) и дают искомую связь действия и канонических уравнений. Отметим разные знаки в правых частях указанных формул.

Приложение 3

Задача Эйнштейна

Считается, что именно он составил ее и предположил, что лишь 2% решающих смогут ее осилить. Нетрудно убедиться, что характер “обработки данных” весьма похож на тот одно-временный учет многих обстоятельств, одни из которых принимаются во внимание прямо сейчас, а другие возникнут на следующем шаге, который возникает при построении теории.



Есть пять домов, каждый своего цвета. В 5 домах живет 5 человек 5 разных национальностей. Национальности не повторяются. Каждый жилец пьет только 1 напиток, курит только 1 сорт сигарет и имеет 1 домашнее животное. Напитки, марки сигарет и животные не повторяются.

Условия:

1. Англичанин живёт в красном доме.
2. Швед держит собаку.
3. Датчанин пьёт чай.
4. Зелёный дом стоит слева от белого.

5. Жилец зеленого дома пьет кофе.
6. Тот, кто курит Pall Mall, имеет птицу.
7. Жилец жёлтого дома курит Dan Hill.
8. Обитатель среднего дома пьет молоко.
9. Любитель Malbogo живёт около человека, который владеет кошкой.
10. Первый дом - норвежца.
11. Тот, кто курит Malbogo, живет по соседству с мужчиной, который пьет воду.
12. Счастливый обладатель лошади живёт возле курильщика сигарет Dan Hill.
13. Тот, кто портит себе здоровье Vinfield'ом, ещё и пиво пьет.
14. Норвежец живёт возле голубого дома.
15. Немец курит Rottmanns.

Вопрос: КОМУ ПРИНАДЛЕЖИТ РЫБА??? (это одно из домашних животных)

Система условий полная, что предполагает одно ЕДИНСТВЕННОЕ правильное решение. При этом в предъявленном решении необходимо продемонстрировать выполнение всех указанных условий.

Приложение 4

Опτικο-метрический параметрический резонанс как инструмент исследования геометрических свойств пространства-времени

Эффект ОМПР является частным случаем явления оптико-механического параметрического резонанса. Оно состоит в том, что, если компонента скорости атома, параллельная волновому вектору поля, по какой-либо причине меняется с определенной частотой, то на определенной частоте, близкой к частоте атомного перехода будет происходить периодическое усиление и ослабление сигнала с частотой колебаний атома. Это явление связано с перераспределением энергии по частотам в результате параметрического резонанса. В обычных условиях такой сигнал наблюдаться не будет из-за усреднения по времени, но его можно выделить с помощью специальной обработки. Оказывается, что величина такого сигнала сопоставима с величиной пика обычного сигнала, характеризующего взаимодействие атома и резонансного поля.

Воспользуемся хорошо известной в теоретической спектроскопии полуклассической моделью взаимодействия атома с электромагнитным полем для исследования воздействия ГВ на атом космического мазера.

Сначала рассмотрим двухуровневый атом в монохроматическом квазирезонансном сильном поле с частотой Ω , близкой к атомной частоте ω . “Сильное” поле означает, что доми-

нируют вынужденные переходы. Такая система, с одной стороны, описывается в терминах матрицы плотности, а с другой – поле представляется классической волной. В результате получаем систему уравнений Блоха для компонент матрицы плотности в виде

$$\begin{aligned} \frac{d}{dt}\rho_{22} &= -\gamma\rho_{22} + 2i\alpha_1 \cos(\Omega t - k_1 y)(\rho_{21} - \rho_{12}) \\ \left[\frac{\partial}{\partial t} + v \frac{\partial}{\partial y} \right] \rho_{12} &= -(\gamma_{12} + i\omega)\rho_{12} - 2i\alpha_1 \cos(\Omega t - k_1 y)(\rho_{22} - \rho_{11}) \\ \rho_{22} + \rho_{11} &= 1 \end{aligned} \quad (71)$$

Здесь ρ_{22} , ρ_{11} – населенности уровней; ρ_{12} , ρ_{21} – поляризационные члены; γ , γ_{12} – продольная и поперечная постоянные распада (считая нижний уровень основным, получаем $\gamma_{12} = \gamma/2$); $\alpha_1 = \frac{\mu E}{\hbar}$ – параметр (частота) Раби, пропорциональный интенсивности электромагнитной волны (ЭМВ); μ – дипольный момент; E – электрическая напряженность; $\hbar = 1.05 \cdot 10^{-27}$ эрг·с – постоянная Планка; k_1 – волновой вектор ЭМВ; v – скорость атома вдоль оси Oy ; $\gamma \ll \alpha_1$ – условие сильного поля.

Переходя к исследованию астрофизической системы, заметим, что в космическом пространстве известны источники монохроматического электромагнитного излучения. Это – космические мазеры, атомы которых находятся в основных состояниях, а переходы осуществляются с метастабильных уровней, т.е. двухуровневая модель атома в данном случае вполне пригодна. Насыщенные космические мазеры полностью реализуют условия сильного поля. С другой стороны, можно предположить, что имеется причина, по которой расстояние, а, значит, и скорость заданного атома мазера в направлении на детектор сигнала на Земле будет периодически меняться. Это – периодическое воздействие ГВ, испускаемое источником, расположенным, как показано на Рис.6. Воздействие ГВ сказывается на расположении

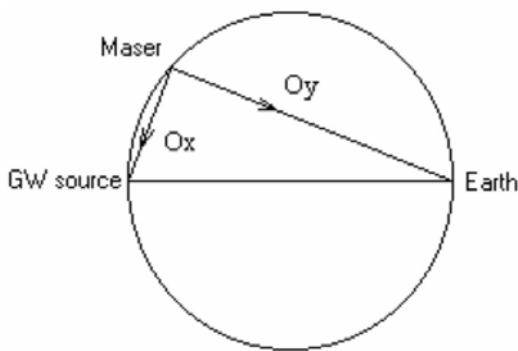


Рис. 6

уровней атома, на ЭМВ мазера и на положении атома. Можно показать, что первый эффект пренебрежимо мал по сравнению с двумя остальными. Влияние ГВ на монохроматическую ЭМВ учитывается в результате решения уравнения для эйконала

$$g^{ik} \frac{\partial \psi}{\partial y^i} \frac{\partial \psi}{\partial y^k} = 0 \quad (72)$$

Закон движения атома должен быть получен из уравнения геодезической

$$\frac{d^2 x^i}{ds^2} + \Gamma_{kl}^i \frac{dx^k}{ds} \frac{dx^l}{ds} = 0, \quad (73)$$

(а не из уравнения геодезического отклонения, как в соответствующих расчетах относительного смещения частей лабораторных установок). Уравнения (refeq71-refeq71) описывают поведение двухуровневого атома насыщенного космического мазера в поле ГВ. Решая их и налагая требования параметрического резонанса, можно рассчитать соответствующий сигнал. При этом обсуждаемый эффект имеет нулевой порядок, и оказывается, что его детектирование на Земле возможно с помощью уже существующих радиотелескопов, позволяющих наблюдать сигнал космического мазера.

Изотропное возмущение метрики Минковского

Будем считать геометрию Римановой и пользоваться обычными уравнениями Эйнштейна в приближении слабого поля вдали от масс $g^{ik} = g^{(0)ik} + h^{ik}$. Поправки к метрическому тензору плоского пространства-времени удовлетворяют волновому уравнению. В простейшем случае плоских волн оно имеет вид

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) h_i^k = 0 \quad (74)$$

В качестве решений естественно рассмотреть выражения

$$h_i^k = \text{Re}[A_i^k \exp(ik_\alpha x^\alpha)] \quad (75)$$

которые удовлетворяют уравнению, если $k_\alpha k^\alpha = 0$, т.е. k^α - светоподобный вектор. Поэтому сам метрический тензор можно записать

$$g^{ik} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 + h \cos \frac{D}{c}(x^0 - x^1) & 0 \\ 0 & 0 & 0 & -1 - h \cos \frac{D}{c}(x^0 - x^1) \end{pmatrix} \quad (76)$$

где h – безразмерная амплитуда ГВ ($h \ll 1$), D – частота ГВ.

Решение уравнения (48) с учетом (51) показывает, что воздействие ГВ приводит к фазовой модуляции ЭМВ. В силу малости h последняя может быть представлена в виде суперпозиции

$$E(t) = E \cos(\Omega t - ky) + E \frac{\omega}{4D} h [\cos((\Omega - D)t - ky) - \cos((\Omega + D)t - ky)] \quad (77)$$

Решение уравнения (49) с учетом (51) дает

$$y(t) \sim h \frac{c}{D} \sin(Dt + k_g x) \quad (78)$$

где k_g – волновой вектор ГВ. Зависимость (78) позволяет записать выражение для компоненты скорости атома, направленной в сторону Земли

$$\begin{aligned} v &= v_0 + v_1 \cos Dt \\ v_1 &= hc \end{aligned} \quad (79)$$

Подставляя (53) и (52) в (47), получим

$$\begin{aligned} \frac{d}{dt} \rho_{22} &= -\gamma \rho_{22} + 2i[\alpha_1 \cos(\Omega t - ky) + \alpha_2 \cos((\Omega - D)t - ky) - \alpha_2 \cos((\Omega + D)t - ky)](\rho_{21} - \rho_{12}) \\ \frac{d}{dt} \rho_{12} &= -(\gamma_{12} + i\omega) \rho_{12} - 2i[\alpha_1 \cos(\Omega t - ky) + \alpha_2 \cos((\Omega - D)t - ky) - \alpha_2 \cos((\Omega + D)t - ky)](\rho_{22} - \rho_{11}) \\ \rho_{22} + \rho_{11} &= 1 \end{aligned} \quad (80)$$

где $\alpha_2 = \frac{\omega h}{4D} \alpha_1$, а соотношение (53) было учтено в выражении для полной производной $\frac{d}{dt} = \frac{\partial}{\partial t} + kv$. Решение системы уравнений (54) проводится методом асимптотического разложения. При этом в случае выполнения определенных условий на параметры задачи можно говорить об оптико-метрическом параметрическом резонансе – ОМПР. Эти условия имеют вид:

- ЭМВ является спектроскопически сильной

$$\frac{\gamma}{\alpha_1} = \Gamma \varepsilon; \Gamma = O(33); \varepsilon \ll 1 \quad (81)$$

- Амплитудное условие ОМПР, связанное с трихроматическим полем

$$\frac{\alpha_2}{\alpha_1} = \frac{\omega h}{4D} = a \varepsilon; a = O(33); \varepsilon \ll 1 \quad (82)$$

- Амплитудное условие ОМПР, связанное с периодическим изменением скорости атома

$$\frac{kv_1}{\alpha_1} = \frac{\omega h}{\alpha_1} = \kappa \varepsilon; \kappa = O(33); \varepsilon \ll 1 \quad (83)$$

- Частотное условие ОМПР

$$(\omega - \Omega + kv_0)^2 + 4\alpha_1^2 = D^2 + O(\varepsilon) \Rightarrow D \sim 2\alpha_1 \quad (84)$$

Если условия (22.1-22.4) выполнены, то в результате решения (54) методом асимптотического разложения по малому параметру ε получим главный член разложения для $\text{Im}(\rho_{21})$ – величины, характеризующей поток энергии рассеянного излучения. На частоте, смещенной на D от центрального пика, он пропорционален ε^0 , и имеет вид

$$\text{Im}(\rho_{21}) \sim \frac{\alpha_1}{D} \cos 2Dt + O(\varepsilon) \quad (85)$$

Отрицательные значения соответствуют усилению, положительные – ослаблению потока энергии на указанной частоте, происходящему из-за перераспределения энергии излучения лазера в условиях ОМПР.

Может, однако, случиться так, что сигнал в предлагаемом эксперименте не будет обнаружен, или будет отличаться от приведенного расчета, или будет различаться для разных точек наблюдения. Это будет указывать на неучтенные, но существенные факторы. В том числе и на геометрические свойства пространства-времени.

Анизотропное возмущение метрики Минковского

Известные рассуждения и выкладки, приводящие к уравнению (50), можно повторить, считая, что выражение для метрики можно записать следующим образом

$$g_{ij}(x) \rightarrow g_{ij}(x, \dot{x}) = \eta_{ij}(x) + h_{ij}(x, \dot{x}) \quad (86)$$

где

$$\eta_{ij}(x) = \eta_{ij}^{(0)}(x) \quad (87)$$

- метрика Минковского для плоского пространства, $h_{ij}(x, \dot{x})$ - малая поправка, такая, что $h_i^k(x, \dot{x}) = \eta^{(0)kj} h_{ij}(x, \dot{x})$. Структура уравнений Эйнштейна при этом принципиально не изменится, и поправка $h_i^k(x, \dot{x})$ будет по-прежнему удовлетворять волновому уравнению типа (50). Но выражение (75) будет выглядеть, например, так

$$h_i^k(x, \dot{x}) = \text{Re}[A_i^k(\dot{x}) \exp(ik_\alpha x^\alpha)] \quad (88)$$

Это означает, что амплитуда ГВ будет различаться в различных направлениях их распространения. С точки зрения наблюдений, основанных на методе ОМПР, эту разницу невозможно заметить непосредственно, поскольку используемый эффект – нулевого порядка. Но она будет проявлять себя косвенно, например, выполнением условий ОМПР на различных расстояниях в различных направлениях от одного источника ГВ. Так, условия (82) и (83) теперь станут такими

- Амплитудное условие ОМПР, связанное с трихроматическим полем

$$\frac{\alpha_2 \zeta_1(\dot{x})}{\alpha_1} = \zeta_1(\dot{x}) \frac{\omega h}{4D} = a \zeta_1(\dot{x}) \varepsilon; a = O(33); \varepsilon \ll 1 \quad (89)$$

- Амплитудное условие ОМПР, связанное с периодическим изменением скорости атома

$$\frac{kv_1 \zeta_2(\dot{x})}{\alpha_1} = \zeta_2(\dot{x}) \frac{\omega h}{\alpha_1} = \kappa \zeta_2(\dot{x}) \varepsilon; \kappa = O(33); \varepsilon \ll 1 \quad (90)$$

где функции $\zeta_1(\dot{x}), \zeta_2(\dot{x})$ связаны с выражениями для амплитуд ГВ $A_i^k(\dot{x})$.

Этот эксперимент можно использовать для определения геометрических свойств пространства-времени следующим образом. Теоретические выражения, которые следует сравнивать с результатами эксперимента, естественно, следует получать при различных предположениях о геометрии пространства-времени. То из предположений, которое приведет к наиболее близкому соответствию с экспериментом, и будет соответствовать геометрическим свойствам пространства-времени.

Обобщенные n -арные законы композиции в алгебре и их связь с ассоциированными метрическими формами

В.М. Чернов

Институт систем обработки изображений РАН, Самара, Россия.

E-mail: vche@smr.ru

В работе рассматривается задача полилинеаризации норм алгебры $\mathbf{H}_4$. Вводятся новые бинарные, а также тернарные и кватернарные операции в алгебре $\mathbf{H}_4$ с изотропным базисом ("умножение Цассенхауза"). Показывается что квадратичная норма Минковского элемента алгебры, норма Бервальда-Моора, ассоциированная с формой четвертой степени, а также рассмотренная в предыдущей работе автора норма, ассоциированная с кубической формой, совпадают со значениями введенной бинарной, кватернарной и тернарной операций при равных значениях элементов – "сомножителей Цассенхауза".

1 Введение

1.1 Классические композиционные алгебры и алгебры $\mathbf{H}_n$

Наиболее широкое распространение в физике, механике, информатике и других приложениях получили *композиционные алгебры*, то есть алгебры без делителей нуля с единицей, на векторных пространствах которых определены невырожденные квадратичные формы $N(x)$ (*нормы*) с условием $N(xy) = N(x)N(y)$. Рекурсивный метод их построения и классификация над различными полями связана с существованием в алгебрах, получаемых на каждом шаге рекурсии (анти)автоморфизма $x \mapsto \bar{x}$ второго порядка, продолжаемого рекурсивно для последующих шагов построения и порождающих указанные выше формы $N(x)$. За желание иметь алгебры, отличные от $\mathbf{R}$ и $\mathbf{C}$, но с аналогами вещественной или комплексной нормы приходится расплачиваться некоммутативностью и/или неассоциативностью таких алгебр. Кроме того, рекурсивный процесс Кэли-Диксона построения композиционных алгебр уже на третьем шаге приводит к неассоциативности и не может быть продолжен дальше [1]-[3]. Помимо собственно композиционных алгебр, процесс Кэли-Диксона приводит к явному описанию алгебр ряда алгебр и с делителями нуля (например, алгебры двойных чисел, изоморфной $\mathbf{R} \oplus \mathbf{R}$, алгебры (2×2) -матриц $\mathbf{M}_2(\mathbf{R})$), но векторных пространствах которых определена также квадратичная мультипликативная форма $N(x)$, уже не являющаяся невырожденной [1].

Некоммутативная четырехмерная алгебра кватернионов, например, успешно используется при решении задач механики, машинного зрения, в физике. Это связано как и с наличием нормы, так и, например, с элегантным представлением ортогональных преобразований трехмерного пространства не на "внешнем" матричном языке, а в терминах внутренних операций алгебры кватернионов, то есть, "бескоординатно".

Далее, например, представление элемента $\mathbf{X}$ четырехмерной алгебры, изоморфной алгебре (2×2) -матриц $\mathbf{M}_2(\mathbf{R})$ в "клиффордовом" базисе с правилом умножения базисных элемен-

тов в форме

$$\begin{aligned} \mathbf{e}_0^2 &= \mathbf{e}_0, \mathbf{e}_1^2 = \mathbf{e}_0, \mathbf{e}_2^2 = \mathbf{e}_0, \mathbf{e}_3^2 = -\mathbf{e}_0; \\ \mathbf{e}_1\mathbf{e}_2 &= -\mathbf{e}_2\mathbf{e}_1, \mathbf{e}_1\mathbf{e}_3 = -\mathbf{e}_3\mathbf{e}_1, \mathbf{e}_2\mathbf{e}_3 = -\mathbf{e}_3\mathbf{e}_2; \mathbf{e}_1\mathbf{e}_2 = \mathbf{e}_3, \end{aligned}$$

и инволютивным отображением (т.н. *симплектическая инволюция*)

$$\mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \bar{\mathbf{X}} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

естественным вложением $\mathbf{R} \rightarrow \mathbf{M}_2(\mathbf{R})$ с $y \mapsto y\mathbf{e}_0$, $y \in \mathbf{R}$ позволяет записать определяющее квадратное уравнение для элемента $\mathbf{X}$ с коэффициентами, выраженными в бескоординатной форме в терминах нормы $N(\mathbf{X}) = \mathbf{X} \cdot \bar{\mathbf{X}} = \det \mathbf{X}$ и следа $Tr(\mathbf{X}) = \mathbf{X} + \bar{\mathbf{X}} = a + d$:

$$\mathbf{X}^2 - Tr(\mathbf{X})\mathbf{e}_0\mathbf{X} + N(\mathbf{X})\mathbf{e}_0 = 0, \quad (1)$$

то есть, фактически, в форме частного (двумерного) случая теоремы Гамильтона-Кэли. Отсюда при надлежащей интерпретации следует, что все результаты "линейной" геометрии двумерной плоскости могут быть получены как следствия алгебраических свойств четырехмерной алгебры $\mathbf{M}_2(\mathbf{R})$.

В отличие от композиционных алгебр, произвольные ассоциативно-коммутативные конечномерные алгебры уже не являются квадратичными алгебрами над полем $\mathbf{R}$. Поэтому алгебраические уравнения для их элементов естественным образом могут быть ассоциированы с автоморфизмами более высоких порядков, что дает основание предполагать возможность анализа свойств геометрических интерпретаций этих алгебр, выражаемых в терминах группы "симметрий" более высокого порядка, чем второй.

Первым шагом к реализации отмеченного выше комплекса идей и пониманию роли автоморфизмов высокого порядка при создании геометро-физических моделей пространства-времени является, по мнению автора, определение инвариантных характеристик уравнений, которым удовлетворяют элементы ассоциативно-коммутативных алгебр. То есть, определение аналогов форм $N(\mathbf{X}) = \mathbf{X} \cdot \bar{\mathbf{X}} = \det \mathbf{X}$ и $Tr(\mathbf{X}) = \mathbf{X} + \bar{\mathbf{X}} = a + d$ соотношения (1). Таким первым шагом в указанном направлении явились работы автора [3],[4], в которых, в частности был получен следующий результат (теорема 2.1).

Рассмотрим алгебру $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \cong \mathbf{H}_4$ с базисом $\{E_1, E_2, E_3, E_4\}$ и с правилом умножения базисных элементов, задаваемым таблицей 1.1. (Такой базис мы далее будем называть *изотропным* базисом).

Таблица 1.1.

$\times$	E_1	E_2	E_3	E_4
E_1	E_4	0	0	0
E_2	0	E_2	0	0
E_3	0	0	E_3	0
E_4	0	0	0	E_4

Мультипликативно нейтральным элементом (единицей алгебры) в этом базисе является элемент $\mathbf{I} = E_1 + E_2 + E_3 + E_4$, а поле $\mathbf{R}$ канонически вкладывается в алгебру $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$:

$$\mathbf{R} \rightarrow \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \cong \mathbf{H}_4, \quad x \mapsto x\mathbf{I}, \quad x \in \mathbf{R}.$$

Теорема. Алгебра $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ является алгеброй четвертой степени над $\mathbf{R}$, то есть, любой элемент $\mathbf{w} \in \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ удовлетворяет алгебраическому уравнению степени не выше четвертой с вещественными коэффициентами.

Действительно, пусть $\mathbf{w} = aE_1 + bE_2 + cE_3 + dE_4 \leftrightarrow (a, b, c, d)$. Рассмотрим четыре отображения алгебры $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ в себя, являющихся, очевидно, автоморфизмами:

$$\begin{aligned}\tau_0 : \mathbf{w} &= aE_1 + bE_2 + cE_3 + dE_4 \mapsto aE_1 + bE_2 + cE_3 + dE_4, \\ \tau_1 : \mathbf{w} &= aE_1 + bE_2 + cE_3 + dE_4 \mapsto bE_1 + cE_2 + dE_3 + aE_4, \\ \tau_2 : \mathbf{w} &= aE_1 + bE_2 + cE_3 + dE_4 \mapsto cE_1 + dE_2 + aE_3 + bE_4, \\ \tau_3 : \mathbf{w} &= aE_1 + bE_2 + cE_3 + dE_4 \mapsto dE_1 + aE_2 + bE_3 + cE_4.\end{aligned}\tag{2}$$

Другими словами, отображения переставляют циклическим образом компоненты (a, b, c, d) элемента $\mathbf{w}$ алгебры. Нетрудно показать, что элемент $\mathbf{w}$ является корнем многочлена

$$\Phi(\xi; \mathbf{w}) = (\xi - \tau_0(\mathbf{w}))(\xi - \tau_1(\mathbf{w}))(\xi - \tau_2(\mathbf{w}))(\xi - \tau_3(\mathbf{w})),\tag{3}$$

а также то, что коэффициенты полинома $\Phi(\xi; \mathbf{w})$ вещественные. Действительно, непосредственным вычислением получаем:

$$\Phi(\xi; \mathbf{w}) = \xi^4 - s_1(\mathbf{w})\xi^3 + s_2(\mathbf{w})\xi^2 - s_3(\mathbf{w})\xi + s_4(\mathbf{w}),\tag{4}$$

где вещественные коэффициенты $s_\nu(\mathbf{w})$ являются однородными симметричными формами компонент

(a, b, c, d) элемента $\mathbf{w}$ алгебры:

$$\begin{aligned}s_1(\mathbf{w}) &= s_{14}(\mathbf{w}) = a + b + c + d, \\ s_2(\mathbf{w}) &= s_{24}(\mathbf{w}) = ab + ac + ad + bc + bd + cd, \\ s_3(\mathbf{w}) &= s_{34}(\mathbf{w}) = bcd + acd + abd + abc, \\ s_4(\mathbf{w}) &= s_{44}(\mathbf{w}) = abcd.\end{aligned}\tag{5}$$

Многочлен $\Phi(\xi; \mathbf{w})$ минимальной степени с вещественными коэффициентами и с коэффициентом при старшей степени ξ , равным единице, такой, что $\Phi(\xi; \mathbf{w})|_{\xi=\mathbf{w}} = 0$ будем называть *определяющим многочленом элемента $\mathbf{w}$* , а его коэффициенты – *определяющими формами*.

Замечание 1.1. Формы (5) инвариантны относительно любой перестановки $\sigma \in \mathbf{S}_4$ четырех компонент (a, b, c, d) элемента $\mathbf{w}$ алгебры. Следовательно, так как $\sigma \in \mathbf{S}_4$ является автоморфизмом алгебры $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ над полем $\mathbf{R}$, то многочлен четвертой степени $\Phi(\xi; \mathbf{w})$ наряду с корнем $\mathbf{w}$ имеет своими корнями еще, по крайней мере, 2^3 корня $\sigma(\mathbf{w})$, $\sigma \in \mathbf{S}_4$, то есть, все его автоморфные образы относительно автоморфизмов $\sigma \in \mathbf{S}_4$.

Замечание 1.2. В базисе $\{E, I, J, K\}$ алгебры $\mathbf{H}_4 \cong \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ с законом умножения базисных единиц, задаваемых следующей таблицей Кэли.

элемент $\omega \in \mathbf{H}_4$ представляется в форме $\omega = tE + xI + yJ + zK$ ($t, x, y, z \in \mathbf{R}$), E - мультипликативно нейтральный элемент (единица) алгебры. Связь между изотропным базисом $\{E_1, E_2, E_3, E_4\}$ и базисом $\{E, I, J, K\}$ осуществляется посредством линейного преобразования с ортогональной матрицей Адамара

$$had_4 = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Таблица 1.2.

$\times$	E	I	J	K
E	E	I	J	K
I	I	E	K	J
J	J	K	E	I
K	K	J	I	E

Совершенно очевидно, что каждому автоморфизму, определяемому действием элемента группы подстановок $\mathbf{S}_4$, то есть, переставляющему компоненты элемента $\omega \in \mathbf{H}_4$, представленного в базисе $\{E_1, E_2, E_3, E_4\}$, соответствует линейное преобразование компонент t, x, y, z элемента $\omega = tE + xI + yJ + zK$, и также реализующее некоторый автоморфизм алгебры $\mathbf{H}_4$. В частности, такими автоморфизмами являются:

$$\begin{aligned}
\mu_0 : \quad \omega &\mapsto \mu_0(\omega) = tE + xI + yJ + zK, \\
\mu_1 : \quad \omega &\mapsto \mu_1(\omega) = tE + xI - yJ - zK, \\
\mu_2 : \quad \omega &\mapsto \mu_2(\omega) = tE - xI + yJ - zK, \\
\mu_3 : \quad \omega &\mapsto \mu_3(\omega) = tE - xI - yJ + zK.
\end{aligned} \tag{6}$$

При представлении элемента $\omega \in \mathbf{H}_4$ в базисе $\{E_1, E_2, E_3, E_4\}$ этой четверке преобразований соответствует действие на компоненты элемента некоторой подгруппы четвертого порядка (изоморфной прямому произведению $\mathbf{C}_2 \times \mathbf{C}_2$ двух циклических групп второго порядка) группы $\mathbf{S}_4$.

Записывая в этом случае определяющий многочлен, непосредственным вычислением получаем:

$$\begin{aligned}
\Phi(\xi; \omega) &= (\xi - \mu_0(\omega))(\xi - \mu_1(\omega))(\xi - \mu_2(\omega))(\xi - \mu_3(\omega)) = \\
&= (\xi^4 - S_1(\omega)\xi^3 + S_2(\omega)\xi^2 - S_3(\omega)\xi + S_4(\omega))E,
\end{aligned}$$

где

$$\begin{aligned}
S_{14}(\omega) &= 4t, \\
S_{24}(\omega) &= 6t^2 - 2x^2 - 2y^2 - 2z^2, \\
S_{34}(\omega) &= -4tx^2 - 4y^2t - 4z^2t + 4t^3 + 8yzx = 4t(t^2 - x^2 - y^2 - z^2) + 8yzx, \\
S_{44}(\omega) &= x^4 + y^4 + z^4 - 2t^2x^2 - 2t^2y^2 - 2t^2z^2 + t^4 + 8txyz - 2x^2y^2 - 2x^2z^2 - 2y^2z^2,
\end{aligned} \tag{7}$$

и $S_{24}(\omega)$ с точностью до нормирующих множителей совпадает с (псевдо)метрической формой Минковского, а $S_{44}(\omega)$ имеет "стандартный бервальд-мооровский" вид.

1.2 Основные идеи и определения

Как ни парадоксально, но идея использования дополнительных (или переопределенных) операций на той или иной конечномерной алгебре, ассоциированных с автоморфизмами порядка выше второго, с целью алгебраической поддержки решения геометрических задач имеет почтенную историю. Существует достаточно экзотическая и малоизвестная алгебраическая структура, известная как "конечные почти-поля Цассенхауза"[6] (см. также [7], Глава 20). Именно, в поле $\mathbf{F}_q$, $q = p^m$, (p - простое) вводится операция $x * y$ ($x, y \in \mathbf{F}_q$), выражающаяся через операцию умножения в поле $\mathbf{F}_q$ по закону $x * y = y \cdot \eta(x)$, где η - автоморфизм

Фробениуса специального вида. Эта операция некоммутативна и неассоциативна (последнее – в силу теоремы Веддербёрна [7],[8]). Естественно, что конечность полей $\mathbf{F}_q$, следовательно, и почти-полей Цассенхауза, ограничивает круг задач, решаемых с применением такой техники, исключительно конфигурационными задачами конечной геометрии [7]. Идея рассмотрения "полилинейных в отличие от классических "билинейных скалярных произведений в ассоциативно-коммутативных алгебрах с целью создания адекватных геометро-физических моделей принадлежит, по всей видимости, Д.Г.Павлову [9], [10].

Рассмотрим пример алгебры $\mathbf{R} \oplus \mathbf{R}$ с базисом $\{E_1, E_2\}$ и с правилом умножения базисных элементов, задаваемым таблицей 1.3 и мультипликативно нейтральным элементом (единицей) $\mathbf{I} = E_1 + E_2 = (1, 1)$.

Таблица 1.3.

$\times$	E_1	E_2
E_1	E_4	0
E_2	0	E_2

Типичный элемент $\mathbf{x} = x_1 E_1 + x_2 E_2$ будем далее обозначать для краткости $\mathbf{x} = (x_1, x_2)$.

Пример 1.1. Рассмотрим две подстановки

$$\sigma^1 \leftrightarrow \sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad \sigma^2 \leftrightarrow \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}. \quad (8)$$

Определим действие операторов σ^1, σ^2 , ассоциированных с подстановками σ_1, σ_2 на элементы $\mathbf{x} = (x_1, x_2)$ алгебры

$\mathbf{R} \oplus \mathbf{R}$:

$$\sigma^1(x) = (x_1, x_2), \quad \sigma^2(x) = (x_2, x_1),$$

то есть, операторы σ^1, σ^2 переставляют компоненты элемента $\mathbf{x} = (x_1, x_2)$ в соответствии с нижними строками подстановок σ_1, σ_2 .

Введем на $\mathbf{R} \oplus \mathbf{R}$ новую бинарную операцию $[\mathbf{x}, \mathbf{y}]$ ("умножение Цассенхауза"):

$$[\mathbf{x}, \mathbf{y}] = \sigma^1(\mathbf{x}) \cdot \sigma^2(\mathbf{y}),$$

где символом $(\cdot)$ обозначено обычное "покомпонентное" умножение элементов из $\mathbf{R} \oplus \mathbf{R}$.

Непосредственно проверяются следующие соотношения.

$$[\mathbf{x}, \mathbf{y}] = (x_1, x_2) \cdot (y_2, y_1) = (x_1 y_2, x_2 y_1) \doteq (\xi_2, \xi_1), \quad (9)$$

$$[\mathbf{x}, \mathbf{x}] = (x_1, x_2) \cdot (x_2, x_1) = (x_1 x_2, x_2 x_1) = x_1 x_2 \mathbf{I} \doteq N(\mathbf{x}) \mathbf{I}. \quad (10)$$

Заметим, что введенная в (10) функция $N(\mathbf{x})$ совпадает с традиционной нормой элемента алгебры $\mathbf{R} \oplus \mathbf{R}$, выраженной в терминах изотропных координат (компонент).

Далее, справедливы равенства:

$$\begin{aligned} N([\mathbf{x}, \mathbf{y}]) &= [[\mathbf{x}, \mathbf{y}], [\mathbf{x}, \mathbf{y}]] = \sigma^1([\mathbf{x}, \mathbf{y}]) \cdot \sigma^2([\mathbf{x}, \mathbf{y}]) = \\ &= (x_1 y_2, x_2 y_1) \cdot (x_2 y_1, x_1 y_2) = (x_1 y_2 x_2 y_1, x_1 y_2 x_2 y_1) = \\ &= (x_1 y_2, x_1 y_2) \cdot (x_2 y_1, x_2 y_1) = (x_1 y_2) \mathbf{I} \cdot (x_2 y_1) \mathbf{I} = \xi_1 \xi_2 \mathbf{I} = \\ &= (x_1 x_2) \mathbf{I} \cdot (y_1 y_2) \mathbf{I} = N(\mathbf{x}) N(\mathbf{y}). \end{aligned} \quad (11)$$

Пример 1.2. Остановимся еще на одной иллюстрации "цассенхаузового умножения". Рассмотрим множество двумерно индексированных $(n \times n)$ массивов с покомпонентным сложением. Введем операцию $*$ "умножения отличную от обычного матричного умножения. Пусть $A = \{a_{ij}\}$, $B = \{b_{ij}\}$, $D = \{d_{ij}\} = A \odot B$. Тогда, по определению:

$$d_{ij} = \sum_{k=1}^n a_{ik} b_{jk},$$

или, в неформальных терминах, " (i, j) -й элемент $*$ -произведения" равен скалярному произведению i -ой строки на j -ую строку". Преобразование массивов $\tau : A \rightarrow \tau(A)$ по правилу $\tau : a_{ij} \mapsto a_{ji}$, есть "обычное" транспонирование матриц, но которое является автоморфизмом второго порядка по отношению к операции $*$: $\tau(A * B) = \tau(A) * \tau(B)$, тогда как для обычного матричного умножения транспонирование является антиавтоморфизмом: $(AB)^t = B^t A^t$. В терминах операции $*$ традиционное матричное умножение является просто "умножением Цассенхауза" относительно операции $*$: $AB = A * \tau(B)$. Рассмотрение, например, трехмерных массивов $A = \{a_{ijk}\}$, $B = \{b_{ijk}\}$, $C = \{c_{ijk}\}$ и автоморфизм третьего порядка $\tau : a_{ijk} \mapsto a_{jki}$ приводит к "цассенхаузовому умножению" трехмерных массивов $X = [A, B, C]$, где

$$x_{pqr} = \sum_{i,j=1}^n a_{pij} b_{iqj} c_{ijr}$$

и так далее.

Таким образом, из соотношения (11) следуют:

- равенство $N([\mathbf{x}, \mathbf{y}]) = N(\mathbf{x}) N(\mathbf{y})$, то есть "соотношение мультипликативности" для нормы $N(\mathbf{x})$ относительно введенной операции $[\mathbf{x}, \mathbf{y}]$, причем эта норма совпадает с традиционной нормой элемента алгебры $\mathbf{R} \oplus \mathbf{R}$, выраженной в терминах изотропных координат (компонент);
- равенство

$$\sigma^1([\mathbf{x}, \mathbf{y}]) \cdot \sigma^2([\mathbf{x}, \mathbf{y}]) = \xi_1 \xi_2 \mathbf{I}, \quad (12)$$

где (ξ_1, ξ_2) есть компоненты "цассенхаузовского произведения" $[\mathbf{x}, \mathbf{y}]$, причем упорядоченная пара верхних индексов операторов σ^1 , σ^2 совпадает с упорядоченной парой нижних индексов компонент (ξ_1, ξ_2) .

Последнее соображение является основой для обобщения понятия закона композиции для алгебр с n -арными операциями, а первое (мультипликативность нормы) является мотивацией для выбора именно такого принципа обобщения.

Введем некоторые формальные определения и понятия.

Пусть $\sigma_* \in \mathbf{S}_n$ - некоторая подстановка:

$$\sigma_* = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma_*(1) & \sigma_*(2) & \dots & \sigma_*(n) \end{pmatrix}.$$

Ассоциированным с подстановкой $\sigma_* \in \mathbf{S}_n$ оператором $\sigma^* : \mathbf{H}_n \rightarrow \mathbf{H}_n$ будем называть оператор, переставляющий изотропные координаты элемента из алгебры $\mathbf{H}_n$:

$$\sigma^* : \mathbf{x} = (x_1, x_2, \dots, x_n) \mapsto \sigma^*(\mathbf{x}) = (x_{\sigma_*(1)}, x_{\sigma_*(2)}, \dots, x_{\sigma_*(n)}).$$

Пусть $(\sigma^1, \sigma^2, \dots, \sigma^m)$ есть упорядоченное семейство ассоциированных операторов $(m \leq n)$. Определим действие m -семейства $(\sigma^1, \sigma^2, \dots, \sigma^m)$ на m -семейство элементов $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \subset \mathbf{H}_n$ "тензорным образом":

$$(\sigma^1 \otimes \sigma^2 \otimes \dots \otimes \sigma^m)(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) = \sigma^1(\mathbf{x}_1) \cdot \sigma^2(\mathbf{x}_2) \cdot \dots \cdot \sigma^m(\mathbf{x}_m).$$

Пусть $A \subset \mathbf{Z}^m$ - некоторое множество индексов, $\tilde{\mathbf{S}}_n^m = \{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}); (a_1, \dots, a_m) \in A\}$ некоторое "отмеченное" индексным множеством $A \subset \mathbf{Z}^m$ множество m -семейств ассоциированных операторов.

Определение 1.1. m -арной операцией на алгебре $\mathbf{H}_n$ $(m \leq n)$, порожденной семейством

$\tilde{\mathbf{S}}_n^m = \{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}); (a_1, \dots, a_m) \in A\}$, будем называть операцию, определенную равенством (здесь и далее $\lambda_m \in \mathbf{R}$)

$$\begin{aligned} [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m] &= \\ &= \lambda_m \sum_{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}) \in \tilde{\mathbf{S}}_n^m} (\sigma^{a_1} \otimes \sigma^{a_2} \otimes \dots \otimes \sigma^{a_m})(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) = \\ &= \lambda_m \sum_{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}) \in \tilde{\mathbf{S}}_n^m} \sigma^{a_1}(\mathbf{x}_1) \cdot \sigma^{a_2}(\mathbf{x}_2) \cdot \dots \cdot \sigma^{a_m}(\mathbf{x}_m) \end{aligned}$$

Определение 1.2. m -арную операцию на алгебре $\mathbf{H}_n$ $(m \leq n)$, порожденную семейством $\tilde{\mathbf{S}}_n^m = \{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}); (a_1, \dots, a_m) \in A\}$, будем называть *нормируемой*, если

$$N(\mathbf{x}) = \lambda_m \sum_{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}) \in \tilde{\mathbf{S}}_n^m} (\sigma^{a_1} \otimes \sigma^{a_2} \otimes \dots \otimes \sigma^{a_m})(\mathbf{x}, \mathbf{x}, \dots, \mathbf{x}) \in \mathbf{R}. \quad (13)$$

Семейство ассоциированных операторов $\tilde{\mathbf{S}}_n^m$ в этом случае будем называть *нормирующим семейством*, а функцию $N(\mathbf{x}) \in \mathbf{R}$ будем называть $\tilde{\mathbf{S}}_n^m$ - (или просто, если из контекста ясно, какое семейство $\tilde{\mathbf{S}}_n^m$ имеется в виду).

Определение 1.3. Пусть $\mathbf{S}_n^m = \{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}); (a_1, \dots, a_m) \in A\}$ есть множество подстановок, ассоциированных с нормирующим семейством операторов

$\tilde{\mathbf{S}}_n^m = \{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}); (a_1, \dots, a_m) \in A\}$, $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ есть m -арная операция, порожденная семейством $\tilde{\mathbf{S}}_n^m$. Пусть в изотропных координатах элемент $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ имеет координаты $(\xi_1, \xi_2, \dots, \xi_m)$:

$$[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m] = (\xi_1, \xi_2, \dots, \xi_m).$$

Будем говорить, что семейство $\tilde{\mathbf{S}}_n^m$ порождает обобщенный m -арный закон композиции, если выполняется равенство

$$\begin{aligned} N([\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]) &= \underbrace{[[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m], \dots, [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]]}_{m \text{ раз}} = \\ &= \lambda_m \sum_{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}) \in \tilde{\mathbf{S}}_n^m} (\sigma^{a_1} \otimes \sigma^{a_2} \otimes \dots \otimes \sigma^{a_m}) \left(\underbrace{[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m], \dots, [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]}_{m \text{ раз}} \right) \cdot \\ &= \lambda_m \sum_{(\sigma^{a_1}, \sigma^{a_2}, \dots, \sigma^{a_m}) \in \tilde{\mathbf{S}}_n^m} \xi_{\sigma^{a_1}(1)} \xi_{\sigma^{a_2}(2)} \dots \xi_{\sigma^{a_m}(m)} \mathbf{I} \end{aligned}$$

Пример 1.3. В обозначениях Примера 1.1 рассмотрим одноэлементное множество $\mathbf{S}_2^2$ состоящее из одной пары подстановок $\{(\sigma_1, \sigma_2)\}$, определенных (8). Равенство (12) означает, что бинарная операция, порожденная семейством $\tilde{\mathbf{S}}_2^2$ ассоциированных операторов $\tilde{\mathbf{S}}_2^2 = \{(\sigma^1, \sigma^2)\}$, порождает бинарный закон композиции $\sigma^1([\mathbf{x}, \mathbf{y}]) \cdot \sigma^2([\mathbf{x}, \mathbf{y}]) = \xi_1 \xi_2 \mathbf{I}$, совпадающий с точностью до обозначений с мультипликативным законом композиции

$$N([\mathbf{x}, \mathbf{y}]) = [[\mathbf{x}, \mathbf{y}], [\mathbf{x}, \mathbf{y}]] = N(\mathbf{x}) N(\mathbf{y}) \text{ и } \lambda_2 = 1.$$

Замечание 1.4. Не всякое множество $\tilde{\mathbf{S}}_n^m$ является нормирующим множеством операторов. Нетрудно убедиться, что в обозначениях Примера 1.1 одноэлементное множество $\tilde{\mathbf{S}}_2^2 = \{(\sigma^1, \sigma^1)\}$ не является нормирующим множеством операторов, так как $N(\mathbf{x}) = [\mathbf{x}, \mathbf{x}] \notin \mathbf{R}$. В связи с этим возникает Вопрос 1: каковы необходимые и достаточные условия, сформулированные в теоретико-групповых терминах, того, чтобы множество операторов, ассоциированное с $\mathbf{S}_n^m = \{(\sigma_{a_1}, \sigma_{a_2}, \dots, \sigma_{a_m})\}$, было бы нормирующим?

Замечание 1.5. Естественно, что наибольший интерес вызывают те нормирующие множества $\tilde{\mathbf{S}}_4^m$ ассоциированных операторов множества $\mathbf{S}_4^m$, для которых $N(\mathbf{w})$ при всех $\mathbf{w} = (a, b, c, d) \in \mathbf{H}_4$ совпадала бы с одной из форм (5). В связи с этим возникает Вопрос 2: каковы необходимые и достаточные условия, сформулированные в теоретико-групповых терминах, того, что норма, порожденная множеством операторов $\tilde{\mathbf{S}}_n^m$ совпадала бы с одной из форм (5) для $\mathbf{H}_4$?

Настоящая работа посвящена получению достаточных условий в Вопросе 2. Другими словами, получению ответа на вопрос: какова должна быть бинарная, тернарная или кватернарная операция на алгебре $\mathbf{H}_4$, чтобы

$$N(\mathbf{x}) = \underbrace{[\mathbf{x}, \mathbf{x}, \dots, \mathbf{x}]}_{m \text{ раз}} \in \{s_{24}(\mathbf{w}), s_{34}(\mathbf{w}), s_{44}(\mathbf{w})\}$$

Замечание 1.6. Доказательства утверждений в следующих двух разделах работы сводятся к рутинной, хотя и громоздкой, проверке тождеств, аналогично Примеру 1.1.. Поэтому в работе приводятся только формулировки соответствующих теорем.

2 Обобщенные n -арные законы композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$

2.1 Обобщенный кватернарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$

Пусть кватернарная операция в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ определена равенством

$$[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4] = \lambda_4 (\sigma^1 \otimes \sigma^2 \otimes \sigma^3 \otimes \sigma^4) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \lambda_4 \sigma^1(\mathbf{x}_1) \cdot \sigma^2(\mathbf{x}_2) \cdot \sigma^3(\mathbf{x}_3) \cdot \sigma^4(\mathbf{x}_4) \quad (14)$$

где $\lambda_4 = 1$ и

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}. \quad (15)$$

Теорема 2.1. Одноэлементное семейство четверки операторов $\tilde{\mathbf{S}}_4^4 = \{(\sigma^1, \sigma^2, \sigma^3, \sigma^4)\}$, ассоциированное с одноэлементным множеством $\mathbf{S}_4^4$, состоящим из четверки (циклических) подстановок $\{(\sigma_1, \sigma_2, \sigma_3, \sigma_4)\}$ является нормирующим множеством, порождающим обобщенный

кватернарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ в (мультипликативной) форме

$$\begin{aligned} N_4([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]) &= [[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]] = \\ &= (\sigma^1 \otimes \sigma^2 \otimes \sigma^3 \otimes \sigma^4) ([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]) = \\ &= \xi_{\sigma_1(1)} \xi_{\sigma_2(2)} \xi_{\sigma_3(3)} \xi_{\sigma_4(4)} \mathbf{I} = \\ &= (x_{11}x_{12}x_{13}x_{14}) (x_{21}x_{22}x_{23}x_{24}) (x_{31}x_{32}x_{33}x_{34}) (x_{41}x_{42}x_{43}x_{44}) \mathbf{I} = \\ &= N_4(\mathbf{x}_1) \mathbf{I} \cdot N_4(\mathbf{x}_2) \mathbf{I} \cdot N_4(\mathbf{x}_3) \mathbf{I} \cdot N_4(\mathbf{x}_4) \mathbf{I}, \end{aligned} \quad (16)$$

где $(\xi_1, \xi_2, \xi_3, \xi_4)$ есть четверка координат элемента $[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]$ - результата определенной выше кватернарной операции в изотропном базисе: $(\xi_1, \xi_2, \xi_3, \xi_4) = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]$.

Кроме того, справедливо равенство

$$\begin{aligned} N_4(\mathbf{x}) &= [\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}] = (\sigma^1 \otimes \sigma^2 \otimes \sigma^3 \otimes \sigma^4) (\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) = \\ &= \xi_{\sigma_1(1)} \xi_{\sigma_2(2)} \xi_{\sigma_3(3)} \xi_{\sigma_4(4)} \mathbf{I} = (x_1x_2x_3x_4) \mathbf{I} = s_{44}(\mathbf{x}) \cdot \mathbf{I}, \end{aligned} \quad (17)$$

Замечание 2.1. Из Теоремы 2.1 следует, в частности, что введенное в определении 1.2. соотношением (1.14) понятие нормы действительно представляет собой мультипликативную (псевдо)норму (5) Бервальда-Моора $s_{44}(\mathbf{x})$, выраженную в изотропном базисе.

2.2 Обобщенный тернарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$

Пусть тернарная операция в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ определена равенством

$$[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3] = (\sigma^2 \otimes \sigma^3 \otimes \sigma^4 + \sigma^4 \otimes \sigma^1 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 \otimes \sigma^1 + \sigma^3 \otimes \sigma^1 \otimes \sigma^2) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \quad (18)$$

где $\lambda_3 = 1$ и

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}. \quad (19)$$

Теорема 2.2. Четырехэлементное семейство троек операторов

$$\tilde{\mathbf{S}}_4^3 = \{(\sigma^2, \sigma^3, \sigma^4), (\sigma^4, \sigma^1, \sigma^3), (\sigma^2, \sigma^4, \sigma^1), (\sigma^3, \sigma^1, \sigma^2)\},$$

ассоциированное с семейством $\mathbf{S}_4^4$ троек подстановок

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}, \quad (20)$$

является нормирующим множеством, порождающим обобщенный тернарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ в форме

$$\begin{aligned} N_3([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]) &= [[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3], [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]] = \\ &= (\sigma^2 \otimes \sigma^3 \otimes \sigma^4 + \sigma^4 \otimes \sigma^1 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 \otimes \sigma^1 + \sigma^3 \otimes \sigma^1 \otimes \sigma^2) ([\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3], \dots, [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]) = \\ &= (\xi_{\sigma_2(1)} \xi_{\sigma_3(2)} \xi_{\sigma_4(3)} + \xi_{\sigma_4(1)} \xi_{\sigma_1(2)} \xi_{\sigma_3(3)} + \xi_{\sigma_2(1)} \xi_{\sigma_4(2)} \xi_{\sigma_1(3)} + \xi_{\sigma_3(1)} \xi_{\sigma_1(2)} \xi_{\sigma_2(3)}) \mathbf{I}, \end{aligned} \quad (21)$$

где $(\xi_1, \xi_2, \xi_3, \xi_4)$ есть четверка координат элемента $[\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ - результата определенной выше тернарной операции в изотропном базисе:

$$(\xi_1, \xi_2, \xi_3, \xi_4) = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3],$$

и, кроме того, выполняются равенства

$$\begin{aligned} N_3(\mathbf{x}) &= [\mathbf{x}, \mathbf{x}, \mathbf{x}] = \\ &= (\sigma^2 \otimes \sigma^3 \otimes \sigma^4 + \sigma^4 \otimes \sigma^1 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 \otimes \sigma^1 + \sigma^3 \otimes \sigma^1 \otimes \sigma^2)(\mathbf{x}, \mathbf{x}, \mathbf{x}), \end{aligned} \quad (22)$$

$$N_3(\mathbf{x}) = s_{34}(\mathbf{x}) \mathbf{I}. \quad (23)$$

2.3 Обобщенный бинарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$

Пусть бинарная операция в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ определена равенством

$$[\mathbf{x}_1, \mathbf{x}_2] = \lambda_2 (\sigma^1 \otimes \sigma^2 + \sigma^1 \otimes \sigma^3 + \sigma^1 \otimes \sigma^4 + \sigma^2 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 + \sigma^3 \otimes \sigma^4)(\mathbf{x}_1, \mathbf{x}_2) \quad (24)$$

где $\lambda_2 = 1/2$ и

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \sigma_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}. \quad (25)$$

Теорема 2.3. Шестиэлементное семейство пар операторов

$$\tilde{\mathbf{S}}_4^2 = \{(\sigma^1, \sigma^2), (\sigma^1, \sigma^3), (\sigma^1, \sigma^4), (\sigma^2, \sigma^3), (\sigma^2, \sigma^4), (\sigma^3, \sigma^4)\},$$

ассоциированное с семейством $\mathbf{S}_4^2$ пар подстановок

$$\mathbf{S}_4^2 = \{(\sigma_1, \sigma_2), (\sigma_1, \sigma_3), (\sigma_1, \sigma_4), (\sigma_2, \sigma_3), (\sigma_2, \sigma_4), (\sigma_3, \sigma_4)\} \quad (26)$$

является нормирующим множеством, порождающим обобщенный бинарный закон композиции в алгебре $\mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R} \oplus \mathbf{R}$ в форме

$$\begin{aligned} N_2([\mathbf{x}_1, \mathbf{x}_2]) &= [[\mathbf{x}_1, \mathbf{x}_2], [\mathbf{x}_1, \mathbf{x}_2]] = \\ &= \frac{1}{2} (\sigma^1 \otimes \sigma^2 + \sigma^1 \otimes \sigma^3 + \sigma^1 \otimes \sigma^4 + \sigma^2 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 + \sigma^3 \otimes \sigma^4)([\mathbf{x}_1, \mathbf{x}_2], [\mathbf{x}_1, \mathbf{x}_2]) = \\ &= \frac{1}{2} (\xi_{\sigma_1(1)} \xi_{\sigma_2(2)} + \xi_{\sigma_1(1)} \xi_{\sigma_3(2)} + \xi_{\sigma_1(1)} \xi_{\sigma_4(2)} + \xi_{\sigma_2(1)} \xi_{\sigma_3(2)} + \xi_{\sigma_2(1)} \xi_{\sigma_4(2)} + \xi_{\sigma_3(1)} \xi_{\sigma_4(2)}) \mathbf{I}, \end{aligned} \quad (27)$$

где $(\xi_1, \xi_2, \xi_3, \xi_4)$ есть четверка координат элемента $[\mathbf{x}_1, \mathbf{x}_2]$ - результата определенной выше тернарной операции в изотропном базисе: $(\xi_1, \xi_2, \xi_3, \xi_4) = [\mathbf{x}_1, \mathbf{x}_2]$.

Кроме того, справедливы равенства

$$\begin{aligned} N_2(\mathbf{x}) &= [\mathbf{x}, \mathbf{x}] = \\ &= \frac{1}{2} (\sigma^1 \otimes \sigma^2 + \sigma^1 \otimes \sigma^3 + \sigma^1 \otimes \sigma^4 + \sigma^2 \otimes \sigma^3 + \sigma^2 \otimes \sigma^4 + \sigma^3 \otimes \sigma^4)(\mathbf{x}, \mathbf{x}), \end{aligned} \quad (28)$$

$$N_2(\mathbf{x}) = s_{24}(\mathbf{x}). \quad (29)$$

Замечание 2.3.. Из Теоремы 2.3 следует, в частности, что введенное в определении 1.2. соотношением (1.14) понятие нормы действительно представляет собой (псевдо)норму соответствующую метрике Минковского в форме $s_{44}(\mathbf{x})$, выраженную в изотропном базисе.

Замечание 2.4. В отличие от множества (17) подстановок в Теореме 2.2, которое группой не является, множество (21) подстановок в Теореме 2.3 является четырехэлементной нециклической группой, а в Теореме 2.1 множество подстановок есть циклическая группа.

3 Обобщения и открытые проблемы

1. В связи с Вопросами 1 и 2 п.1 возникает проблема классификации обобщенных законов композиции для пространства $\mathbf{H}_n$ произвольной размерности.

Проблема 1. Каковы необходимые и достаточные условия, сформулированные в теоретико-групповых (или комбинаторных) терминах, того, чтобы множество операторов, ассоциированное с семейством подстановок $\mathbf{S}_n^m = \{(\sigma_{a_1}, \sigma_{a_2}, \dots, \sigma_{a_m})\}$, было бы нормирующим; как связана порожденная норма с коэффициентами определяющего уравнения элемента алгебры $\mathbf{H}_n$?

Отметим, что групповая структура множества подстановок $\mathbf{S}_n^m$ является, скорее всего, достаточным условием того, чтобы ассоциированное множество операторов было бы нормирующим (Теоремы 2.1 и 2.3). Однако необходимость этого неочевидна, как показывает Теорема 2.2. Скорее всего, вопрос классификации множеств $\mathbf{S}_n^m$ в Проблеме 1 является комбинаторным, но не исключительно теоретико-групповым.

2. Исчерпывающая классификация ассоциативно-коммутативных алгебр без нильпотентных элементов содержится в теореме Вейерштрасса [2]: любая ассоциативно-коммутативная конечномерная алгебра без нильпотентных элементов над $\mathbf{R}$ изоморфна прямой сумме алгебр $\mathbf{R}$ и $\mathbf{C}$.

Из этой теоремы легко следует, что существует не более трех неизоморфных четырехмерных алгебр этого класса, а именно: $\mathbf{H}_4$, $\mathbf{H}_2 \oplus \mathbf{C}$, $\mathbf{C} \oplus \mathbf{C}$. В связи с этим возникает проблема экстраполяции теорем 2.1 – 2.3 на случай указанных четырехмерных алгебр и на случай произвольной ассоциативно коммутативной конечномерной алгебры $\mathbf{A}_n$.

Проблема 2. Каковы необходимые и достаточные условия того, чтобы множество операторов, ассоциированное с семейством автоморфизмов алгебры $\mathbf{A}_n$, было бы нормирующим; как связана порожденная норма с коэффициентами определяющего уравнения элемента алгебры $\mathbf{A}_n$?

3. Из соотношения (16) и Замечания 2.1 следует чисто алгебраический факт мультипликативности нормы Бервальда-Моора, выраженной в терминах изотропного базиса. Конечно, соответствующие (индуцированные) соотношения инвариантности формы $S_{44}(\mathbf{x})$ остаются справедливыми и в "физическом" базисе $\{E, I, J, K\}$ с Таблицей 2.1 умножения базисных элементов. Но соотношение мультипликативности для формы $s_{44}(\mathbf{x})$, а именно $s_{44}(\mathbf{x}_1 \cdot \mathbf{x}_2 \cdot \mathbf{x}_3 \cdot \mathbf{x}) = s_{44}(\mathbf{x}_1) s_{44}(\mathbf{x}_2) s_{44}(\mathbf{x}_3) s_{44}(\mathbf{x}_4)$ может, по всей видимости, интерпретироваться как "масштабная инвариантность" свойств четырехмерного пространства-времени с метрикой Бервальда-Моора и/или как его "пространственно-временная изотропность". В отличие от Бервальд-Мооровской метризации четырехмерного пространства-времени, четырехмерное пространство с метрикой Минковского, с точностью до масштабирования совпадающего с $\mathbf{H}_4$, снабженным метрической формой $S_{24}(\mathbf{x})$, обладает только свойством "пространственной но не "пространственно-временной" изотропией. Группа (линейных) изометрий пространства $\mathbf{H}_4$ с метрикой Минковского хорошо известна и вне связи с Теоремой 2.3.

Проблема 3. Могут ли быть получены преобразования Лоренца из соотношений (22) и (23), то есть, как прямые следствия явных соотношений Теоремы 2.3?

4. Пусть $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ есть "какой-то" базис в $\mathbf{H}_4$, изотропный, например. Пусть $\mathbf{B}$ оператор, не обязательно линейный, действующий из $\mathbf{H}_4$ в $\mathbf{H}_4$, так что

$$\mathbf{B}\mathbf{x} = \mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3 + y_4\mathbf{e}_4 \in \mathbf{H}_4.$$

Условие того, что $\mathbf{B}$ сохраняет форму $s_{24}(\mathbf{x})$, то есть норму $N_2(\mathbf{x})$, может быть записано в виде

$$N_2(\mathbf{x}) = [\mathbf{x}, \mathbf{x}] = [\mathbf{Bx}, \mathbf{Bx}] = [\mathbf{y}, \mathbf{y}] = N_2(\mathbf{y}) = N_2(\mathbf{Bx}). \quad (30)$$

Но, в силу линейности операции $[\mathbf{w}_1, \mathbf{w}_2]$, соотношение (30) можно переписать в виде

$$[\mathbf{x}, \mathbf{x}] = \sum_{i,j=1}^4 x_i x_j [\mathbf{e}_i, \mathbf{e}_j] = [\mathbf{y}, \mathbf{y}] = \sum_{i,j=1}^4 y_i y_j [\mathbf{e}_i, \mathbf{e}_j].$$

Если оператор $\mathbf{B}$ линейный, то массив $\{b_{ij} = [\mathbf{e}_i, \mathbf{e}_j]; i, j = 1, 2, 3, 4\}$, как известно, полностью определяет действие оператора $\mathbf{B}$ на всем $\mathbf{H}_4$. Более того, для линейных изометрических операторов условия на числа b_{ij} могут быть получены из Теоремы 2.3. Если требовать того, чтобы $\mathbf{B}$ сохраняло форму $s_{34}(\mathbf{x})$, то есть норму $N_3(\mathbf{x})$, то также в силу линейности операции $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$, получаются условия изометричности (нелинейного) оператора $\mathbf{B}$ в терминах трехмерного массива $\{b_{ijk} = [\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k]; i, j, k = 1, 2, 3, 4\}$ и явных соотношений Теоремы 2.2.

Проблема 4. Пользуясь Теоремой 2.2, описать операторы $\mathbf{B}$, сохраняющие форму $s_{34}(\mathbf{x})$, то есть норму $N_3(\mathbf{x})$.

Работа выполнена при поддержке некоммерческого фонда развития исследований по финслеровой геометрии.

Список литературы

- [1] Общая алгебра. (Под ред. Л.А.Скорнякова). М.: Наука, 1990
- [2] Allenby, R.B.J.T. Rings, Fields and Groups: An Introduction to Abstract Algebra. 2nd edition, 1991
- [3] Кантор И.Л., Солодовников А.С. Гиперкомплексные числа. М.: Наука, 1973.
- [4] Чернов В.М. Об определяющих уравнениях элементов ассоциативно-коммутативных конечномерных алгебр и ассоциированных метрических формах // Гиперкомплексные числа в геометрии и физике, №4, 2005, с.57-74.
- [5] Chernov V.M. On defining equations for the elements of associative and commutative algebras and on associated metric forms // Space-Time Structure, Moscow, TETRU, 2006, pp.7-24
- [6] Zassenhaus H. Über endliche Fastkörper // Abh. Math. Sem. Hamburg, 11, 1936, 187-220.
- [7] Холл М. Теория групп. ИЛ, 1962.
- [8] Ленг С. Алгебра. М.:Мир, 1968.
- [9] Павлов Д.Г. Обобщение аксиом скалярного произведения // Гиперкомплексные числа в геометрии и физике, №1, 2004, с.5-19.
- [10] Pavlov D.G. Generalisation of scalar product axioms // Space-Time Structure, Moscow, TETRU, 2006, pp.7-24.

Экспериментальные исследования и астрофизические наблюдения, свидетельствующие о неоднородности и анизотропии реального пространства-времени.

В.А. Панчелюга

Институт гиперкомплексных систем в геометрии и физике, Фрязино, Россия.

Институт теоретической и экспериментальной биофизики РАН, Пущино, Россия

E-mail: panvic333@yahoo.com

Задача данного курса – рассмотреть эксперименты, которые можно на том или ином основании рассматривать, как свидетельство неоднородности и анизотропии реального пространства-времени. Какие экспериментальные феномены могут рассматриваться, как имеющие отношение к свойствам пространства-времени? После работ А. Эйнштейна принято считать, что пространственно-временная структура вселенной связана с проявлениями гравитационного взаимодействия. Поэтому, одно из правил отбора экспериментальных работ для нашего курса – работы, связанные с гравитационным воздействием.

Из известных физических взаимодействий только два – электромагнитное и гравитационное – играют значительную роль на больших расстояниях. При этом, условно говоря, электромагнитное взаимодействие является определяющим на планетарных расстояниях, а для межпланетных и межзвездных расстояний более важным является гравитационное. Поэтому, если наблюдаемое в эксперименте воздействие имеет космофизическую природу, то оно может рассматриваться в нашем курсе. Понятно, что здесь можно “прихватить лишнего”. В реальности ситуация обстоит немного иначе: мы имеем космофизическое воздействие, на сложную тест-систему. Механизм этого воздействия неизвестен. Но наличие определенных периодов (звездный, суточный, лунные, солнечные периоды) позволяют, в первом приближении, говорить о космофизической природе и относить эти эксперименты к нашей тематике. Хотя во втором приближении – нужно разбираться в действующих механизмах.

С другой стороны, если не касаться упомянутой связи между геометрией и физическими взаимодействиями, то необходимо отметить, что воздействие связанное со свойствами пространства-времени должно иметь наиболее универсальный характер – оно будет отражаться в динамике систем различной качественной природы. Отсюда – следующее правило отбора: мы будем рассматривать экспериментальные исследования, в которых обнаруживается подобная универсальность.

Представленные ниже эксперименты подбирались так, чтобы удовлетворять описанным правилам.

Для данного издания были отобраны, по преимуществу, те экспериментальные работы информация о которых разбросана в многочисленных журнальных публикациях, зачастую труднодоступных. Мы намерено не приводим здесь широко известных работ, таких как, например, по анизотропии реликтового излучения, т.к. о них можно легко получить самую различную информацию от популярной до специализированной. Предполагается, что читатель, при желании, может сделать это самостоятельно. Представленный ниже текст примерно соответствует 2-4 лекциям (в зависимости от степени детализации), читаемым на Летней школе.

Введение

За последние десятилетия появился ряд сообщений о необычных эффектах, причины которых трудно интерпретировать на основании традиционных представлений. К ним относятся эффекты, наблюдаемые при исследованиях “слабых” и “сверхслабых” воздействий на биологические и физико-химические объекты (системы). Наиболее популярным объяснением возможных причин эффектов “слабых” и “сверхслабых” воздействий является специфическое действие электромагнитных полей. В тоже время, имеется ряд сообщений о возможной гравитационной природе некоторых из этих эффектов. Такое объяснение представляется парадоксальными при сопоставлении этих воздействий с амплитудой термодинамических флуктуаций (kT). Однако, многие феномены при исследованиях “слабых” и “сверхслабых” воздействий могут приобрести рациональную интерпретацию при учете пространственной неоднородности и анизотропии нашего мира. Основанием для этого мнения являются работы последних десятилетий. В предлагаемой статье представлен их краткий обзор.

Как правило, исходным материалом для изучения внешних влияний на исследуемую систему являются достаточно длинные временные ряды, получаемые в результате многократных, высокостандартизированных измерений некоторой величины. Общеизвестны методы обработки и анализа временных рядов (методы статистической обработки результатов измерений, поиска периодичностей, вычисления корреляций, оценок достоверности умозаключений). При этом свидетельством зависимости измеряемой величины от различных факторов является корреляция изменений этой величины с изменениями “силы” воздействия. На этом основаны работы с поиском корреляций исследуемых величин с изменениями солнечной активности, напряженности межпланетного магнитного поля, температуры, давления, концентрации различных веществ и т.п. Поиск таких корреляций существенно облегчается в случаях, когда “сила” воздействий изменяется периодически. Сходство периодов изменений изучаемой величины и изучаемого агента (фактора), обычно рассматривается как доказательство зависимости данной величины от этого фактора. В работах всей этой группы исследований речь идет о “воздействиях”, “влияниях” изучаемого фактора на исследуемую величину.

В нашей лаборатории, при исследовании различных эффектов, получил развитие метод сравнения тонкой структуры распределений результатов измерений, тонкой структуры соответствующих гистограмм, построенных по коротким, непересекающимся отрезкам временных рядов [3, 6, 10]. *При этом, обнаружены феномены, могущие быть следствием не прямого, специфического влияния некоторого фактора на исследуемую величину, а неспецифического изменения пространства-времени, сходным образом отражающегося в процессах совершенно разной природы.* Среди причин этих неспецифических изменений могут быть как изменения “масштаба” нашего мира, обусловленного движением в неоднородном гравитационном поле, так и изменения направлений в анизотропном пространстве [1-12].

1 Об особенностях во временных рядах флуктуаций результатов измерений.

Студентов учат делать не менее двух измерений. Их называют “параллельными”. Если они не совпадают – делают еще одно. Выбирают два близких... Если делать много, на самом деле последовательных, во времени измерений “при прочих равных условиях”, оказывается,

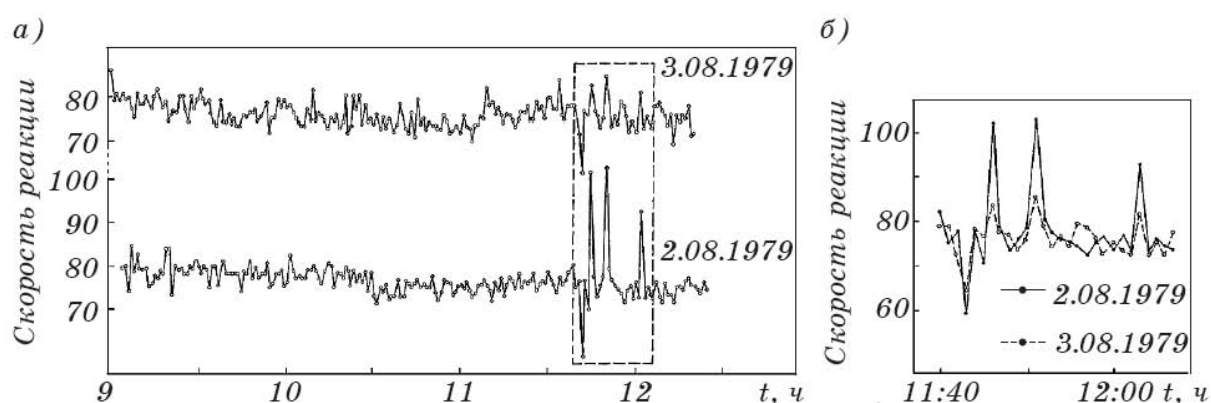


Рис. 1: **а)** Макроскопические флуктуации скорости реакции АК и ДХФИФ. ББС МГУ, 2 и 3 августа 1979. В рамках выделены соответствующие сигналы. По оси ординат – скорость реакции в относительных единицах, по оси абсцисс – время в часах. **б)** Сопоставление сигналов 2 и 3 августа. [13]

что “выбросы” неустранимы. Много лет назад в нашей лаборатории были начаты кругло-суточные, последовательные измерения флуктуаций в ходе различных процессов. Вначале это были многолетние наблюдения флуктуаций скорости протекания различных биохимических реакций, после – мониторинг флуктуаций скорости α -распада. Тонкая структура гистограмм, построенная по результатам этих измерений, оказалась очень информативной. В ней проявились закономерности, не выявляемые обычными методами [1-12]. Однако, наряду с исследованием формы гистограмм, внимание привлекли возможные закономерности реализации таких “выбросов” – резких отклонений от среднего уровня измеряемых величин. Возникло предположение, что эти выбросы, при соответствующем временном разрешении, представляют собой “сигналы”, форма которых может быть вполне закономерной.

Специально форму таких “сигналов” попытались исследовать в 1978-1979 г.г. В много-часовых опытах, в разных географических пунктах измерялись флуктуации скорости ферментативной реакции ($\text{АТФ} + \text{креатин} \leftrightarrow \text{АДФ} + \text{фосфокреатин}$), катализируемой креатинкиназой, и скорости реакции аскорбиновой кислоты (АК) с дихлорфенолиндофенолом (ДХФИФ). Результаты этих исследований частично были опубликованы в 1981 году [13].

В этих исследованиях [13] при синхронных измерениях скорости креатинкиназной реакции и скорости взаимодействия аскорбиновой кислоты с дихлорфенолиндофенолом в Пушкино, Симферополе, Алма-Ате, Пояконде (ББС МГУ) было обнаружено сходство форм низкочастотных компонент – трендов – при сдвиге временных рядов, соответствующем различию местного времени. При этом, исходные, несглаженные временные ряды флуктуаций оставались некоррелированными. Кроме сходства трендов были обнаружены воспроизводимые временные последовательности – “сигналы”. Эти “сигналы” наблюдались на фоне низкоамплитудных флуктуаций, трудно отличимых от ошибок метода измерений, как относительно высокоамплитудные флуктуации продолжительностью $\sim 20 - 25$ мин.

На рис. 1 а) и рис. 2 взятом из [13], в качестве примера, пунктирными рамками выделены “сигналы” во временных рядах изменений скорости реакции АК и ДХФИФ, выполненных в деревянном “немагнитном” доме на Беломорской биостанции МГУ в августе 1979 года. На рис. 1 б) “сигналы”, полученные 2 и 3 августа 1979 г., рис. 1 а), наложены друг на друга

для удобства сравнения. Очевидно, практически полное совпадение формы для высокоамплитудной части “сигналов”. То же можно отметить и в отношении “сигналов”, показанных на рис. 2. “Сигналы”, зарегистрированные 2 и 3 августа (рис. 1 а) совпадали не только по форме, но и по времени. Тогда как, сходные по форме временные последовательности, зарегистрированные 4-6 августа (рис. 2) наблюдались в разное время суток и в разное время от начала опытов.

“Сигналы”, аналогичные показанным на рис. 1 и рис. 2, наблюдались также при измерениях скорости реакции АК+ДХФИФ в опытах проведенных в Пущино, Алма-Ате и других географических пунктах. Аналогичные явления наблюдались и при измерениях скорости креатинкиназной реакции. Тогда не удалось установить какие-либо воспроизводимые зависимости проявления этих сигналов от времени суток, сезонов или географических координат.

За последующие годы был опубликован ряд сообщений об аналогичных явлениях. Представленные в этих сообщениях наблюдения, как правило, не поддавались интерпретации на основании традиционных представлений. Сама публикация таких сообщений была сопряжена с известными трудностями. Однако и, будучи опубликованными, многие сообщения очень редко входили в “научный метаболизм” и подвергались забвению. Нам кажется, что в настоящее время актуальным является возобновление интереса к работам такого рода – во многих из них содержатся основания для принципиально новых направлений исследований. В связи с этим мы попытались дать, представленный ниже, краткий обзор этих работ.

1.1 Космофизическая обусловленность флуктуаций зрительного порога человека.

Этот обзор мы хотели бы начать с упоминания об исследованиях механизмов адаптации органов чувств человека, проводившихся начиная с 1913 г. академиком П.П. Лазаревым [14].

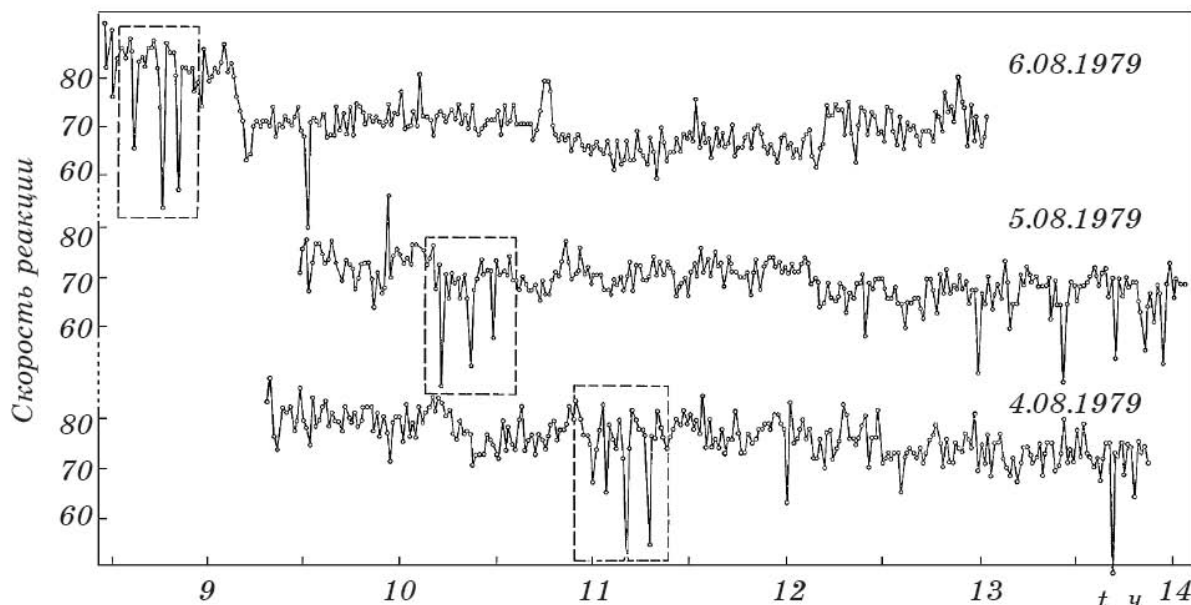


Рис. 2 Макроскопические флуктуации скорости реакции АК и ДХФИФ. ББС МГУ, 4-6 августа 1979. В рамках выделены соответствующие сигналы. По оси ординат – скорость реакции в относительных единицах, по оси абсцисс – время в часах. [13]

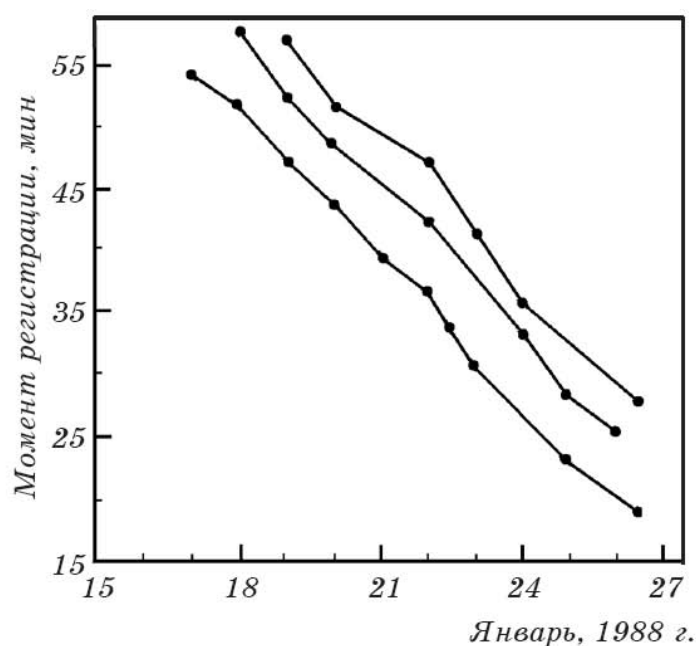
В ходе исследований колебаний чувствительности для периферического зрения, в зависимости от разного рода факторов, им было обнаружено: существование суточной цикличности с максимумом около двух часов дня и минимумом после полуночи, сезонные колебания с максимумами весной и осенью и минимумом зимой, зависимость чувствительности от широты местности и от высоты над уровнем моря. И хотя сам П.П. Лазарев, после серии исследований, объяснил перечисленные зависимости, в основном, влияниями изменений температуры и давления, эти работы послужили отправной точкой для дальнейших исследований, в которых исходным экспериментальным материалом служили временные ряды флуктуаций чувствительности зрительного порога человека.

Одним из таких исследований являются работы **Вячеслава Евгеньевича Жвирблиса** [15, 16] по наблюдению за дрейфом нулевой точки визуального полутеневого поляриметра. Им было обнаружено, что нулевая точка испытывает сложнопериодический дрейф, стационарная компонента которого удовлетворительно коррелирует с гелиогеофизическими индексами. На основании этого было высказано предположение, что дрейф вызывается космофизическими факторами.

В серии работ **А.Д. Сизова** также использовались временные ряды значений чувствительности (дифференциального зрительного порога) глаза человека. Так, в [17] проводились синхронные определения положения фотометрического равновесия двумя независимыми операторами с использованием поляриметра и нефелометра. Общим в этих приборах является использование для измерений визуального уравнивания яркостей смежных полей зрения. В одних опытах операторы находились в одном и том же помещении, в других на расстоянии ~ 10 км друг от друга [17]. Корреляционный анализ результатов измерений обнаружил их воспроизводимость при сопоставлении в одни и те же моменты по местному звездному времени. На основании этого была выдвинута гипотеза об “экзогенном факторе”, который синхронно действует на операторов, но при этом ими не осознается.

При сопоставлении результатов несинхронных, выполненных одним оператором, измерений во временных рядах были обнаружены характерные экстремальные значения, сильно выступающие за уровень среднего. Анализ наблюдений показал, что эти экстремальные значения иногда появляются в виде групп (или серий) состоящих из 3-5 аномальных значений, продолжающихся в течение 5-7 мин. Некоторые из этих групп устойчиво наблюдаются на протяжении нескольких последовательных дней, в одно и то же звездное время [18]. На рис. 3, взятом из [18], показаны тренды моментов регистрации трех экстремальных значений в последовательные сутки опыта. Сдвиг трендов близок к значению 4 мин/сутки, что эквивалентно утверждению о том, что исследованная группа из 3-х аномальных значений появляется примерно в один и тот же момент по звездному времени. Также отмечается, что величины интервалов времени между последовательными моментами регистрации экстремальных значений образуют дискретный ряд значений [20].

Эти результаты привели к предположению о космофизическом воздействии на человека, как на “приемник”, связанном с определенными положениями небесной сферы относительно испытуемого. Возникла идея сопоставить координаты некоторых звезд с моментами регистрации экстремальных значений зрительного порога. Т.к., экстремальные значения появлялись нерегулярно, то для первоначального анализа были выбраны звезды, имеющие переменные характеристики, т.е., переменные звезды-цефеиды. Предпосылки обращения к звездам, как причине получаемой феноменологии представлены в [20].



Дата обнаружения экстремального значения

Рис. 3 Тренд моментов регистрации экстремальных значений дифференциального зрительного порога глаза человека в последовательные сутки опыта. Кривые 1-3 соответствуют трем аномальным значениям. По оси абсцисс – дата обнаружения экстремального значения. По оси ординат – момент регистрации, мин. [18].

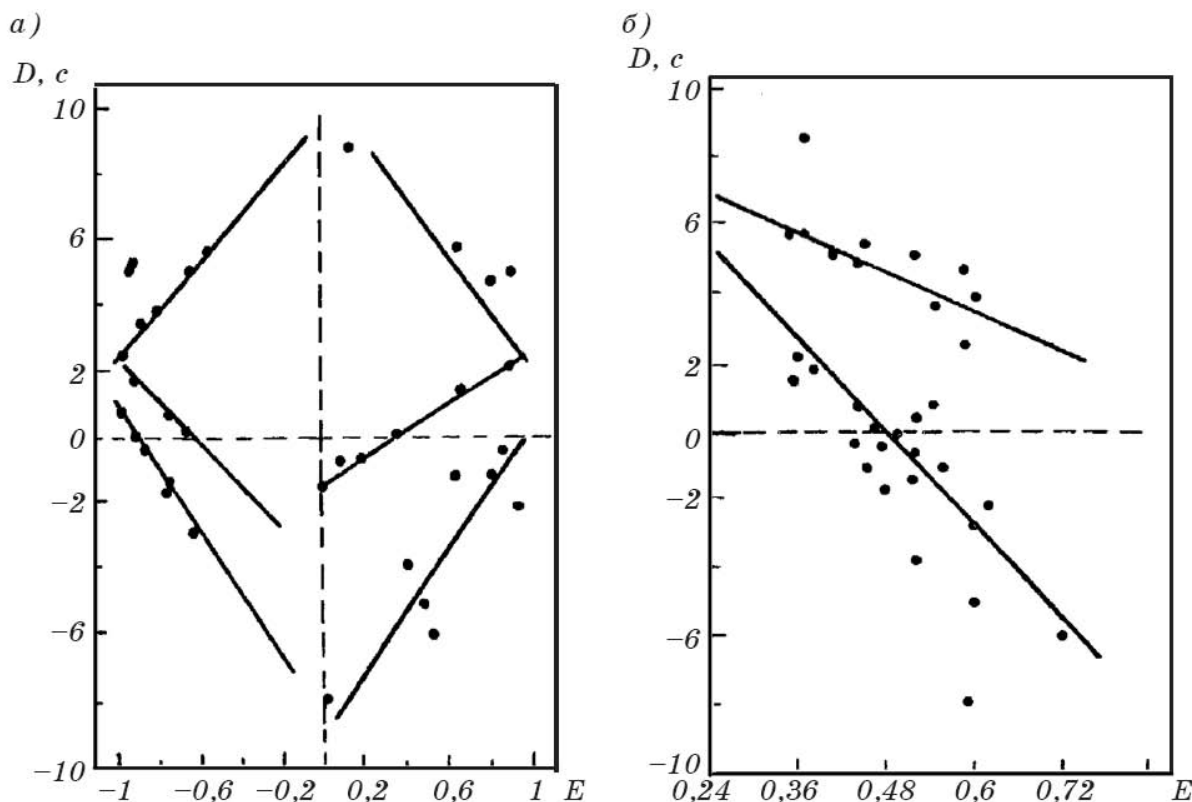


Рис. 4 **а)** Сопоставление средних значений D и E (пояснения в тексте). **б)** Сопоставление средних значений D и усредненных за период опыта, значений фазы периода изменения блеска (пульсации) цефеид (ось абсцисс). [18].

©В.А. Панчелюга Экспериментальные исследования и астрофизические наблюдения, свидетельствующие о неоднородности и анизотропии реального пространства-времени.

В [18, 21-25] проанализированы моменты регистрации экстремальных значений дифференциального зрительного порога человека исходя из предположения об их связи с эклиптическими координатами переменных звезд. Измерения порога чувствительности, на основе которых выполнялся анализ, были организованы так, что они включали моменты кульминации цефеид. При этом, если в пределах 30 сек до и после момента кульминации звезды наблюдалось аномальное значение порога чувствительности глаза, то рассчитывалась разность D между временем регистрации t аномального значения порога чувствительности и моментом кульминации t_0 звезды:

$$D = t - t_0. \quad (1)$$

Для каждой звезды по всем результатам опыта рассчитывалось среднее значение D , которое сопоставлялось со средним значением произведения косинусов эклиптических координат E :

$$E = \cos(L - L_0) \cdot \cos(b), \quad (2)$$

где L и L_0 - эклиптическая долгота звезды и гелиоцентрическая долгота Земли на момент регистрации сигнала, b - эклиптическая широта звезды. Результаты этого сопоставления, рис. 4), оказались симметричными относительно нулевого значения E . Можно видеть, что D тем меньше чем ближе звезда к плоскости эклиптики и направлению Солнце-Земля или D тем больше, чем ближе находится звезда к полюсу эклиптики, или чем меньше разность эклиптических долгот Земли и звезды отличаются от 90° .

Разность $L - L_0$ зависит от орбитального движения Земли, следовательно, величина D должна иметь сезонную зависимость. Действительно, значения D обнаруживают тренд уже в пределах 30-40 суточного непрерывного опыта, при этом направление тренда зависит от положения звезды относительно вектора скорости орбитального движения Земли [18]. Следовательно, значение D должно зависеть также от относительной скорости "приемника" и источника [22]. Т.к. скорость пульсационного движения газовых масс цефеид связана с периодом изменения их блеска, то должна наблюдаться зависимость значений D от фазы блеска цефеид. На рис. 4) суммированы результаты, свидетельствующие о существовании такой зависимости [18].

Также исследовалась зависимость длительности экстремальных значений дифференциального зрительного порога от фазы изменения блеска цефеид [26]; отклик зрительной системы в зависимости от расстояния до звезд [27]; зависимость от суточного вращения Земли [29]. Было сформулировано предположение о том, что действующим агентом, вызывающим наблюдаемые эффекты является монохроматическое излучение звезд, воздействующее на сенсорную систему человека в актах резонансного поглощения γ -квантов [28]. Вследствие доплеровского смещения частоты, вызванного относительным движением "приемника" и источника, условие резонансного поглощения может нарушаться. Поэтому, число звезд, которые могут влиять в данный момент времени на "приемник" относительно мало. Это, по мнению А.Д. Сизова, объясняет возможность наблюдения откликов от отдельных звезд, при том, что их общее число очень велико.

Представленные выше результаты, получены с использованием измерений дифференциального зрительного порога человека. Для возможности осуществления непрерывного многосуточного мониторинга была поставлена задача поиска физической системы, флуктуации

в которой были бы подобны флуктуациям дифференциального зрительного порога человека [18, 30]. Основную идею такого поиска дала внешняя аналогия последовательных во времени измерений дифференциального зрительного порога с записью фликкер-шума от полупроводниковых приборов. В результате А.Д. Сизовым, в качестве детектора был выбран заключенный в термостабилизированный объем неуравновешенный мостик Уитстона, в одно из плеч которого был включен фоторезистор, а в две диагонали моста – источник ЭДС и чувствительный индикатор-самописец.

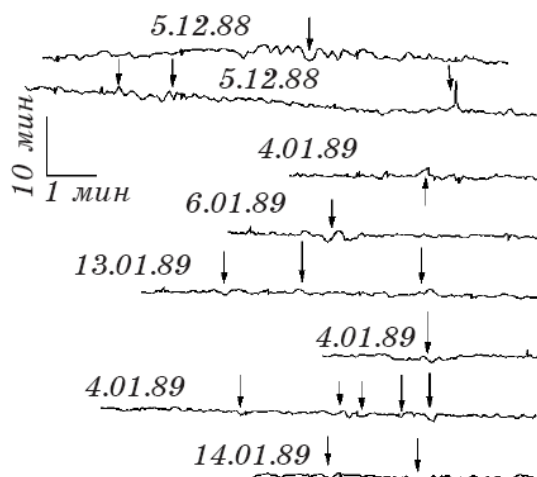


Рис. 5 Формы сигналов, записанные с мостика Уитстона вблизи моментов кульминации (показаны стрелками) звезд. Рисунок взят из [18].

В процессе регистрации были получены записи, рис. 5, внешне напоминающие запись автоколебаний, частота которых изменялась вблизи момента кульминации звезды – картина напоминала запись биений при интерференции волн мало отличающихся по частоте. Моменты времени, которые сопоставлялись моментам кульминации звезд, показаны на рис. 5. стрелками.

Результаты исследований флуктуаций тока в мостике Уитстона представлены в [31]. Анализировалась 7-месячная непрерывная запись, в которой было зарегистрировано свыше 50000 флуктуаций длительностью более 3 с, но лишь 397 флуктуаций длительностью не менее 400 с, которые условно назывались сигналами. На рис. 6 а) приведена гистограмма распределения числа длительных сигналов за время опыта. Рис. 6 б) дает распределение числа длительных сигналов по времени суток. Обнаружено, что максимум сигналов приходится на время местного полдня (9-11 ч всемирного времени). Это ведет к предположению о выделенности плоскости небесного меридиана и о том, что источники регистрируются приемником в течение 0.5 часа от момента верхней кульминации светила. Исследования сигналов, регистрируемых мостиком Уитстона, обнаружили зависимость продолжительности “отклика” в зависимости от величины проекции единичного вектора по направлению звезда-Солнце на линию Солнце-Земля [32].

В работе [33] анализируются влияния планет на регистрирующую систему. Было показано, что экстремальные значения параметров сигналов чаще всего наблюдаются при одновременной реализации особых планетных конфигураций, когда угловые расстояния между планетами Солнечной системы в пределах ошибки в несколько градусов равны: 0, 90, 180 или

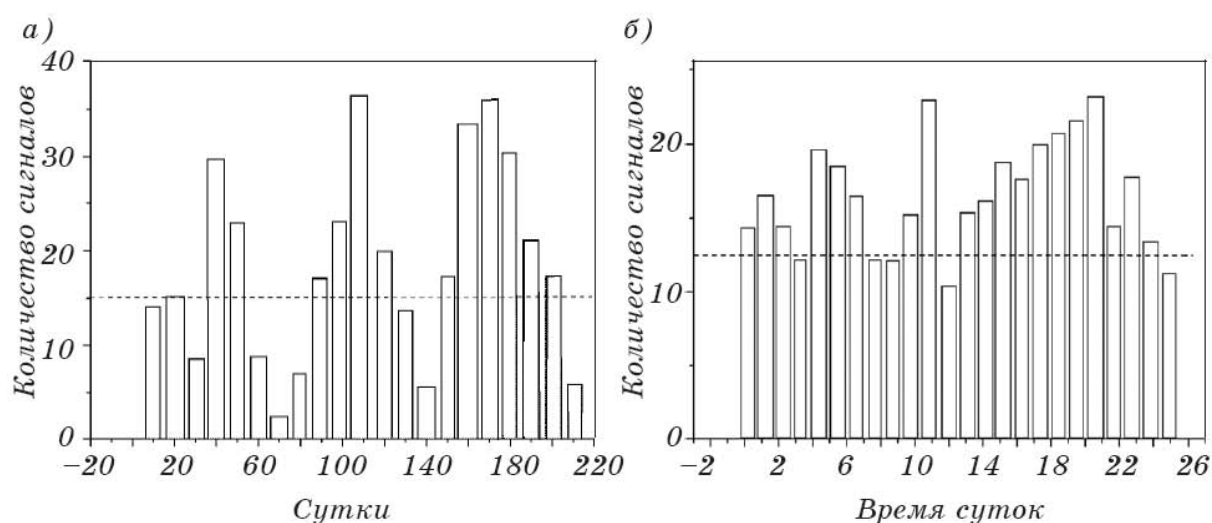


Рис. 6 Гистограммы распределения числа длительных сигналов за время опыта **а)**, по времени суток, **б)**. Пунктир - ожидание. Рисунки взяты из [31].

270 градусов. Также отмечается, что длительность сигналов возрастает в моменты верхней кульминации Луны или Солнца. Автор полагает, что влияние планет объясняется экранированием ими определенных звезд-источников.

Исходя из полученных результатов, А.Д. Сизов предлагает методы определения межзвездных расстояний без применения телескопов [34, 35], а также определение скорости пульсационного движения газовых оболочек и лучевых скоростей звезд [36].

1.2 Н.А. Козырев и последователи.

Работы А.Д. Сизова в части использования мостика Уитстона перекликаются с исследованиями Н.А. Козырева [37, 38], с тем отличием, что регистрирующая система А.Д. Сизова является ненаправленной, а Н.А. Козырев и В.В.Насонов использовали телескопы или системы зеркал, позволяющие селективно исследовать те или иные астрономические объекты. Это позволило, по-видимому, впервые зарегистрировать эффект неизвестного воздействия высокой проникающей способности, исходящего от космических источников (звезды, галактики). Условия эксперимента Н.А. Козырева позволяют говорить о неэлектромагнитной природе регистрируемого воздействия, поскольку оптический вход телескопа перекрывался металлическим экраном [38].

Исследования Козырева были повторены новосибирской группой под руководством М.М. Лавренътева [39-41]. В одном из экспериментов, где предполагаемым источником воздействия служило Солнце, было показано, что физический и биологический датчики реагируют сходным образом только при наведении телескопа с перекрытым входным отверстием на так называемое "истинное" Солнце, которое, находясь на траектории движения видимого Солнца, опережает его на $2^{\circ} 4/6''$. Подобное возможно, если воздействие от "истинного" Солнца распространяется со скоростью значительно превышающей скорость света.

А.Ф. Пугач [11], проведя серию наблюдений по методике Козырева в Киеве и Крыму, подтвердил реальность неэлектромагнитного воздействия космических источников на физические системы, отметив, что источники проявляют свойства транзиев, т.е. эпизодически появляются и исчезают за порогом чувствительности аппаратуры. Этими же авторами ста-

вится задача экспериментального определения скорости распространения воздействия.

1.3 В.А. Дубровский: спектры микросейсм и распределение звездных объектов.

С перечисленными выше исследованиями по методике Н.А. Козырева и работами А.Д. Сизова перекликаются работы В.А. Дубровского, для которых определяющим также является влияние звезд на процессы, происходящие на Земле [43-46]. В этих работах, в ходе измерений зависимости амплитуды микросейсмического фона от частоты, проводившихся с 1987 г. по 1993 г. на севастопольской сейсмической станции Таврического университета с использованием лазерного интерферометра, установленного в штольне на глубине 20 м, было обнаружено шесть пиков в районе частот 2.3, 1.0, 0.9, 0.6, 0.4, 0.2 Гц [43]. Т.к. не было найдено ни техногенных ни естественных сейсмогенных факторов, ответственных за пики на указанных частотах, то была проанализирована их возможная космофизическая обусловленность. Оказалось, что существуют массивные звездные объекты на расстояниях L_i равных: 1.3, 2.7, 3.5, 5.0, 8.1, 11.1 (парсек) [47, 48]. Все эти шесть расстояний L_i можно взаимно однозначно сопоставить с шестью частотными пиками ν_i с помощью простой формулы:

$$\nu_i = \frac{C}{L_i}. \quad (3)$$

При этом, каждому частотному пику сопоставляется соответствующее расстояние примерно с одной и той же константой C , имеющей размерность скорости и равной $C \sim 3 \cdot 10^{19}$ /. На рис. 7 показано расчетное распределение гравитационного потенциала звезд, а также приведены значения ν_i и L_i . В.А. Дубровский отмечает, что найденная экспериментально скорость этого взаимодействия находится в хорошем согласии с не опровергнутой до сих пор оценкой Лапласа нижнего предела скорости гравитационного взаимодействия, а также теоретическими и экспериментальными оценками, приведенными в [44, 45]. Этот результат можно рассматривать, как согласующийся с результатами Н.А. Козырева [38] и М.М. Лаврентьева [39], отмечавших, что скорость взаимодействия, связанного с воздействием звезд на регистрирующую систему должна превышать скорость света.

В ходе дальнейших исследований была проделана та же самая, что и для севастопольских данных, процедура сравнения микросейсмического фонового спектра с распределением звездных масс, но для данных, полученных в Аризоне, для более низкочастотной части микросейсмического спектра. Найдено, что севастопольские и аризонские результаты достаточно хорошо перекрываются, совпадая на трех пиках. [46]

Исходя из полученных результатов, выдвигается гипотеза о резонансно-волновом взаимодействии в системе “Земля-звезды”, осуществляемом при помощи гравитационных волн и приводящем к условию связи частоты фоновых микросейсм ν с расстоянием до звезд r :

$$\nu \sim \frac{C}{r}. \quad (4)$$

Амплитуда микросейсм P должна быть пропорциональна количеству звездных масс, находящихся на расстоянии r . Поскольку, это количество растет с расстоянием как r^2 , то с учетом (4):

$$P \sim r^2 \sim \frac{1}{\nu^2}. \quad (5)$$

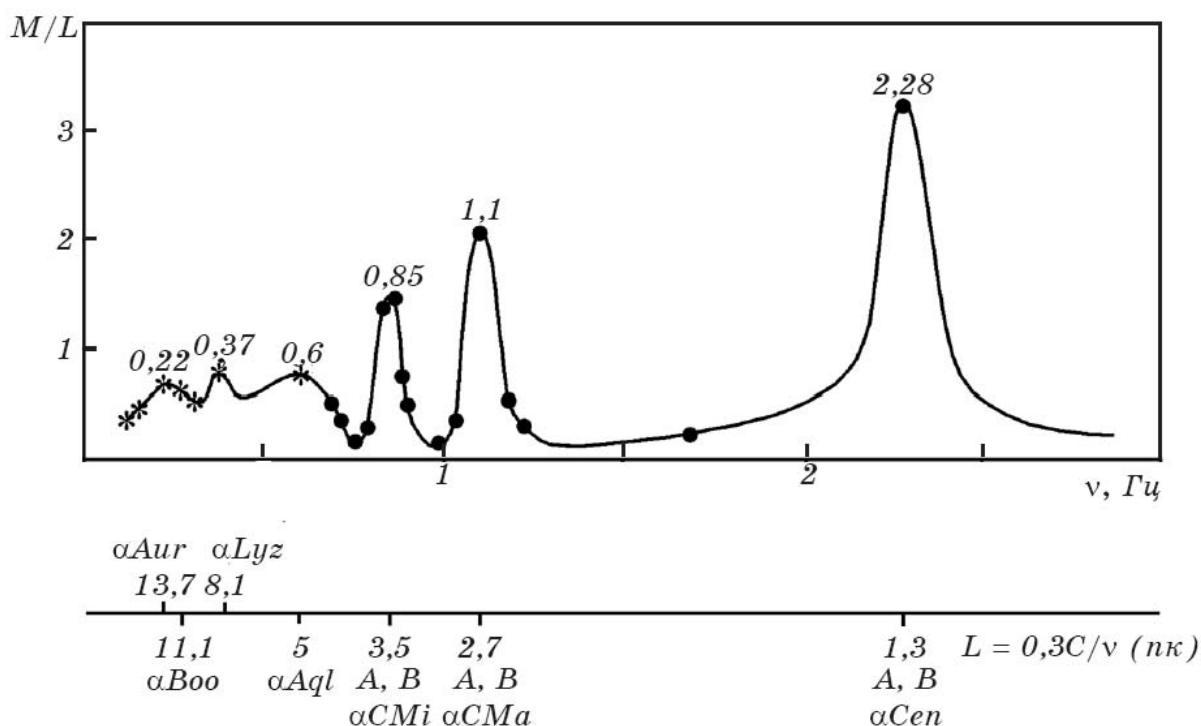


Рис. 7 Расчетное распределение гравитационного потенциала звезд. Жирными точками показаны все ближайшие звезды на расстоянии $L < 4$ парсеков ($1 \cong 3.1 \cdot 10^{18}$ см), а звездочками все ближайшие ярчайшие звезды для которых $L > 4$ парсеков. Масса M дана в массах Солнца. αAur – αCen обозначают α -звезды созвездий в стандартных астрономических обозначениях [43, 44]. A, B обозначают звездные дублеты. Числа около точек экстремумов кривых обозначают частоты. Соответствующие этим частотам значения $L = 0.3C/\nu()$ являются расстояниями до наиболее массивных звезд или их скоплений в созвездиях. Рисунок взят из [41].

Т.е., как следует из (5), амплитуда микросейсм возрастает на низких частотах обратно пропорционально квадрату частоты. В силу быстрого увеличения с расстоянием количества звездных объектов, индивидуальное влияние каждого из них нивелируется и вливается в общий гладкий фон. Это, в итоге, приводит к гладкой кривой зависимости амплитуды микросейсм от частоты в области малых частот.

Т.о., выявляется единая картина соотношения микросейсмического фона с распределением звездных масс как вблизи Солнечной системы (высокочастотная составляющая микросейсм) так и в дальнем космосе (гладкая низкочастотная часть спектра), т.е. выявляется механизм, определяющий общие черты микросейсмического спектра.

Отмечается, что в случае планет Солнечной системы должны существовать резонансные пики, соответствующие волновому взаимодействию Земли с Луной (~ 238 МГц), Солнцем (~ 0.6 МГц), Венерой (0.36-2.2 МГц), Юпитером (100-146 кГц), Сатурном (58-71.6 кГц). Пики, соответствующие Венере, Юпитеру или Сатурну должны изменять свое частотное положение в указанных пределах в зависимости от изменения расстояния между Землей и этими планетами в процессе орбитального движения вокруг Солнца.

Для представленных выше исследований характерно наличие связи между регистрируемыми особенностями временных рядов и координатно-временным положением звезд или

производным от этого положения гравитационным потенциалом в точке наблюдения. Представленная ниже серия исследований в своих истоках, будучи совершенно независимой, от рассмотренных выше работ, по мере своего развития привела к результатам, во многом совпадающим с теми, которые были описаны выше.

1.4 Связь флуктуаций и выбросов в различных процессах с ходом λ_D функции.

Эти работы были начаты в начале 1970-х годов, когда **Виктор Владимирович Соколовский** предложил измерять скорость окисления унитиола в качестве теста при исследованиях корреляций с изменениями солнечной активности и аналогичными феноменами [49]. Несмотря на стандартизированные условия измерений, были обнаружены многочисленные случаи значительного ускорения или замедления этой реакции, коррелирующие с флуктуациями солнечной активности [50, 51]. На примере 21-го цикла солнечной активности была установлена однозначная (обратная) функциональная зависимость между среднегодовыми флуктуациями времени полуокисления унитиола τ_u и уровнем солнечной активности (по числам Вольфа), что свидетельствует о влиянии Солнца на кинетику данной реакции. Этот результат совпадает с результатами наших исследований кинетики биохимических и химических процессов [52], проводившихся в 1957-1985 г.г.

На рис. 8, взятom из [52], в процентном отношении представлены временной ход чисел Вольфа, W , и относительная амплитуда $\sigma \%$ (отношение среднеквадратичного отклонения σ к среднему значению измеряемой величины, выраженное в процентах) разброса результатов измерений. Обращает на себя внимание очевидная зеркальная корреляция этих величин. Из известных нам космофизических процессов в такой же зеркальной противофазности к солнечной активности находится интенсивность потока нейтронной компоненты галактических космических лучей [53].

Такое совпадение результатов независимых исследований наводит на мысль об универсальности обратной зависимости амплитуды флуктуаций в протекании различных процессов от уровня солнечной активности.

В то же время, как отмечается в обеих работах [50, 52], при более детальном (поемечасном и посуточном) рассмотрении временных соотношений между уровнем флуктуаций и солнечной активностью, часто наблюдается их несогласованность. В [50] было высказано предположение о наличии десинхронизирующих агентов иной природы, ответственных за эту несогласованность. С целью их выявления был проведен почти годовой цикл ежесуточных измерений τ_u в период глубокого минимума солнечной активности (с 15 июля 1996 г. по 1 июня 1997 г.) на полярной станции Мирный в экологически чистых условиях Антарктики [50]. На рис. 9, точками, представлены результаты этих измерений. Видно, что τ_u меняется нерегулярным образом от 560 до 190 мин. Было высказано предположение, что такой ход τ_u может быть обусловлен наложением флуктуаций с разным периодом. На рис. 9 и) показаны кривые, полученные усреднением τ_u по 5 и 11 точкам, соответственно. Анализ полученных таким образом кривых выявил два типа флуктуаций: со средним периодом около 35 дней и квазидвухнедельные флуктуации с периодом около 15 дней. Анализ долгопериодного тренда показал наличие полугодового и годового периодов [50].

Поиск физической интерпретации для найденных периодов и годового тренда показал корреляцию (> 0.9) с неравномерным поступательным движением Земли на орбите под влиянием Солнца, которое, наряду с годовым периодом, включает полугодовую волну в измене-

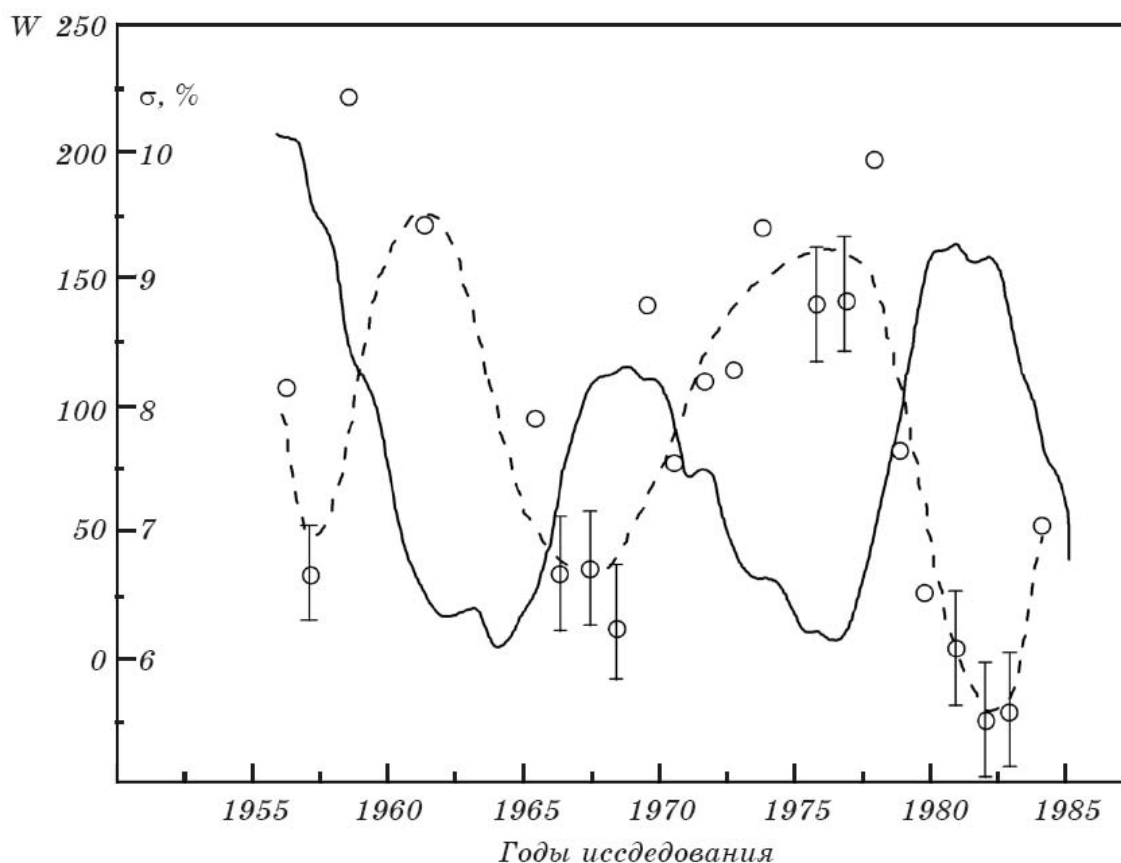


Рис. 8 Сопоставление изменений относительной амплитуды разброса результатов измерений (σ %) с ходом солнечной активности, W . Кружками представлены среднегодовые значения σ %, пунктир – их наилучшая аппроксимация полиномом по методу наименьших квадратов, сплошная линия – среднемесячные сглаженные значения чисел Вольфа W . Отмеченные 95% доверительные интервалы для среднегодовых значений σ % показывают, что эти значения в годы максимумов и минимумов солнечной активности достоверно отличаются. [52].

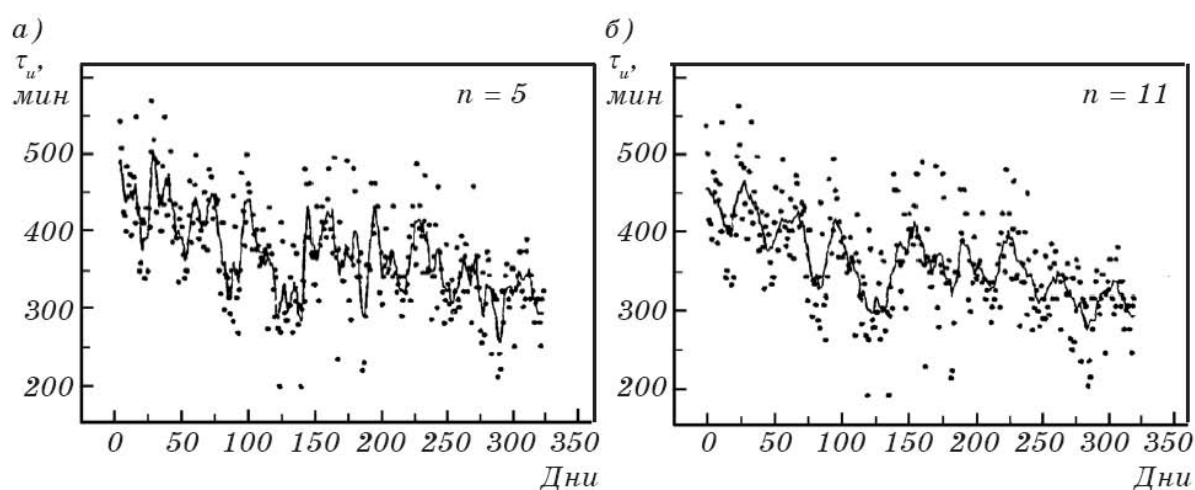


Рис. 9 Ход ежесуточных значений скорости окисления унитиола (точки) и результат усреднения (сплошная кривая) по 5 дням, *а*), и по 11 дням, *б*). Рисунки взяты из [50].

нии хода поступательной скорости. Более короткие периоды были связаны с неравномерностью вращательного движения Земли под действием Солнца и Луны. Это сложное движение раскладывается на прецессионное (вековое) с периодом около 26 тыс. лет и нутационное с периодами от 4 дней до 18.6 лет [54]. Оказалось, что найденные экспериментально периоды лучше всего соответствуют вариации (14.8 дня) и эвекции (31.8 дня) – нутационным периодам, в которых в наибольшей степени проявляется вклад элонгации, т.е. разницы средних долгот Луны и Солнца.

В работе [55] в течение года (2001-2002 г.г.) на антарктической станции Восток исследовалась скорость окисления унитиола в моче. Полученные результаты обнаружили спектр периодов, соответствующих солнечной и геомагнитной активности, а также периоды равные периодам эвекции и нутации. Интересно отметить, что флуктуации скорости реакции в данном эксперименте коррелировали не только с солнечной и геомагнитной активностью, но также и с со скоростью окисления унитиола *in vitro*.

Представленные в [56] результаты мониторинга содержания в крови лимфоцитов и сегментоядерных нейтрофилов, определение величин лейкоцитарного коэффициента и скорости оседания эритроцитов у хронических больных, проанализированные за 10.5 месяцев, также выявили периоды 13-16 дней и 31-32 дня близкие к периодам вариации и эвекции. Кроме того, были обнаружены периоды 3-5, 7, 9-10, 18-22 дня.

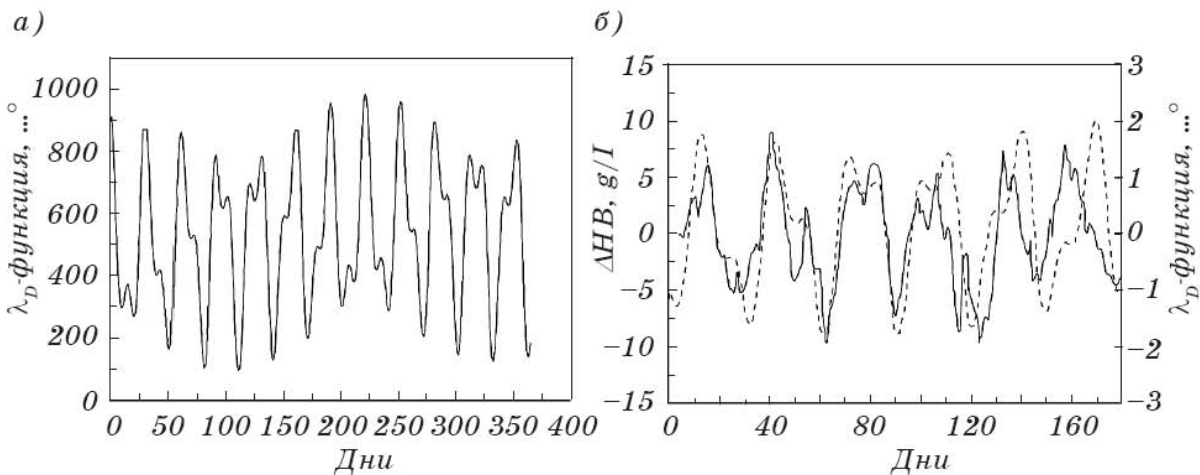


Рис. 10 **а)** Пример временного хода λ_D – функции для 2000 г., **б)** Временной ход короткопериодных вариаций содержания гемоглобина ΔHb (сплошная линия) и временной ход λ_D – функции (точечная линия). Рисунки взяты из [59] – а) и [58] – б).

В статье [57] исследуется связь между динамикой биохимических реакций и ходом λ_D – функции, являющейся суперпозицией трех основных нутационных периодов: 14.8 дней (вариация), 31.8 дней (нутация) и 182.6 дней (годовое неравенство), рис. 10). Данная функция не связана с лунными фазами и, следовательно, с приливными эффектами, и отражает периодические вариации гравитационного поля, обусловленные совместным влиянием Солнца и Луны [58]. Основной итог, рассмотренных выше работ [57, 58], состоит в том, что периоды длительностью 14.8 и 31.8 дней, описывающие короткопериодические нутационные движения Земли, являются типичными для различных физико-химических и биологических процессов в природе, т.к., флуктуации в протекании этих процессов хорошо согласуются с ходом λ_D – функции. Рис. 10) иллюстрирует временной ход короткопериодных вариаций

содержания гемоглобина ΔHb (сплошная линия) и временной ход λ_D – функции (точечная линия) [58]. Видно, практически детальное совпадение хода обеих кривых. Из этого следует, что вариации гравитационного поля, судя по всему, являются определяющими в воздействии на ритмы любых физико-химических и биологических процессов на Земле. На этом фоне солнечная активность, действует как десинхронизирующий фактор, характеризующийся высокой степенью нерегулярности и спонтанности. Те же выводы можно распространить и на чисто физические системы. Так в [59] исследуется годовой ряд (с 1 февраля 1998 г. по 1 февраля 1999 г., антарктическая станция Восток) флуктуаций хода компьютерных часов относительно временных меток GPS-системы. Для него также получено хорошее согласие с ходом λ_D – функции.

Представленные выше результаты были получены на основе анализа временных рядов флуктуаций результатов измерений. Но, в ходе этих работ, на фоне относительно монотонного хода временных рядов, были также обнаружены резкие выбросы – “импульсные сигналы”. Такие импульсные сигналы были зарегистрированы в процессе колориметрирования смеси растворов унитиола и нитрита натрия с реактивом Элмана на разных стадиях окисления [61]. Систематические наблюдения за динамикой таких сигналов проводили в период с 05.05.97 по 01.06.97 на станции Мирный в Антарктике и во время обратного рейса судна “Академик Федоров” из Антарктиды в Санкт-Петербург с 03.07.97 по 02.09.97. Эти наблюдения показали, что определяющими признаками сигналов являются: высокая проникающая способность, зависимость от долготы Солнца в дневное время, совпадение времени их появления с моментами восходов и кульминаций ряда звезд и радиопульсаров, вероятность появления в широком диапазоне географических широт. Длительность зарегистрированных сигналов варьировалась от нескольких секунд до минуты. Из совокупности этих признаков был сделан вывод, что сигналы представляют собой импульсное неэлектромагнитное излучение космофизической природы.

После обнаружения импульсных сигналов [61] было предпринято исследование тока “пустого” микрофотоколориметра, используемого без рабочего вещества [60]. В этом случае также были зарегистрированы импульсные сигналы. Рис. 11) дает пример регистрируемых сигналов. Была сделана запись, продолжительностью 8.5 месяца (с 4 ноября 2000 г. по 20 июля 2001 г., г. Санкт-Петербург), в ходе которой было зарегистрировано около 3000 статистически значимых сигналов. Рис. 11) дает распределение числа зарегистрированных импульсных сигналов. На распределении выделяются два события 23 марта и 19 апреля 2001 г., которые характеризуются аномально высоким числом зарегистрированных сигналов. Эти события предшествуют мощным солнечным вспышкам, произошедшим 27 марта и 23 апреля 2001 г. Дальнейший анализ данных, представленных на рис. 11) показал их связь с ходом λ_D – функции. Спектральный анализ обнаружил периоды, близкие по величине к вариации и эвекции. Период в 27 суток, типичный для солнечной активности не был найден. Полученные результаты убеждают в том, что регистрируемые сигналы не являются артефактами и скорее всего обусловлены космофизическими воздействиями неэлектромагнитной природы. В этом убеждает их связь с ходом λ_D – функции. Связь сигналов с солнечными вспышками говорит в пользу их солнечного происхождения.

В работе [62] предлагается детектор сверхслабых излучений с повышенной чувствительностью к неэлектромагнитной, и в том числе, к гравитационной, компоненте спектра воздействующих физических полей.

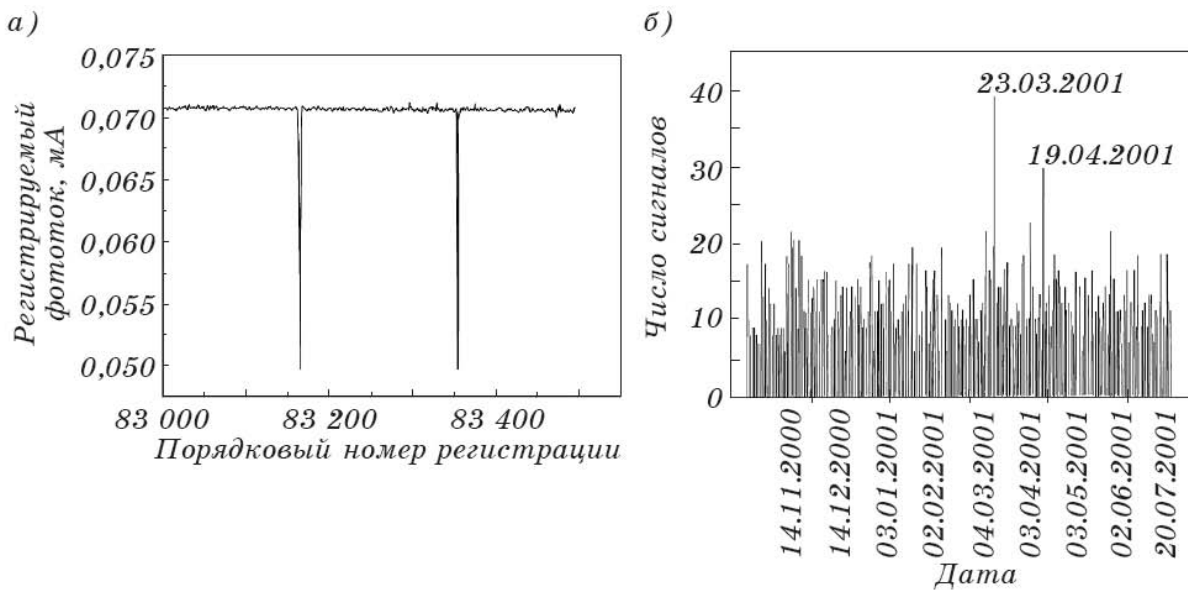


Рис. 11 *а)* Пример импульсных сигналов в фототоке микрофотокolorиметра;
б) Распределение числа импульсных сигналов, зарегистрированных в течение суток за период с 4 ноября 2000 г. по 20 июля 2001 г. (Измерения проводились в г. Санкт-Петербург). Рисунки взяты из [60].

Представленные работы [49-51, 55-62], имеют революционный характер. После многих лет общего убеждения в обусловленности космофизических корреляций влиянием слабых, относительно низкочастотных электромагнитных полей [63], получены, на наш взгляд, убедительные данные о возможной зависимости наблюдаемых явлений от возмущений неэлектромагнитной природы. О регистрации излучения неэлектромагнитной, предположительно гравитационно-волновой, природы сообщается также в работах Н.В. Клочека [64, 65].

1.5 Работы Н.В. Клочека и сотрудников: Т-сигнал и его свойства.

В этих работах в качестве чувствительных элементов использовались прецизионные кварцевые механические резонаторы опорных автогенераторов стандартных электронносчетных частотомеров с рабочей частотой 5 МГц. Эти резонаторы защищены от внешней среды: вакуумированы, термостатированы и экранированы от внешних электромагнитных полей. Кроме того, они были определенным образом ориентированы в пространстве. Для выделения относительноных противофазных изменений резонансных частот один из частотомеров системы работал в режиме датчика секундных меток, другой – приемника, измеряющего период следования этих меток в своей временной шкале [65].

Регистрируемый сигнал автор называет Т-сигналом, очевидно из-за того, что генераторы системы были расположены под углом 90° друг к другу. Такое расположение, как показано в [66], приводит к тому, что эффективность воздействия гравитационной волны на кристалл каждого из генераторов неодинакова и, следовательно, она может проявить себя как относительные противофазные изменения их резонансных частот.

Примеры экспериментальных регистраций Т-сигнала приведены на рис. 12. Здесь, рис. 12), который взят из работы [64] за 1992 г., демонстрирует фрагмент записи сигнала, для которого характерно наличие всплесков очень большой амплитуды, которые, по мнению авторов,

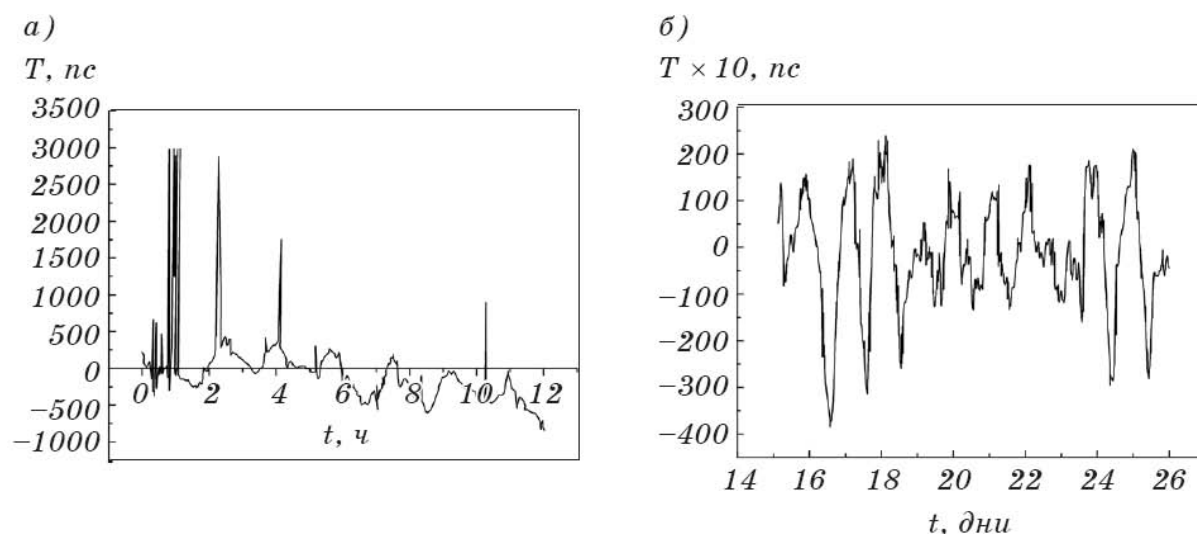


Рис. 12 *а)* Фрагмент временного хода регистрируемого сигнала за период 5-11 августа 1991 г. *б)* Фрагмент записи Т-сигнала 15-25 августа 1991 г. Рисунки *а)* и *б)* взяты из [64] и [65], соответственно.

имеют физический смысл. В более поздней работе [65] за 1995 г. приводится пример сигнала, показанный на рис. 12). В этой работе всплески, подобные приведенным на рис. 12) не рассматриваются. Необходимо отметить, что оба сигнала относятся к близким временным интервалам: 5-11 августа 1991 г и 15-25 августа 1991 г.

Как можно видеть из приведенных экспериментальных записей, для них характерно присутствие вариаций порядка десятков минут и более на фоне четко выделяющегося суточного ритма переменной амплитуды. Достаточно длительный ряд наблюдений (около года) позволил выделить основные черты суточного хода Т-сигнала. Рис. 13 демонстрирует суточный ход Т-сигнала. Обращает на себя внимание увеличение сигнала в утренние часы и в местный полдень, существенное уменьшение сигнала в вечерние часы и устойчивое поведение сигнала в ночное время. Определенно такой характер суточного хода зависит от местного времени, т.е. от положения Солнца относительно регистрирующей системы.

Можно допустить, что уровень Т-сигнала связан с уровнем солнечной активности. Сопоставление плотности потока солнечного радиоизлучения на разных частотах с поведением Т-сигнала вблизи местного полдня показало, что наилучшая корреляция наблюдается для частоты 4995 МГц (6 см). Радиоизлучение соответствующее этой частоте, генерируется в верхней хромосфере, в то время как для частоты 260 МГц (1.15 м), относящейся к уровню верхней короны корреляция с Т-сигналом практически отсутствует. Вместе с тем, изменение конфигурации системы, а именно, увеличение расстояния между резонаторами, привело к изменению суточного хода Т-сигнала, рис. 13). В данном случае наилучшая корреляция наблюдалась с плотностью потока радиоизлучения Солнца на частоте 260 МГц. Следовательно, можно думать, что симметричное относительно момента кульминации возрастание Т-сигнала обязано областям верхней короны. По-видимому, Солнце является источником рассматриваемого неэлектромагнитного волнового излучения в широком диапазоне “длин волн”. Принципиально, что изменение конфигурации резонансной системы дает потенциальную возможность ее настройки на разные диапазоны.

Сравнение спектра мощности Т-сигнала со спектром рентгеновского излучения Солнца

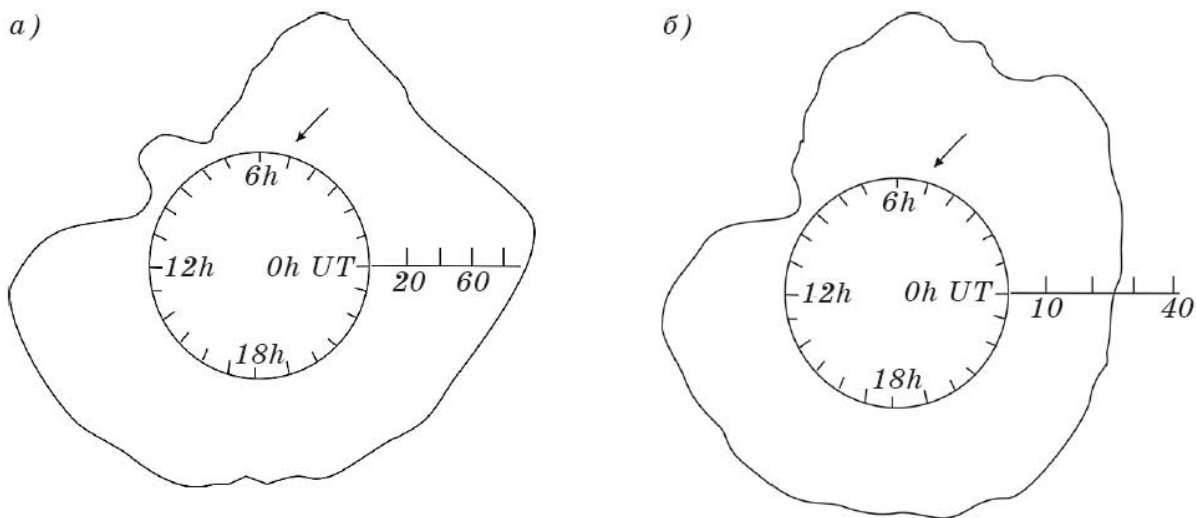


Рис. 13 Суточный ход Т-сигнала, а); то же, после изменения конфигурации системы, б).

показало наличие практически совпадающих около часового периода и периода 36-38 мин. Такие же периодичности наблюдались в работе [67] в вариациях относительной мощности двух стандартов частоты и, при этом, рассматривались как возможный результат воздействия гравитационной волны солнечного происхождения. Отсюда делается вывод, что солнечное электромагнитное излучение и исследуемое излучение неэлектромагнитной природы, по-видимому, имеют одни и те же пространственно-совмещенные источники.

Для исследования биологической активности регистрируемого излучения Н.В. Ключеком были проведены параллельные регистрации Т-сигнала, а сотрудниками НИИ биологии ИГУ (Иркутск) под руководством Д.И. Стомма – измерения интенсивности свечения культуры фотобактерий, находящихся в средоизолирующем боксе,. Оказалось, что изменение светимости фотобактерий, определяемое колориметрическим методом, и поведение относительной суточной дисперсии Т-сигнала во временном интервале 18 марта - 5 апреля 1991г. весьма сходны. Коэффициент корреляции составил 0.9, что является весомым доводом в пользу биологической активности регистрируемого излучения.

Также была показана хорошая корреляция Т-сигнала с К-индексом геомагнитной активности для среднеширотной зоны и изменениями электрокожного сопротивления человека. Отмечается опережение дисперсии Т-сигнала по отношению к геомагнитной активности. Опережение составляет 3-5 суток и тем больше, чем больше возмущение геомагнитного поля.

Н.В. Ключек [64, 65] заключает, что имеются многочисленные теоретические предпосылки и экспериментальные свидетельства в пользу того, что реально существует космофизическое воздействие неэлектромагнитной, скорее всего гравитационно-волновой природы, источником которого являются космические объекты, в том числе Солнце.

Результаты близкие к полученным Н.В. Ключеком были обнаружены также в независимых исследованиях А.Н. Морозова. Его исследования временной зависимости дисперсии флуктуаций подвижности ионов обнаружили ее устойчивый, повторяющийся ход на суточном периоде, близкий к зависимости, показанной на рис. 13. Автором делается вывод, что флуктуации подвижности ионов могут быть обусловлены пространственной анизотропией [68].

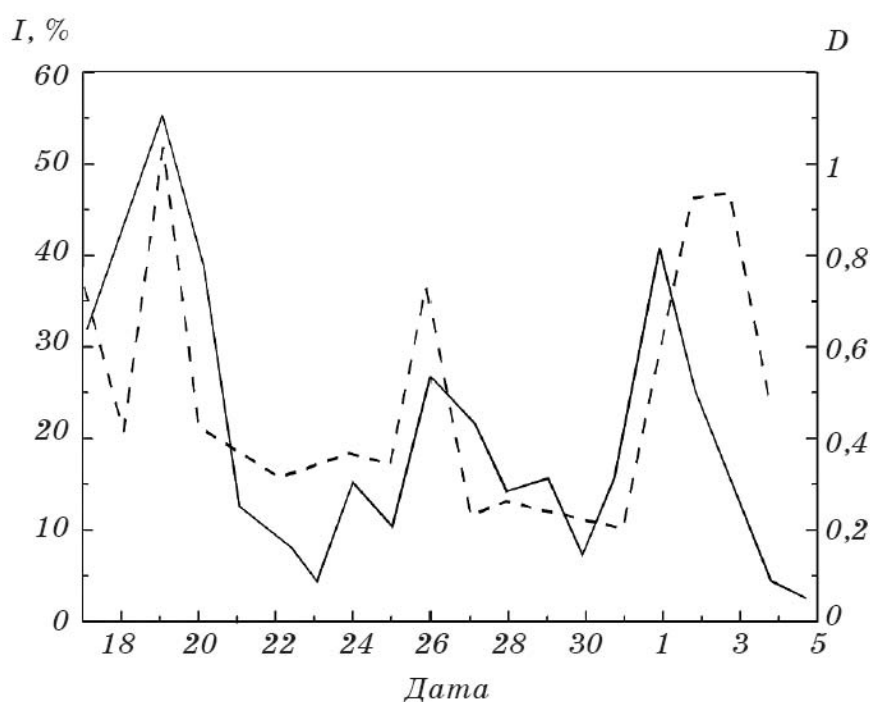


Рис. 14 Интенсивность свечения культуры фотобактерий (1) и относительная суточная дисперсия Т-сигнала 18.03-5.04.1991 (2). [65].

1.6 Исследования В.Н. Смирнова: детектор гравитационных воздействий.

Как и для многих из уже упоминавшихся работ, истоки идеи исследований В.Н. Смирнова также можно проследить в экспериментальных работах Н.А. Козырева [69]. В данном случае речь идет об исследованиях с применением гироскопов. Но, в отличие от работ Н.А. Козырева, где опыты по воздействию некоторых космофизических событий на гироскопическую регистрирующую систему иногда получались, а иногда – нет, В.Н. Смирнову удалось создать устройство, позволяющее уверенно регистрировать ряд событий, представленных ниже. Необходимо отметить, что исследования с использованием устройства В.Н. Смирнова нельзя рассматривать, как законченные. Поэтому, несомненно, что представленный список событий с течением времени будет пополняться.

Центральным элементом устройства В.Н. Смирнова, как и у Н.А. Козырева, является быстро вращающийся волчок. Но на этом, конструктивные аналогии с экспериментами Н.А. Козырева заканчиваются. Устройство В.Н. Смирнова является самостоятельной, оригинальной конструкторской разработкой.

Детальное описание устройства В.Н. Смирнова дано в [70-74]. Его базовым элементом является латунный волчок весом 265 г, установленный на магнитной платформе, которая, в свою очередь, подвешена в сильном магнитном противопололе. Скорость вращения волчка 3600 об/мин. Одна из главных особенностей устройства заключена в специальном режиме вращения: на каждом обороте производится импульсное торможение, длительность которого находится в пределах 18-30% от периода вращения и в каждом конкретном случае подбирается экспериментально. Только при соблюдении этого условия появляется искомая реакция вращающейся массы на координатно-временное положение небесных тел, состоящая в изменении угловой скорости вращения. Чувствительность устройства зависит от выбора

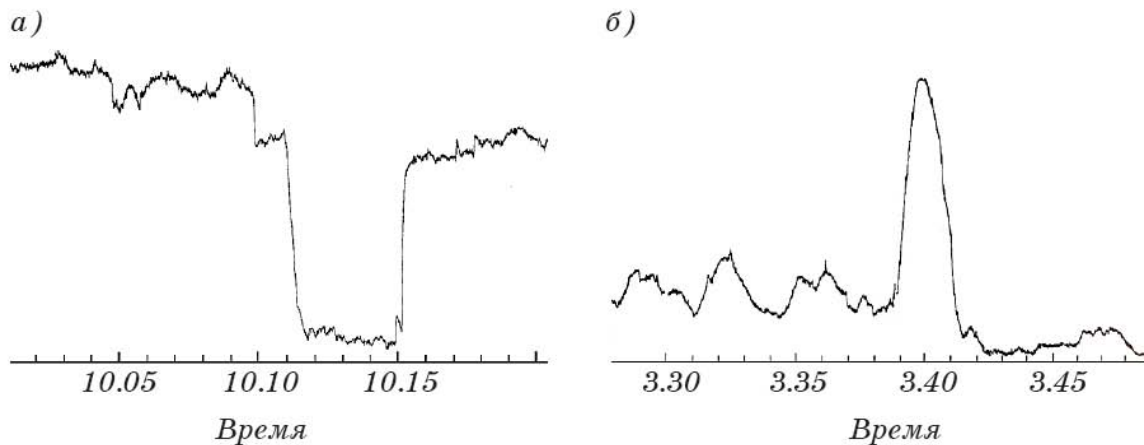


Рис. 15 Примеры регистраций солнечного затмения: **а)** Начало солнечного затмения в 10 ч. 06 мин. 29.03.06 на восточном выступе Южной Америки (Бразилия); **б)** кольцеобразное солнечное затмение 22.08.1998. Геоцентрическое соединение в 3ч 38мин 23 сек. (Примеры регистраций предоставлены автором)

необходимого импульсного торможения и скорости вращения. Она также зависит от момента подачи тормозящего импульса, определяющего пространственную ориентацию вектора торможения. Момент подачи импульса задается поворотом специального азимутального диска и позволяет управлять пространственным положением плоскости, задаваемой вектором торможения и осью вращения волчка, и имеющей, по сути, смысл диаграммы направленности. Устройство наиболее эффективно регистрирует события, которые попадают в эту плоскость. Информация о физическом состоянии всей подвесной системы и, следовательно, самого волчка поступает со специального индукционного датчика. Из низкочастотного спектра колебаний, возбуждаемых в датчике, селективным усилителем выделяется первая гармоника, равная частоте вращения волчка, которая затем преобразуется в аналоговый сигнал и поступает на графопостроитель [70-74].

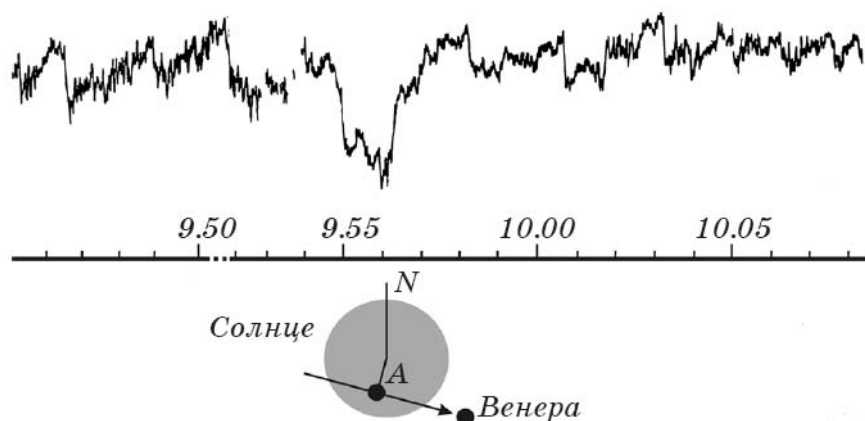


Рис. 16 Прохождение Венеры по диску Солнца 8.06.2004. Пример регистрации взят из [72]

Представление о чувствительности устройства дают следующие оценки. Изменения энер-

гии волчка при регистрации различных событий (заходы Солнца и Луны, момент наступления новолуния, моменты наступления перигея и апогея Луны) находятся в пределах 0.028-0.36 Дж [70]. При средней частоте вращения волчка 65 Гц измерение частоты при заходе Солнца составляет 0.82 Гц, а при лунном затмении – 13 Гц [71].

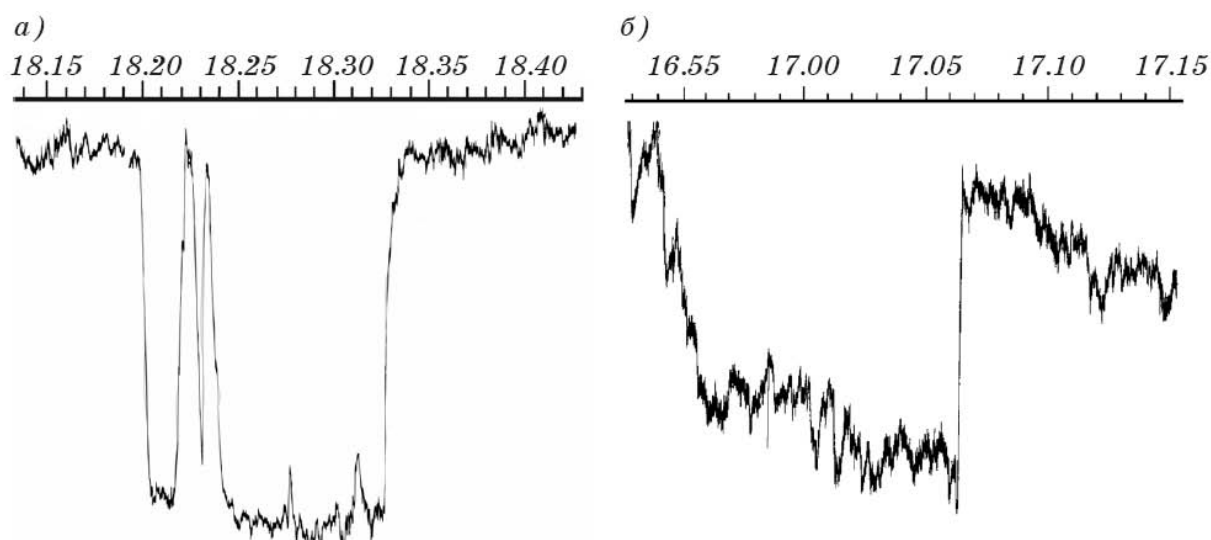


Рис. 17 *а)* Сигнал, интерпретируемый, как предвестник землетрясения в западном Иране 1-2.04.06. Регистрация произведена 29.03.06. Азимут ориентации устройства 9° относительно точки юга. Время московское. *б)* То же на острове Суматра 5.07.05. Регистрация произведена 30.06.05. Сила землетрясения 6.75 балла. Азимут ориентации устройства 53° на восток относительно точки юга. Время московское. Примеры регистраций предоставлены автором.

В [73] волчок рассматривается как резонаторная система, в которой, при помощи, специального режима торможения, возбуждаются упругие стоячие волны, приводящие к деформации его структуры и, следовательно, диссипации энергии при преодолении внутреннего трения. Отмечается, что физика волчка в устройстве В.Н. Смирнова адекватна физике, заложенной в теорию волноводных твердотельных гироскопов, только там компенсация энергии осуществляется специальным электрическим полем, а в устройстве В.Н. Смирнова внешним воздействием, имеющим, предположительно, гравитационно-волновую природу. Поэтому, такая система рассматривается как возможный регистратор гравитационно-волнового воздействия.

Рассмотрим события, регистрируемые устройством.

События в системе Солнце-Луна-Земля: моменты восходов Солнца и Луны и их заходы за линию горизонта, моменты наступления лунного и солнечного затмений, причем последние могут регистрироваться и тогда, когда они происходят в южном полушарии. Кроме того, к регистрируемым ближним астрономическим событиям относятся моменты наступления полнолуния, новолуния, прохождения Луной точек апогея и перигея. Устройство реагирует на моменты наступления зимнего и летнего солнцестояний (когда происходит резкое изменение величины проекции солнечного диска на направление угловой скорости Земли), а также афелий Земли – моменты прохождения Землей экстремальной (т.е. наиболее удаленной от Солнца точки) [70, 73].

Рис. 15 дает примеры регистраций солнечного затмения. Полярность пиков, показанных

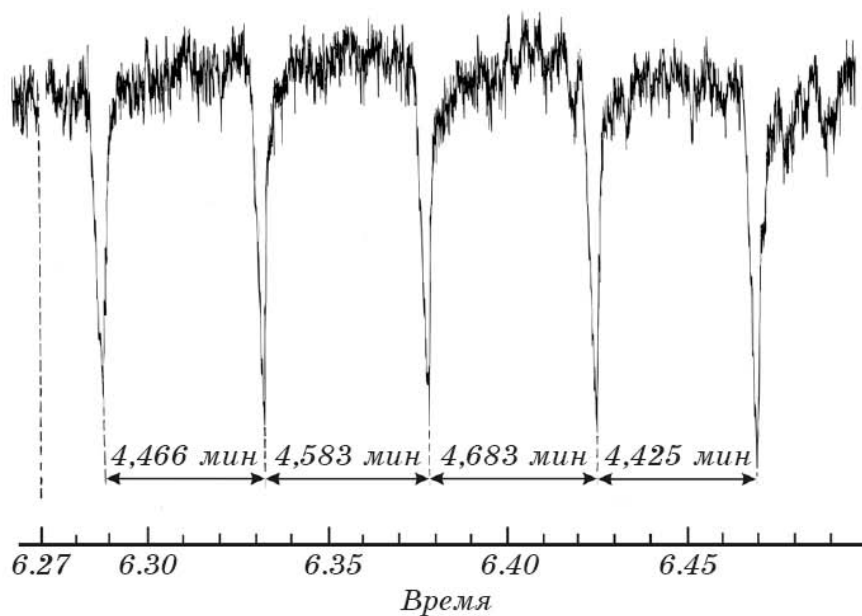


Рис. 18 Фрагмент записи периодических импульсов, зарегистрированных 31.05.2003, установкой В.Н. Смирнова. Рисунок взят из [72].

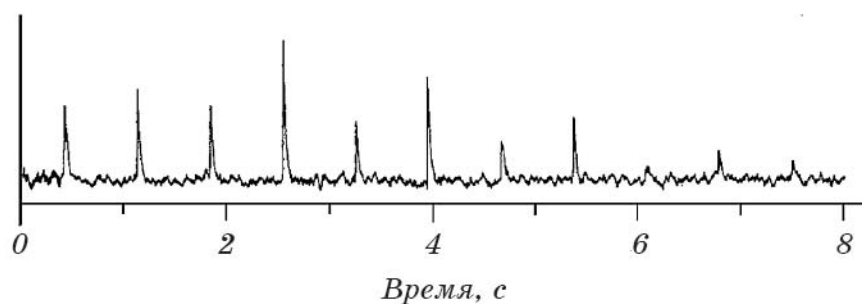


Рис. 19 Пример записи отдельных импульсов пульсара PSR 0329+54. Период следования импульсов составляет около 0.714 сек. Рисунок взят из [86].

на рис 15) и рис 15) в данном случае, значения не имеет, т.к., она связана с особенностями настройки регистрирующей аппаратуры. На рис. 15) дан пример регистрации начала солнечного затмения в 10 ч. 06 мин 29.03.06 видимого на восточном выступе Южной Америки (Бразилия). Регистрация проводилась в г. Москве, при этом, азимут устройства был выставлен по направлению 60° на запад относительно направления на юг. Регистрация, представленная на рис. 15), получена во время кольцеобразного солнечного затмения 22.08.1998 в окрестности момента геоцентрического соединения, произошедшего в 3 ч 38 мин и 23 сек. Как можно видеть, оба пика на рис. 15 несколько запаздывают по отношению к началу регистрируемых событий.

Из событий, связанных с орбитальными конфигурациями планет Солнечной системы, регистрировались: момент захода Юпитера, прохождение Венеры по диску Солнца [72]. На рис. 16 дан пример регистрации связанной с прохождением Венеры по диску Солнца 8.06.2004. Точка “А” на рис. 16 обозначает момент, когда угловое расстояние между центрами Солнца и Венеры достигает минимума в 9 ч 51 мин.

Кроме чисто космических событий устройство регистрирует события, связанные с подвижками тектонических масс. События, которые были зарегистрированы, обычно предшествуют землетрясениям на 5-10 дней и могут рассматриваться как предвестники.

Ряд событий, регистрируемых устройством, пока не поддается интерпретации. Так, 31.05.2003 была осуществлена регистрация периодических импульсов (Рис. 18). Аналогичные импульсы наблюдались и в октябре 2000 г. Т.о., Земля при движении по орбите как бы пересекала ось между созвездиями Телец (май) и Дева (октябрь) [72].

Для сравнения, на рис. 19 показаны импульсы радиоизлучения пульсара PSR 0329+54. Полярность импульсов показанных на рис. 18, как уже отмечалось, значения не имеет и определяется настройкой аппаратуры. Обращает на себя внимание подобие формы импульсов, показанных на рис. 18 и рис. 19. Скважность импульсов практически совпадает. В то же время, период следования импульсов, показанных на рис. 18 , в среднем 4.5 мин, в то время как средний период следования для импульсов пульсара PSR 0329+54 – 0.714 сек [86]. Как известно, пульсары рассматриваются как один из источников гравитационно-волнового излучения.

Как легко видеть из рисунков 15÷ 18, события, регистрируемые устройством В.Н. Смирнова, всегда имеют вид четко различимых пиков с амплитудой в несколько раз превышающей среднюю амплитуду флуктуаций временного ряда. В этом – одно из главных отличий исследований В.Н. Смирнова от ранее рассмотренных работ, где регистрируемые события имеют амплитуду сопоставимую со средней амплитудой флуктуаций и, поэтому, требуются специальные методы, чтобы эти события выделить. Это обстоятельство говорит в пользу высокой чувствительности метода В.Н. Смирнова. В то же время, более “мелкие” особенности во временных рядах В.Н. Смирнова, как правило, не анализируются. Несомненно, что подобный анализ может выявить реакцию устройства на новые, неизвестные к настоящему времени события.

1.7 К.А. Труханов, Ю.А. Бауров: векторный космологический потенциал.

Работы, рассмотренные ниже, были инициированы исследованиями эффекта Ааронова-Бома. В частности, предположением о физической реальности векторного потенциала. Мы включили их в настоящий обзор т.к. предполагаемые проявления векторного потенциала

могут быть восприняты как воздействия неэлектромагнитной природы. Так, например, поле векторного потенциала не экранируется и может, поэтому, приниматься за проявление гравитационного воздействия, тем более, что в случае торроидальных магнитных полей, обнаружить порождающее векторный потенциал магнитное поле, зачастую, не представляется возможным. Поле векторного потенциала может служить своеобразным “предвестником” скрытых перестроек магнитного поля и в этом случае возможна опережающая, по отношению к магнитному полю, реакция системы, которая напоминает эффект Чижевского-Вельхова, и т.д. Кратко рассмотрим различные точки зрения на эффект Ааронова-Бома.

В работе [76] Я. Ааронов и Д. Бом высказали предположение, что в квантовой механике, в противоположность классической электродинамике, потенциалы электромагнитного поля должны рассматриваться, как первичные физические величины, а напряженности полей, как вторичные, производные понятия. Эффект Ааронова-Бома характеризует влияние внешнего электромагнитного поля, сосредоточенного в области, недоступной для заряженной частицы, на ее квантовое состояние. Такое, исчезающее в классическом пределе, нелокальное воздействие электромагнитного поля на заряженную частицу, показывает, что при квантовомеханическом рассмотрении, оно не сводится к локальному воздействию на нее силы Лоренца [75]. Отмечается [78], что в эффекте Ааронова-Бома частицы демонстрируют присущие квантовым объектам черты специфической целостности, но рассматриваемые в контексте классической физики могут порождать иллюзию своеобразного “действия на расстоянии”: поле действует там, где его нет. “Основой такой иллюзии служит классическое представление о мире, как совокупности локализованных и потому, относительно независимых друг от друга объектов. При этом всякая корреляция в их поведении объясняется, в конечном счете, взаимодействием, т.е. обменом энергией-импульсом” [78, с. 613]. При квантовомеханическом же подходе мир, скорее, рассматривается как целое и для того, чтобы частицы рассматривать как отдельные, необходимо создать специальные экспериментальные условия. После рассмотрения ряда квантовомеханических экспериментов, автор работы [77], также полагает, что эффект Ааронова-Бома обусловлен квантовыми особенностями поведения частицы, и никаких новых свойств потенциала, которых не было бы в классической электродинамике, не возникает.

Приведенные работы показывают, что эффект Ааронова-Бома, как правило, рассматривается как чисто квантовый. В то же время, в работе [79], К.А. Труханов сформировал гипотезу о физической реальности электромагнитных потенциалов и рассмотрел возможности исследования векторного потенциала при помощи классических систем. Им также рассматривается возможность исследования поля векторного потенциала при помощи биологических систем [80-81].

Основная трудность принятия физической реальности векторного потенциала заключается в следующем. Принимая реальность векторного потенциала, мы оказываемся в ситуации, когда одному и тому же значению магнитного поля могут соответствовать различные значения потенциала определенные с точностью до градиента произвольной функции. Поэтому, существующие электродинамические теории калибровочно инвариантны.

Бауров Ю.А., в развитой им теории, предположил локальное нарушение калибровочной инвариантности и ввел одномерные дискретные магнитные потоки, которые, впоследствии, назвал “бюонами”. В определение бюонов входит новая векторная константа - космологический векторный потенциал A_r , равный $A_r \approx 1.95 \cdot 10^{11}$ Гс·см. Направление вектора A_r

определено экспериментально и имеет координаты: прямое восхождение $\alpha \approx 270^\circ$ и склонение $\delta \approx 34^\circ$ (вторая экваториальная система координат) [82]. В более поздней работе [84] приводятся уточненные значения прямого восхождения α и склонения δ , равные $\alpha \approx (280-297)^\circ$ и $\delta \approx (30-40)^\circ$. Отмечается, что это направление близко к направлению движения Солнца относительно ближайших звезд ($\alpha = 270^\circ$ и $\delta \approx 30^\circ$).

В концепции Ю.А. Баурова как векторный, так и скалярный потенциалы электромагнитного поля обладают физическим смыслом и измеримы для практически любых расстояний от 10^{-17} до 10^{28} см [82].

Ю.А. Бауров полагает, что наблюдаемое трехмерное пространство и элементарные частицы возникают в результате минимизации потенциальной энергии взаимодействия бюонов. Одновременно возникают галактические и межгалактические магнитные поля и реликтовое излучение. Из развитых им представлений о физическом пространстве вытекает, что космологический векторный потенциал является причиной глобальной анизотропии Вселенной [83].

Векторные потенциалы всех космических источников совместно с A_r формируют в каждой определенной области пространства некоторый суммарный потенциал A_Σ , который всегда меньше A_r . Предсказывается существование нового взаимодействия в природе, возникающего при изменении A_Σ [82]. Отсюда утверждается о существовании квантового информационного канала, “связывающего все объекты живой и неживой природы в одно общее информационное поле” [84]. Утверждается также принципиальная возможность изменений фундаментальных масштабов пространства через суммарный векторный потенциал [84].

Многочисленные экспериментальные исследования, выполненные Ю.А. Бауровым, не дают возможности их полного и детального рассмотрения в настоящем обзоре. Поэтому, в качестве примера, мы приведем лишь опыты с кварцевыми резонаторами.

Теоретическая возможность изменения фундаментальных масштабов пространства дала идею этих экспериментов [84, 85], позволяющих, по мнению, Ю.А. Баурова возможность исследования A_r . Используемая в этих опытах регистрирующая система состояла из двух кварцевых резонаторов, отличавшихся тем, что один из них был помещен в специальную магнитную систему собранную из постоянных магнитов и создающую поле в магнитопроводе ~ 3500 Гс. Измеряемым параметром в этих опытах является время изменения фазы между колебаниями двух резонаторов на 2π . Примеры экспериментальных записей, полученных в экспериментах с кварцевыми резонаторами, представлены на рис. 20.

Основной результат [84] – обнаружение околосуточных периодов 20 ч, 22.5 ч и 22.17 ч. Наличие этих периодов полагается следствием действия новой силы, обусловленной изменением вектора A_Σ и возникающей при вращении кварцевых резонаторов вместе с Землей за счет воздействия векторного потенциала Земли. Регистрирующая система реагирует на это изменение из-за градиента $\partial A_\Sigma / \partial x$, созданного магнитной системой одного из кварцевых резонаторов [84]. Используемая магнитная система, по мнению авторов, служит своеобразным усилителем сигналов – изменений A_Σ . Отмечается, что обнаруженные периоды не возникают, в случае, если используются два резонатора без магнитного поля. Анализ полученного экспериментального материала позволяет выделить три множества направлений. С одним из них связывается направление действия A_r , а с двумя другими – “другие эффективные направления действия новой силы” [84, с. 826].

Автор, в своих теоретических оценках действия новой силы, рассматривает разность ча-

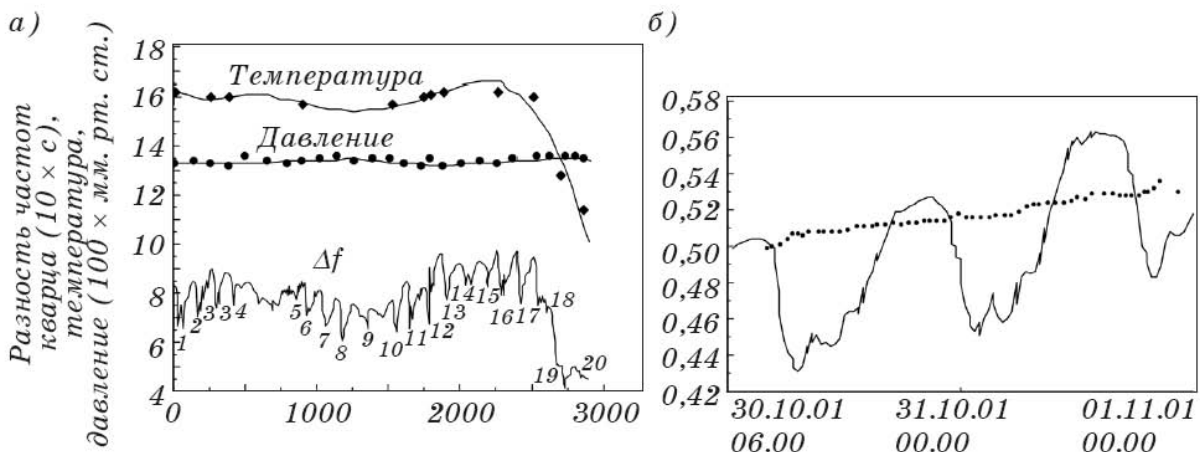


Рис. 20 Пример экспериментальной записи разности фаз для двух кварцевых резонаторов, полученной за период с 1.11.2000 г. по 26.11.2000 г., а); то же для периода в два с половиной дня (6:00-30.10.2001 – 2:00-1.11.2001). Пунктирная кривая на рисунке – результаты одночасовых измерений температуры, б). Рисунок а) взят из [84], б) из [85].

стот кварцевых резонаторов Δf , но при этом, в эксперименте, измеряется время изменения фазы на 2π . Исходя из данного в [84, 85] описания схемы эксперимента, данный параметр не связан однозначно с Δf . Т.к., определение направлений, основано на анализе Δf , то выводы, представленные в [84, 85] нужно, на наш взгляд, принимать с осторожностью. Также представляется недостаточным анализ влияния температуры.

Ю.А. Бауров в [84] отмечает, что в более ранних работах Н.В. Клочка также наблюдался суточный период. Но, необходимо отметить, что в работах Н.В. Клочка этот период возникал за счет различной пространственной ориентации кварцевых резонаторов. В упомянутой работе Ю.А. Баурова отмечается только ориентация кварцевого резонатора, помещенного в магнитное поле. Т.к., о взаимной пространственной ориентации кристаллов кварца в [84] не говорится, то это заставляет предположить, что данный фактор в работе не учитывался. Следовательно, кварцевые резонаторы были, скорее всего, ориентированы по разному, что, как было показано, в рассмотренных выше работах Н.В. Клочка, может приводить к возникновению суточного периода. В то же время утверждение Ю.А. Баурова о том, что околосуточные периоды не возникают, если один из резонаторов не поместить в магнитное поле, находится в противоречии с результатами Н.В. Клочка.

1.8 А.Г. Пархомов: исследование ритмов и флуктуаций в ходе различных процессов.

Основное направление работ А.Г. Пархомова - исследование ритмов и флуктуаций в ходе различных процессов (низкочастотный шум полупроводниковых приборов, генерация колебаний устройствами с кварцевыми резонаторами, альфа и бета распады) в сочетании с анализом хода температуры и радиационного фона. В основном, эти работы выполнены при помощи экспериментальной установки, характерной особенностью которой является большое число одновременно исследуемых процессов (запись информации производится по 20 каналам), а также возможность многолетней непрерывной регистрации сигналов в сочетании с высокой точностью определения времени событий. Особое внимание в установке уделяется точности

измерений времени. Для этого термостатирован кварцевый резонатор часов компьютера, а также на один из входов регистрирующего компьютера каждый час подаются радиосигналы точного времени. Это позволяет при обработке результатов вносить поправки, снижающие погрешность определения абсолютного времени до 0.05 сек. Для минимизации температурных влияний большинство датчиков с источниками питания помещены в шкаф-термостат, где поддерживается температура $31 \pm 0,3$ С, или имеют индивидуальный термостат. Кроме того, для контроля возможных температурных влияний производится запись температуры около установки и в шкафу-термостате [87].

Кратко рассмотрим исследуемые А.Г. Пархомовым процессы и основные, полученные при этом, результаты. Также как и раньше мы, по возможности, будем приводить характерные формы регистрируемых сигналов.

1.8.1 Генераторы с кварцевыми резонаторами.

В ходе измерений использовалась разностная схема с двумя генераторами с резонаторами, имеющими близкие резонансные частоты и температурные коэффициенты. Генераторы были помещены в выравнивающую температуру оболочку. Все устройство было помещено в шкаф-термостат.

Зарегистрированные вариации разности частот имеют амплитуду порядка 0.1 Гц (что составляет 10^{-8} от резонансных частот) и в общих чертах воспроизводят ход температуры в термостате, рис. 21. Величина вариаций примерно соответствует ее оценке по величине предварительно измеренного скомпенсированного температурного коэффициента (около 1 Гц/С в обоих устройствах).

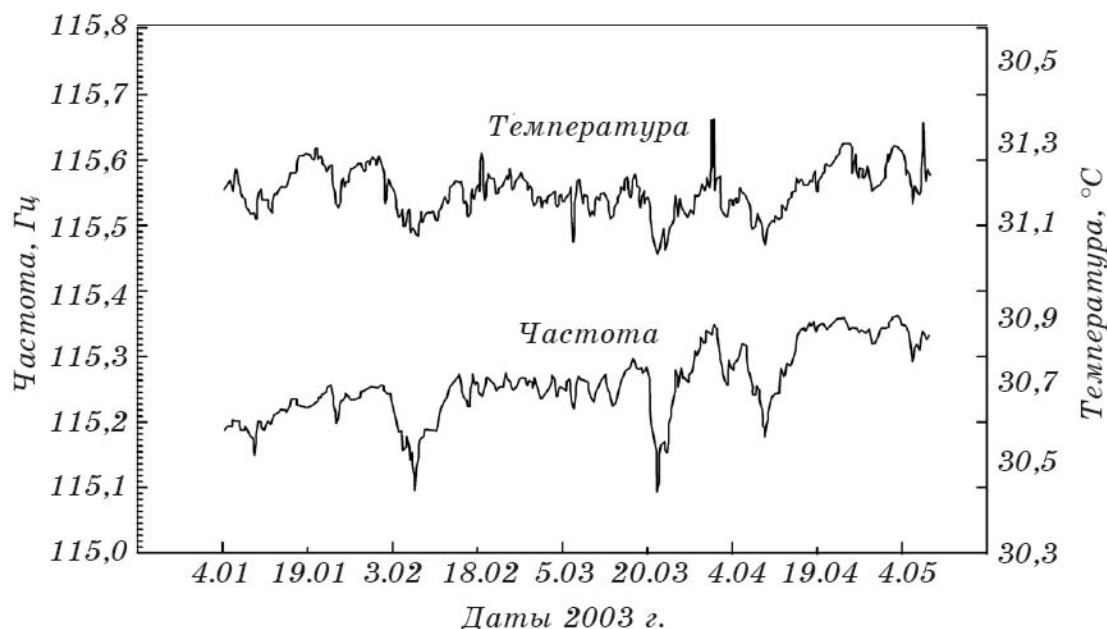


Рис. 21 Ход разности частот двух кварцевых резонаторов, сопоставленный с ходом температуры. [87].

Отмечается, что кварцевые резонаторы использовались для проверки идеи Н.А.Козырева, о том, что процессы с сильным изменением энтропии должны влиять на ход времени и,

следовательно, на частоту используемых кварцевых генераторов. Были проведены специальные эксперименты, в ходе которых процессы со значительными энтропийными изменениями (растворение сахара и плавление льда) происходили около устройств с кварцевыми резонаторами. Изменений в ходе сигналов, кроме тех, которые можно объяснить непостоянством температуры, не обнаружено.

1.8.2 Инфранизкочастотные флуктуации в проводниках.

В 1983-1993 гг. исследование инфранизкочастотного шума в полупроводниках осуществлялось с применением многоканального самописца. Были исследованы фоторезисторы (темновой ток), термостатированные терморезисторы и минерал гематит, обладающий полупроводниковыми свойствами. Помимо этого, в качестве генераторов шума применяли биполярные и МОП-транзисторы, а также микросхемы [88]. В последние годы источниками шума служат кремниевые транзисторы П701А. Шумовой сигнал после усиления операционным усилителем поступает на выпрямитель и фильтр, пропускающий частоты от 0,1 до 1 Гц. Этот сигнал регулирует длительность межимпульсных интервалов, поступающих на входы компьютерного регистратора. Для преобразования амплитуды в длительность используется однопереходный транзистор. Источники шума с усилителями и преобразователями, смонтированные в экранирующей металлической оболочке, размещены в шкафу-термостате. Пример регистрации сигнала приведен на рис. 22. На этом рисунке видны отчетливые пики, которые А.Г. Пархомов называет “события резкого возрастания (СРВ)” и определяет как увеличение амплитуды по сравнению с усредненной величиной трех предшествующих измерений более чем в 1,5 раза.

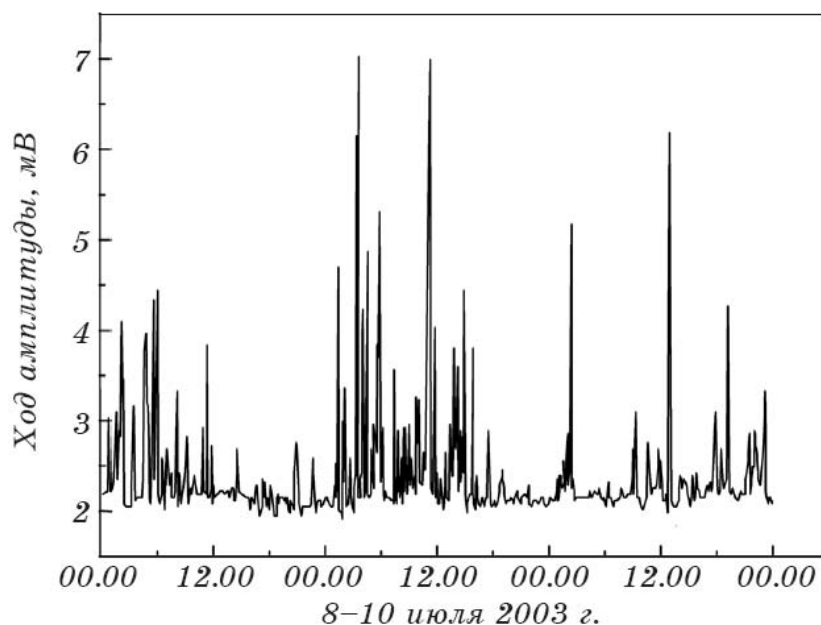


Рис. 22 Фрагмент записи хода амплитуды инфранизкочастотного шума транзистора П701А. [87].

Исследование четырехлетних (2000-2004 гг.) временных рядов методом наложенных эпох показало следующие результаты:

©В.А. Панчелюга Экспериментальные исследования и астрофизические наблюдения, свидетельствующие о неоднородности и анизотропии реального пространства-времени.

1. Устойчивых ритмических изменений вероятности СРВ с периодами солнечных, звездных и лунных суток не обнаружено.
2. Обнаруживается связь вероятности СРВ с изменением солнечной активности в 27-суточном цикле (проанализировано 56 циклов). Видно, что снижение вероятности СРВ происходит как около минимумов, так и около максимумов (10-15 суток до и после минимумов). Наибольшая вероятность СРВ – во время роста и снижения солнечной активности.
3. Ритм синодического лунного месяца на протяжении 49 циклов проявляется отчетливо, но иначе, чем в температурных изменениях: вероятность СРВ около полнолуний примерно в 2 раз выше.
4. Ритм сидерического лунного месяца малозаметен [87].

Результаты приведенных исследований низкочастотного шума транзисторов вполне подтверждают результаты исследований в 1983-1993 гг. [88, 89], проведенных с иными полупроводниковыми источниками шума и другой регистрирующей аппаратурой.

А.Г. Пархомов отмечает, что “поведение” низкочастотного шума в полупроводниках во многом напоминает динамику землетрясений [90]. И там, и здесь проявляются ритмы лунного месяца и солнечной активности, и там и здесь особенно сильные всплески активности происходят после длительных затиший. Но если влияние взаимного положения Земли, Луны и Солнца на ход тектонических процессов можно объяснить действием приливных сил, объяснение лунно-солнечной ритмики в процессах, текущих в полупроводниках, требует иных подходов [91].

1.8.3 Скорость счета α -частиц.

Говоря о скорости счета α -частиц, А.Г. Пархомов отмечает, наличие выбросов, обусловленных радиационными эффектами полупроводниковых детекторов. На рис. 23 приведен пример регистрации скорости счета α -частиц испускаемых ^{239}Pu поверхностно-барьерным полупроводниковым детектором. Видно, что участки с “нормальным” счетом продолжительностью до нескольких месяцев чередуются с участками, где скорость счета резко увеличена. Анализ распределения такого рода всплесков во времени выявляет картину, аналогичную полученной при исследовании низкочастотных флуктуаций в полупроводниках. Иначе говоря, полупроводниковый детектор выступает в роли своеобразного генератора фликкер-шума. В более устойчивых к радиационным воздействиям полупроводниковых детекторах фликкер-шум может проявляться не столь ярко, рис. 24, но полностью исключить его влияние, по мнению А.Г. Пархомова невозможно. Так как фликкер-шум является процессом, чувствительным к различным воздействиям, в том числе воздействиям космической природы, обнаруженные вариации скорости счета в устройствах с ППД нельзя с уверенностью трактовать как вариации скорости радиоактивного распада. Это может быть результатом трудно замечаемого проявления фликкер-шума [87].

Применение к полученным рядам флуктуаций скорости α -распада метода наложенных эпох с использованием в качестве реперов новолуний и минимумов солнечной активности в 27-суточном цикле не выявило достоверных ритмов, превышающих несколько тысячных долей процента. Отмечается, что в измерениях скорости счета альфа-частиц полупроводниковым детектором заметна ритмика с амплитудой не более 0,01% от средней скорости

счета. Возможно, что появление этой ритмики связано не с изменением радиоактивности, а с фликкер-шумом детектора, температурными воздействиями или иными причинами нестациональности регистрирующей аппаратуры. Поэтому приведенную цифру следует считать верхней границей возможных ритмических изменений α -радиоактивности [87].

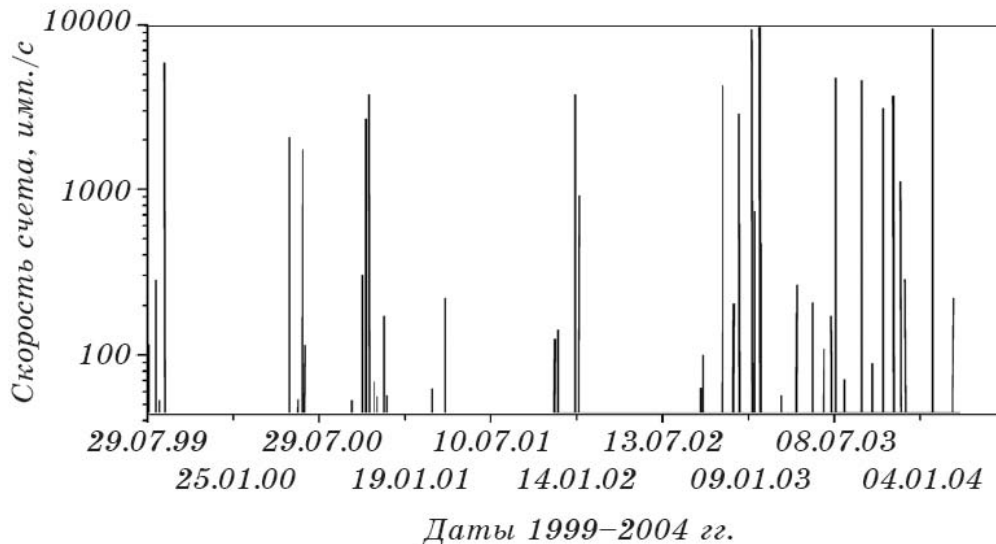


Рис. 23 Всплески скорости счета. Полупроводниковый детектор с источником α -частиц ^{239}Pu . [87].

1.8.4 Скорость счета β -частиц.

Для регистрации β -частиц и сопровождающих β -распады гамма квантов А.Г. Пархомов использовал счетчики Гейгера. Измерения проводились двумя регистраторами использующими в качестве источников β -распада ^{90}Sr - ^{90}Y и ^{60}Co . Анализ полученных временных рядов выявил синхронный и синфазный ритм с периодом в 1 год и амплитудой около 0,3% от средней скорости счета. Хотя ход температуры имеет такой же основной период, однако сдвиг по фазе на 2-3 месяца дает основание А.Г. Пархомову предполагать не термическую причину изменений скорости счета с периодом в год. Также обнаружены месячные ритмы изменений скорости счета и температуры. Но, на основании того, что обе ритмики имеют существенно различный характер, также делается заключение об нетермической причине месячной ритмики скорости счета при β -распаде. Вопрос о том, является ли суточный ритм скорости счета проявлением температурных вариаций или имеет иное происхождение, остается открытым [87].

Наиболее неординарный результат в исследованиях А.Г. Пархомова дал эксперимент с источником ^{60}Co в сочетании со счетчиком Гейгера, расположенным в фокусе параболического зеркала. Этот своеобразный телескоп, за счет суточного вращения Земли более двух лет круглосуточно сканировал небесную сферу. Обычный ход измерений скорости счета, вполне соответствующий статистике Пуассона, время от времени нарушался всплесками протяженностью до нескольких часов, когда скорость счета высокодостоверно возрастала. Эти аномальные участки занимают примерно 1/1000 всего времени наблюдения. Примеры записей таких всплесков показаны на рис. 25. Всплески происходили преимущественно тогда, когда

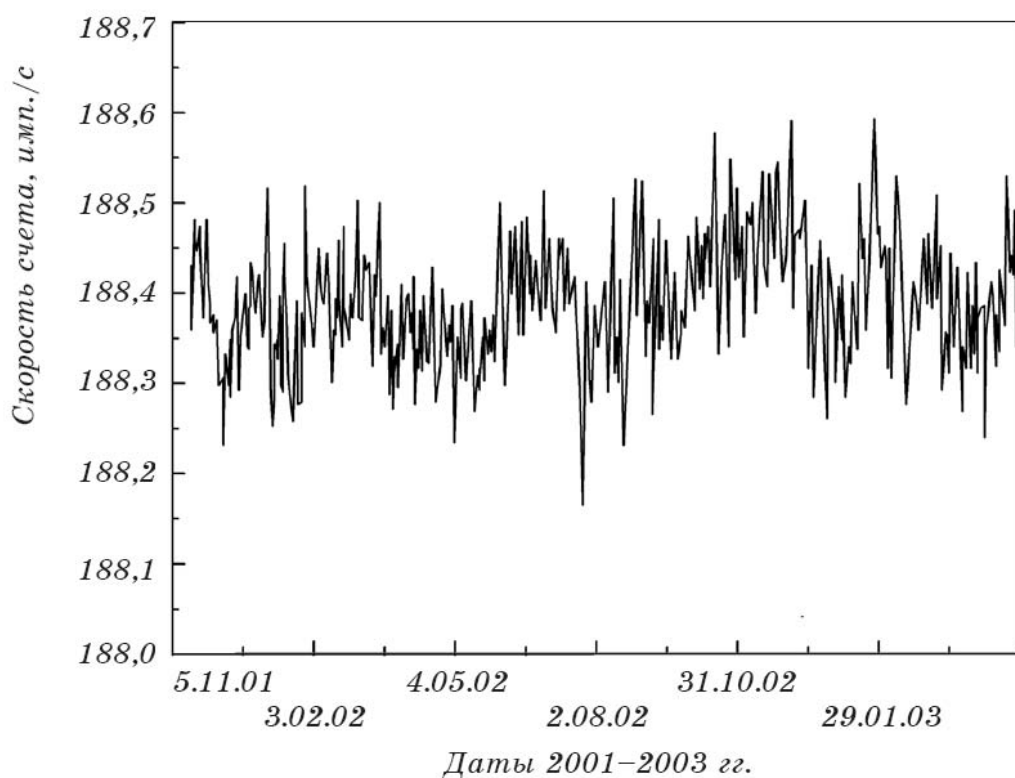


Рис. 24 Ход скорости счета α-частиц источника ^{239}Pu радиационно-стойким полупроводниковым детектором-фотодиодом с ноября 2001 г. по июнь 2003 г. Усреднение 24 часа. [87].

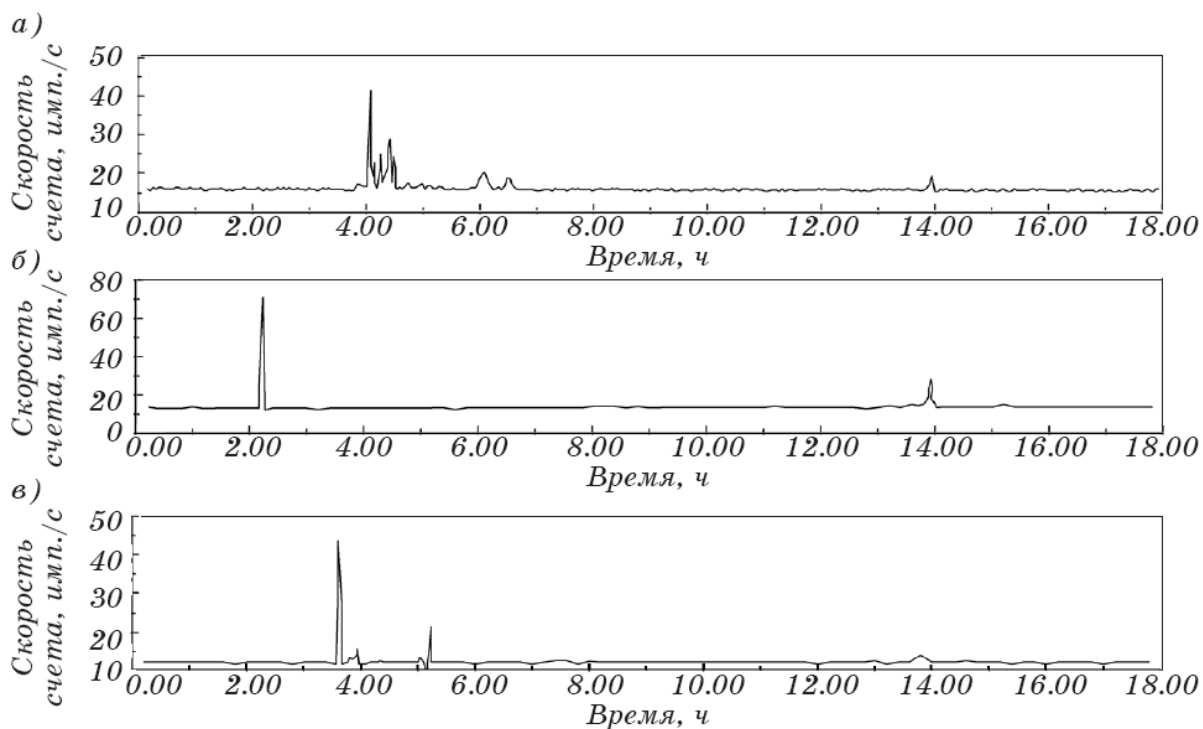


Рис. 25 Примеры зарегистрированных всплесков скорости счета ^{60}Co в фокусе телескопа-рефлектора при одномерном сканировании небесной сферы. [87].

©В.А. Панчелюга Экспериментальные исследования и астрофизические наблюдения, свидетельствующие о неоднородности и анизотропии реального пространства-времени.

телескоп был отклонен от Солнца на угловое расстояние до 30-40 градусов, хотя зарегистрированы всплески и в другое время.

Во втором канале, где использовался такой же счетчик, подключенный к тому же источнику питания, всплески отсутствовали. Всплесков не было, когда этот же регистратор на протяжении многих месяцев работал вне телескопа. На основании этого А.Г. Пархомов делает вывод о том, что возникновение всплесков связано именно с наличием фокусирующего зеркала, концентрирующего потоки некоторого идущего из небесной сферы агента, способного зеркально отражаться и влиять на скорость счета радиоактивного источника.

Во втором варианте этого опыта телескопу было придано колебательное движение перпендикулярно линии сканирования, связанной с суточным вращением Земли. Такое сканирование осуществлялось в течение 49 суток с апреля по сентябрь 2004 г. При такой модификации опыта число регистрируемых всплесков резко повысилась.

Всего было зарегистрировано 280 всплесков, превышающих среднюю скорость счета более чем на 5 стандартных отклонений (средняя скорость счета 5,1 имп/с, определение скорости – по времени набора 256 импульсов). Телескоп был направлен на восток. Число всплесков в сутки и их положение на изображениях не воспроизводятся. Максимальная вероятность всплесков с 8 до 9 и с 19 до 21 час московского времени, минимальная – с 3 до 7 час. По звездному времени резкое возрастание вероятности всплесков зарегистрировано с 17 до 18 час. Вероятность появления всплесков очень мала, когда небо покрыто облаками.

Как отмечает А.Г. Пархомов, эксперименты с фокусирующим зеркалом были проведены целенаправленно для проверки идеи о возможной роли потоков нейтрино очень низких энергий как одного из факторов космофизических взаимодействий [91-93]. По мнению автора, полученные результаты вполне подтверждают предсказания, сделанные в этих работах.

2 Феномен макроскопических флуктуаций и его основные свойства.

Особенности, регистрируемые в рассмотренных выше работах, можно объединить в следующие группы:

- Сигналы – повторные реализации участков временного ряда, форма которых подобна. Необходимо отметить, что речь идет именно о подобии формы, т.к., наряду с внешним воздействием, определяющим форму участка временного ряда (сигнала), всегда присутствует компонента шума, обусловленная внутренними свойствами системы. [13]
- Выбросы – участки временного ряда, для которых величины измеряемых значений существенно превосходят величину дисперсии временного ряда.
- Затухания – участки временного ряда, для которых величина дисперсии σ значительно меньше величины дисперсии всего временного ряда, Σ .
- Флуктуации измеряемых значений, длительность которых существенно превосходит среднюю длительность флуктуаций для данного временного ряда. Под длительностью флуктуаций понимается промежуток времени в течение которого величина $x - \bar{X}$ сохраняет свой знак. Здесь под x подразумевается измеряемая величина, $\bar{X}$ – величина среднего для рассматриваемого временного ряда.
- Изменения среднего: участки временного ряда для которых величина среднего $\bar{x}$ значительно отличается от $\bar{X}$.

Как следует из приведенной классификации, рассмотренные нами работы, объединяет то обстоятельство, что основным анализируемым параметром в них является изменение мгновенных значений относительно среднего. Для обнаружения перечисленных выше особенностей, расположенных нерегулярным образом во временных рядах флуктуаций, они должны иметь достаточную амплитуду. Т.к. означенные особенности обнаруживаются во временных рядах, то порядок измеренных величин во временном ряду не должен быть нарушен.

В этом основное отличие вышерассмотренных работ от исследований нашей лаборатории, где результаты измерений, могут составлять совершенно случайный, с точки зрения существующих на сегодняшний день статистических тестов, временной ряд. Причиной этому является то обстоятельство, что объектом нашего анализа является подобие тонкой структуры пар гистограмм. Исходным материалом для подобного анализа служат последовательности гистограмм, построенных на основе исходных временных рядов. При этом существует взаимно-однозначное соответствие между отрезком временного ряда и получаемой на его основе гистограммой. В то же время, обратный переход от гистограммы к отрезку временного ряда – неоднозначен и может быть произведен $n!$ способами, где n длина отрезка исходного временного ряда, на основе которого строится гистограмма. Так для наиболее часто используемых значений $n = 60$, $n! = 10^{82}$. Т.е. после того, как измерения произведены, любая из $n! = 10^{82}$ перестановок не влияет на форму результирующей гистограммы, которая может рассматриваться как спектр амплитуд по отношению к исходному отрезку временного ряда. Важны только результаты произведенных измерений. В этом также существенное отличие метода макроскопических флуктуаций от рассмотренных выше работ, где последовательность измерений во временном ряду имеет большое значение.

Рассмотрим более детально методику обработки экспериментальных данных, используемую в методе макроскопических флуктуаций, а также краткую историю метода.

2.1 Феномен макроскопических флуктуаций. Краткая история.

Результаты любых многократных, последовательных, высокостандартизованных измерений параметров различных природных процессов сопровождаются неустранимым “разбросом результатов”. Обычно, этот разброс рассматривают как помеху, затрудняющую получение точных значений измеряемых величин и от него стараются избавиться. Существуют общепринятые методы уменьшения “вредного влияния” разброса результатов на достоверность умозаключений, следующих из этих измерений. В то же время, разброс или, по другому, временные ряды флуктуаций в ходе этих процессов, как оказалось, имеют самостоятельную ценность и могут служить источником информации о наиболее глубоких и фундаментальных свойствах нашего мира – о его пространственно-временной структуре. Исследования, позволившие сформулировать подобное утверждение были начаты более 50-и лет тому назад (в 1951-56 г.г.), когда один из авторов этой статьи, С.Э.Шноль, начал систематические исследования причин флуктуаций результатов измерений скоростей биохимических реакций. Поводом для начала этих исследований была излишне, как казалось, большая амплитуда этих флуктуаций. Однако, самым ярким их свойством были сильно изрезанные гистограммы – спектры амплитуд флуктуаций измеряемых величин. В ряде опытов эти изрезанные гистограммы получались удивительно похожими друг на друга. Это явление было названо феноменом макроскопических флуктуаций [101].

Сначала, эта странная похожесть объяснялись следствием особых свойств изучаемых объ-

ектов – сократительных белков мышц [102-103]. Но потом, годы спустя, аналогичные странности были обнаружены при измерениях любых процессов от скоростей химических реакций, до измерений радиоактивного распада.

Изначально феномен макроскопических флуктуаций пытались объяснить внутренними свойствами процесса, флуктуации которого изучались. Эти попытки привели к идее о колебательных режимах в ходе изучаемых химических реакций. Одним из итогов этого периода было исследование аспирантом С.Э. Шноля А.М. Жаботинским, колебательной реакции, открытой Б.П. Белоусовым [104]. Впоследствии, она получила всемирную известность, как реакция Белоусова-Жаботинского или BZ-реакция [105].

По-видимому, впервые мысль о том, что феномен может быть связан не с внутренними свойствами системы возникла в связи с опытами 60-70-ых годов, в которых была обнаружена синхронность колебаний (флуктуаций) ферментативной активности белков актомиозинового комплекса и креатинкиназы в *разных сосудах* [106-111]. Из этих опытов следовало, что “конформационные колебания” зависят от какой-то “внешней силы”. Но тем не менее, в то время причина наблюдаемых явлений искалась в непосредственном взаимодействии молекул белков [103].

Стимулом для постановки первых опытов по измерениям в разных географических пунктах явилось предложение В.Е. Жвирблиса провести с ним синхронные измерения: флуктуаций при настройке визуального поляриметра в Москве и серия измерений флуктуаций ферментативной, креатинкиназной активности в Пущино. Результаты этих опытов в апреле – мае 1978 г. показали сходство гистограмм в одно и то же время при измерении разных процессов и при расстоянии между лабораториями более 100 км [112]. Результаты этих опытов были восприняты “скептически” и, поэтому, даже не были полностью проанализированы. Следующий шаг был сделан осенью 1979 г. – были произведены измерения радиоактивности в Москве – в изотопной лаборатории корпуса А МГУ (впоследствии ин-та им. А.Н. Белозерского) и скорости химической реакции (аскорбиновой кислоты с дихлорфенолиндофенолом) – в Пущино. Также было обнаружено сходство гистограмм. Но, несмотря на это, до декабря 1980 г. эти опыты больше не ставили, ввиду “смущения”. Систематическая работа началась лишь в феврале 1980 г. – одновременные измерения разных процессов в Пущино и измерения α -активности в МИФИ (Н.Б. Хохлов, М.П. Шарапов). Эти синхронные опыты (около 30-и) продолжались до июня и показали высокую вероятность сходства гистограмм в процессах разной природы в одно и то же время в этих опытах [3, 15, 112-113].

Поправка на разницу местного времени первый раз бы сделана при сравнении гистограмм при измерениях в Пущино и в Ленинграде (А.Сунгуров, ^3H -бета-активность) в 1984 году [2, 113]. Однако, более тщательно “эффект местного времени” был исследован только при переходе от часовых гистограмм к одноминутным. Это произошло после создания И.А. Рубинштейном высокостабильных устройств для измерений α -активности в 2000 г. В процессе сравнения гистограмм, полученных при измерениях на корабле в Арктической экспедиции (С.Н. Шаповалов) и в Пущино была обнаружена высокая синхронность по местному времени при расстояниям между “лабораториями” в несколько тысяч километров и разности местного времени 9 часов [8]. Затем этот эффект был показан при измерениях в Антарктике и Пущино. В Германии (Линдау) и Пущино. В Германии (район Дюссельдорфа) и в Москве, В США (г. Каламбус) и в Пущино, В Испании (г. Валенсия) и в Пущино [3, 6, 114-115].

Детальное исследование этого феномена макроскопических флуктуаций продолжается

до сих пор. Тонкая структура, форма гистограмм не зависит от природы процесса, а только от места и от времени проведения измерений. Относительная величина необъясняемой тривиальными причинами амплитуды флуктуаций (величина разброса результатов) различна для процессов разной природы. Эта амплитуда максимальна при измерениях биологических процессов и может составлять 50 – 100% от измеряемой величины. В химических измерениях амплитуда меньше и составляет обычно 1-5% от средней величины. Для радиоактивного распада, согласно статистике Пуассона, средняя амплитуда флуктуаций равна $\sqrt{N}$, где N – среднее число распадов. Амплитуда флуктуаций частоты, например, кварцевого резонатора равна всего 10^{-6} от средней величины, в квантовых генераторах эта величина имеет порядок $10^{-9} - 10^{-10}$. Причины этих характерных различий имеют глубокий физический смысл. Однако, как показано в ходе многочисленных экспериментов, при соответствующей нормировке, детальная форма гистограмм оказывается сходной для разных процессов. Для примера приведем список (не полный) процессов, для которых это утверждение было проверено:

- измерения радиоактивности всех основных видов: α -, β -, γ -, К-захвата, при самых разных методах измерений – счетчиками Гейгера, сцинтилляционными счетчиками, полупроводниковыми детекторами;
- при измерениях движения частиц латекса в электрическом поле;
- при измерениях времени релаксации t_2 протонов воды в переменном магнитном поле;
- при измерениях времени ожидания разряда неоновой лампы;
- при измерениях параметров колебаний в колебательной химической реакции Б.П. Белоусова;
- при измерениях флуктуаций темнового тока в фотоумножителях;
- при измерениях шумов в гравигradientной антенне “Улитка”;
- при измерениях потока нейтронов из земной коры;
- при измерениях шумов в полупроводниковых шумовых генераторах. [116]

Накопленный нами опыт позволяет сделать утверждение о том, что детальная форма тонкой структуры гистограмм не зависит от природы процесса, флуктуации которого исследуются и в каждый момент в данном географическом пункте подобна для процессов различной природы.

2.2 Исходный экспериментальный материал для исследования феномена макроскопических флуктуаций.

Для уяснения сути феномена макроскопических флуктуаций рассмотрим следующий пример [121-123]. Предположим, у нас есть простейшая электрическая цепь, состоящая из сопротивления и источника напряжения. Пусть любые внешние воздействия на рассматриваемую электрическую цепь полностью исключены. Ток в такой цепи с высокой точностью будет равен некоторой постоянной величине. Также предположим, что мы последовательно проводим измерения тока в цепи, причем, каждый раз со все большей чувствительностью. Тогда, при достижении определенной чувствительности, мы заметим, что величина тока, вначале казавшаяся нам постоянной, в действительности подвержена некоторым флуктуациям. Делая

последовательные, однотипные измерения, мы получим временной ряд величин флуктуаций тока в цепи. Такие временные ряды величин флуктуаций в протекании различных природных процессов и являются исходным экспериментальным материалом для исследования феномена макроскопических флуктуаций.

Очевидно, что используя соответствующую измерительную процедуру, возможно зарегистрировать флуктуации, которые, как известно, сопровождают протекание любого природного процесса. Особо отметим, что используются временные ряды флуктуаций именно *природных* процессов. Их получение, в большинстве случаев, достаточно сложная задача, требующая, с одной стороны, построения высокочувствительных, широкополосных измерительных систем, а с другой - принятия специальных мер по устранению внешних воздействий на исследуемый флуктуационный процесс.

Как известно, наиболее общими причинами флуктуаций являются дискретность материи, тепловое движение, случайный характер микроскопических процессов, лежащих в основе исследуемого макроскопического процесса. Все это является причиной того, что временные ряды флуктуаций являются абсолютно случайными, с точки зрения стандартных статистических методов. Поэтому считается, что ни отдельные измерения, ни короткие (десятки точек) отрезки временного ряда флуктуаций не позволяют извлечь к.-л. достоверную информацию. Для этого нужны достаточно длинные (тысячи и более точек) отрезки, дающие возможность точного определения статистических параметров. Оригинальная методика, разработанная в ходе исследований феномена макроскопических флуктуаций, принципиально использует короткие отрезки случайных временных рядов. Она позволяет получать информацию не обнаружимую обычными методами статистического анализа и находить закономерности в, казалось бы, совершенно случайных временных рядах.

2.3 Методика обработки экспериментальных данных.

Рассмотрим методику обработки экспериментальных данных, на которой основано обнаружение основных эффектов феномена макроскопических флуктуаций.

Эту методику условно можно подразделить на два этапа. Первый этап иллюстрирует рис. 26. Здесь, рис. 26 А представляет исходный ряд величин флуктуаций некоторого процесса. Этот ряд, разбивается на короткие отрезки, обычно 30-100 точек. На рис. 26 В, в качестве примера, приведены четыре таких 100-точечных отрезка. Каждый из этих отрезков является исходным материалом для построения гистограммы. На рис. 26 С приведены четыре гистограммы, построенные на основе отрезков, показанных на рис. 26 В. После этого, каждая гистограмма сглаживается несколько раз k -точечным прямоугольным окном. Результаты сглаживания гистограмм, представленных на рис. 26 С показаны на рис. 26 D. Число сглаживаний и ширина окна зависят от свойств исходного ряда и, в первую очередь, от величины его дисперсии. Чаще всего используется величина $k = 4$ и число сглаживаний, лежащее в интервале $10 \div 20$ раз. В результате применения описанной выше процедуры, исходный ряд флуктуаций, рис. 26 А, преобразуется в последовательность гистограмм, рис. 26 D, которая является основным объектом дальнейшего анализа.

Необходимо отметить, что под “гистограммой”, в первую очередь, подразумевается алгоритм определенного упорядочения некоторого короткого отрезка экспериментальных данных, как это показано на рис. 26 С. Подобный алгоритм совершает однозначное отображение n -точечного отрезка временного ряда в соответствующую ему гистограмму. Но, как уже от-

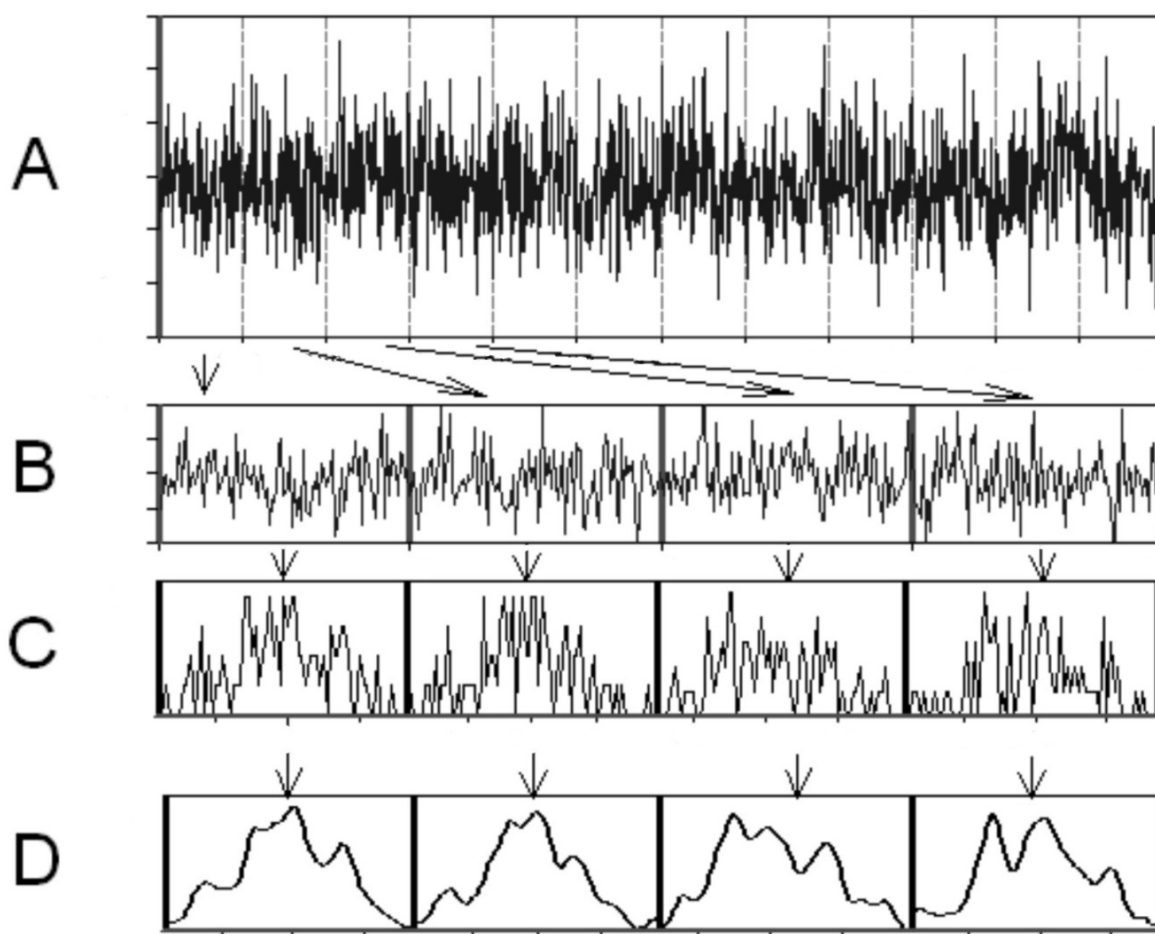


Рис. 26 Методика построения гистограмм.

мечалось, обратный переход от гистограммы к исходному отрезку, неоднозначен, т.к. форма гистограммы не чувствительна к произвольным перестановкам внутри лежащего в ее основе временного ряда и, поэтому, является, в некотором роде, уникальным инструментом для обнаружения скрытого в отрезке временного ряда порядка.

При этом, изучаемые нами гистограммы являются, по сути, спектрами амплитуд флуктуаций. Иногда также используется термин “распределение амплитуд флуктуаций”. Хотя, в строгом статистическом смысле для используемых коротких отрезков временных рядов функция распределения не существует.

Итак, основным объектом нашего анализа является форма тонкой структуры гистограмм, построенных по коротким, непересекающимся отрезкам временных рядов. Ниже приведен алгоритм, иллюстрируемый рис. 27 и используемый для исследования подобия формы гистограмм во временных рядах данных.

2.4 Построение распределения интервалов.

На рис. 27 А дан пример последовательности из $N = 20$ гистограмм, являющейся исходным материалом для процесса экспертного сравнения. Данная последовательность, являясь результатом первого этапа, тождественна последовательности на рис. 26 D. Каждая гисто-

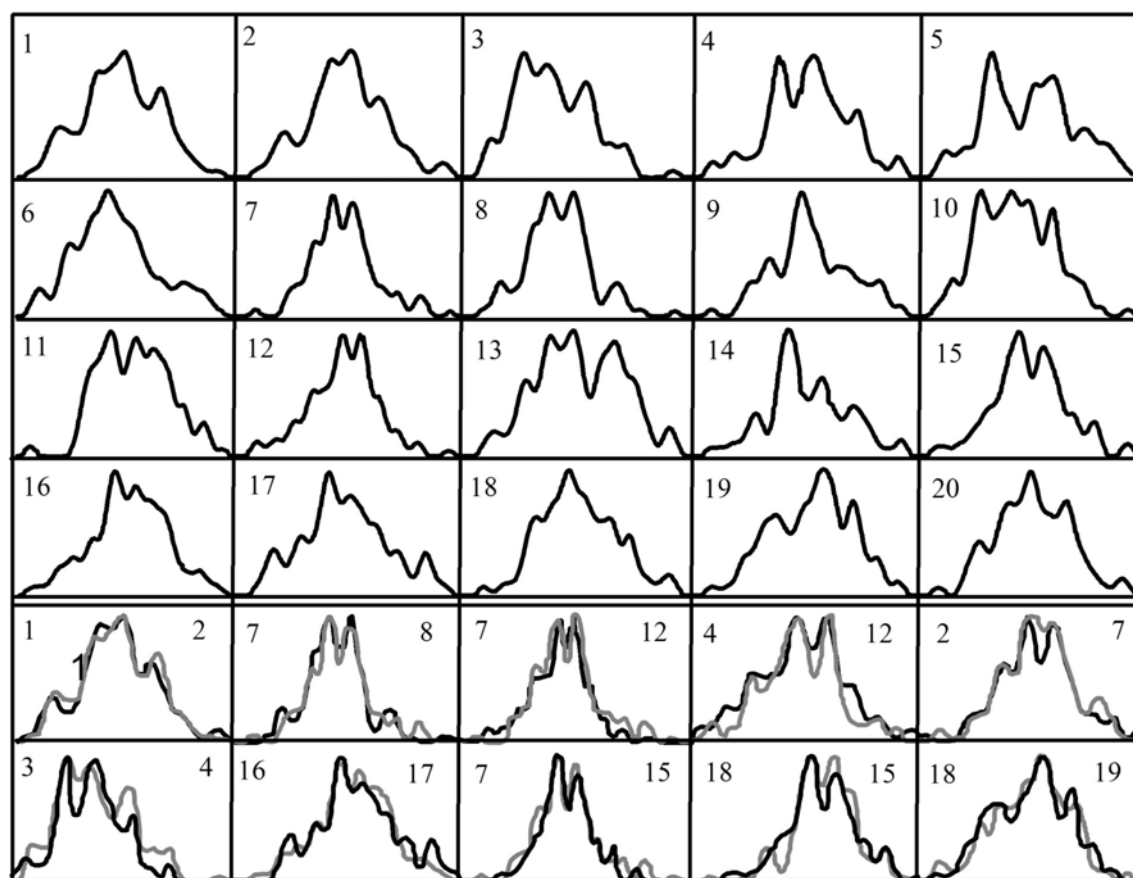


Рис. 27 Исходная последовательность и пары подобных гистограмм.

грамма в ней сравнивается со всеми другими гистограммами этой или другой подобной последовательности. В случае, если гистограммы сравниваются с гистограммами той же последовательности, необходимо произвести $N(N - 1)/2$ попарных сравнений, для различающихся последовательностей одинаковой длины - N^2 сравнений. Так, для исследования последовательности, рис. 27 А, необходимо 190 сравнений. Рис. 27 В дает 10 пар гистограмм из 190 возможных, которые, в результате экспертной оценки, признаны похожими.

Необходимо отметить, что в процессе попарного сравнения тонкой структуры гистограмм допускаются линейные преобразования, не изменяющие их тонкую структуру: сдвиги, растяжения или сжатия и зеркальные повороты. Совпадающие пары гистограмм, показанные на рис.27 В, приведены с учетом упомянутых линейных преобразований.

Как можно видеть из рис. 27 В, процесс экспертного сравнения состоит в оценке подобия формы пар гистограмм. Как показала практика, экспертная оценка является наиболее чувствительной к особенностям формы гистограмм и, как правило, не может быть повторена традиционными методами корреляционного анализа, спектрального анализа, введением различного рода мер сходства и т.д. [3]. Как следует из многолетних попыток создания автоматических алгоритмов сравнения форм гистограмм, полная или частичная автоматизация процесса экспертного сравнения возможна на пути создания комплексных алгоритмов моделирующих отдельные аспекты человеческого восприятия, в особенности его целостную

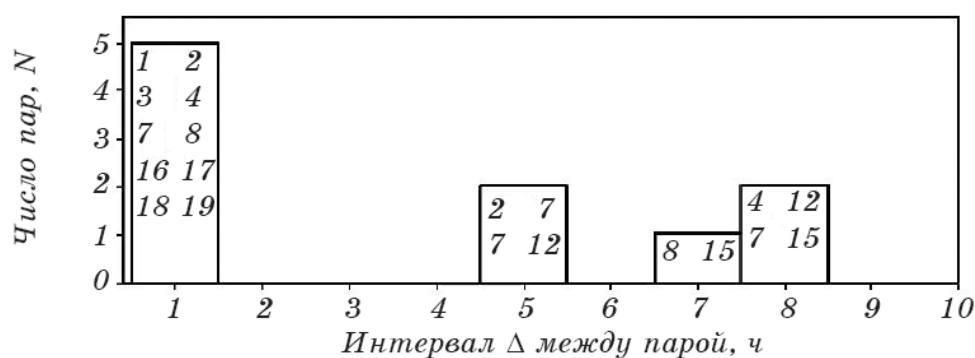


Рис. 28 Пример построения распределения интервалов для последовательности гистограмм, приведенной на рис. 27. N – число похожих пар, Δ – величина интервала, разделяющего пару гистограмм.

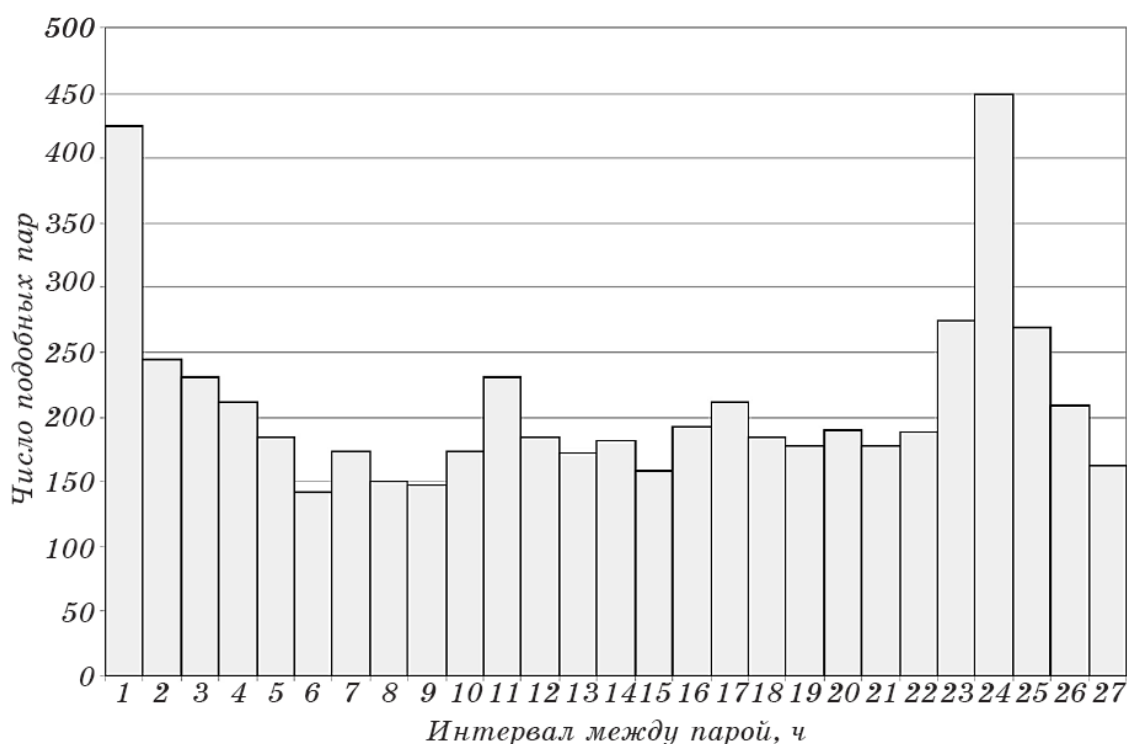


Рис. 29 Иллюстрация эффекта ближней зоны и суточного периода: 1-часовые гистограммы, построенные по 60-ти 1-мин измерениям активности ^{239}Pu . Видно, что наиболее вероятна повторная реализация гистограмм сходной формы в ближайший соседний интервал времени и через 24 часа. [17] Ось X – интервал времени, между парой гистограмм в часах; ось Y – число пар гистограмм, найденных подобными.

природу.

Заключительной стадией второго этапа является построение распределения интервалов между парами подобных гистограмм, показанное на рис. 28. Под интервалом Δ , понимается промежуток времени, разделяющий пару гистограмм во временном ряду. Экспертная оценка, являясь двузначной, имеет своим результатом заключение о подобии или непохожести гистограмм. В первом случае, интервал входит в распределение с весом равным единице, во втором – нулю. Так, например, в последовательности из $N = 20$ гистограмм, представленной

на рис. 27 А, имеются $N - \Delta = 15$ пар гистограмм разделенных интервалом равным пяти. Из них только две пары с номерами №2-№7 и №7-№12 найдены подобными. Следовательно, в результирующем распределении, основанном на последовательности, показанной на рис. 27 А, интервал №5 будет иметь значение равное двум.

Построение распределения интервалов, завершая процесс обработки экспериментальных данных, является основой дальнейшего анализа, в процессе которого получены основные свойства феномена макроскопических флуктуаций, кратко рассмотрены ниже.

На рис. 29 представлен пример распределения интервалов, полученного в результате экспериментального сравнения в эксперименте по измерению флуктуаций скорости α -распада. На нем видны два пика, соответствующих интервалам №1 и №24, которые связаны с наличием эффекта ближней зоны и суточного периода (см. ниже). Наличие этих пиков указывает на то, что наиболее вероятно повторная реализация гистограмм сходной формы в ближайший соседний интервал времени (интервал №1) и через 24 часа (интервал №24). Здесь использованы 1-часовые гистограммы, построенные по 60-ти 1-мин измерениям активности ^{239}Pu . [17]

2.5 Основные свойства феномена макроскопических флуктуаций.

Самый общий итог многолетних исследований эффекта макроскопических флуктуаций – доказательство неслучайности тонкой структуры формы гистограмм, построенных по относительно небольшому числу результатов измерений хода во времени флуктуаций процессов различной природы – от биохимических реакций и шумов в гравитационной антенне до α -распада. Рассмотрим основные черты МФ-эффекта.

2.5.1 Эффект ближней зоны.

Этот эффект состоит в достоверно более высокой вероятности появления сходных гистограмм в ближайших (соседних) не перекрывающихся интервалах рядов результатов измерений, рис. 29. Из “эффекта ближней зоны” естественно следует понятие “время жизни” данной идеи формы. Однако пока не удалось найти столь малое время, чтобы форма гистограмм не успевала измениться. Максимальная вероятность сходства только в первом, ближайшем соседнем интервале времени не изменялась при изменении этого интервала от нескольких часов до миллисекунд. Физический смысл этой своеобразной фрактальности требует дальнейших исследований. [3, 6, 15]

2.5.2 Универсальность феномена макроскопических флуктуаций.

Универсальность феномена макроскопических флуктуаций заключается в высокой вероятности сходства формы гистограмм, построенных по результатам одновременных, независимых измерений флуктуаций в процессах различной качественной природы. Сходство гистограмм при исследовании процессов, в которых диапазоны превращаемой энергии отличаются на десятки порядков (энергия α -распада и шумов в гравитационной антенне – различие около 40 порядков) означает, что этот фактор неэнергетической природы. Ввиду принципиального различия природы процессов и методов их измерений такое сходство также означает весьма общую природу фактора, определяющего форму гистограмм. [3, 6, 16]

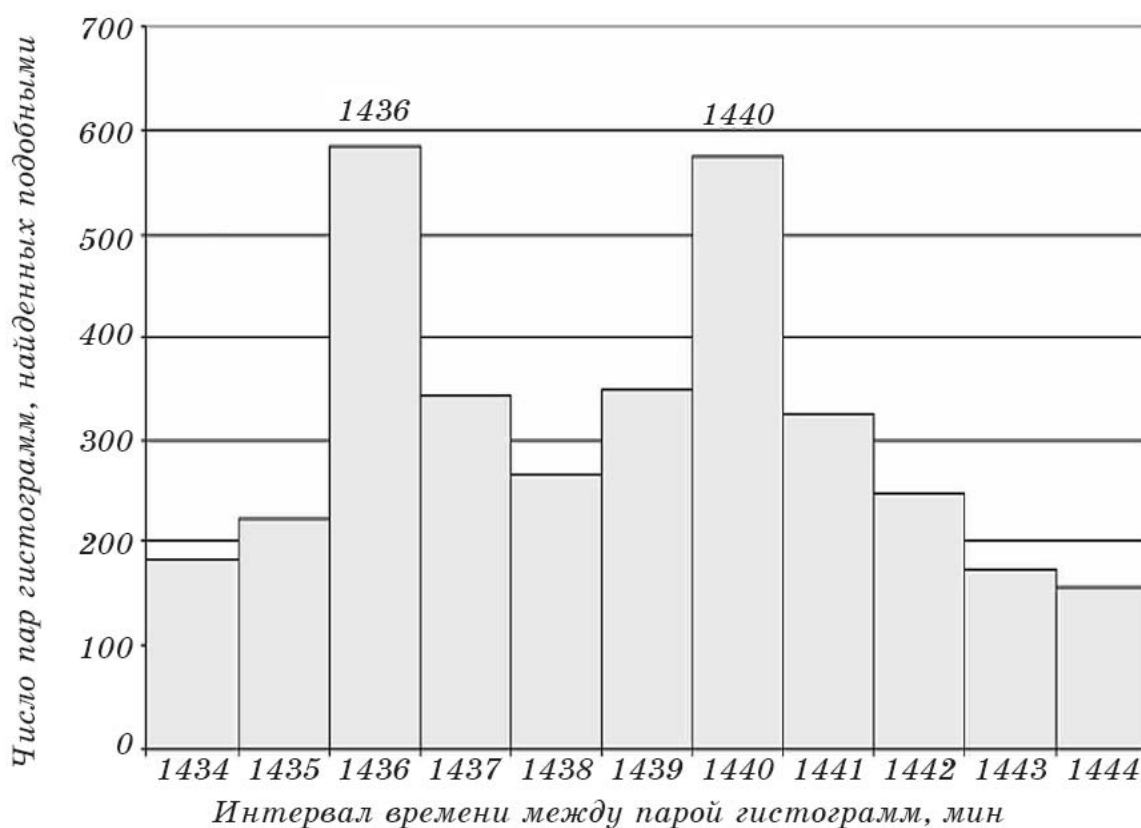


Рис. 30 “Расщепление” 24-часового пика при использовании 1-мин гистограмм, построенным по 60-ти 1-сек измерениям активности ^{239}Pu . [7] Ось X – интервал времени, между парой гистограмм в минутах; ось Y – число пар гистограмм, найденных подобными.

2.5.3 Периодичность проявления феномена макроскопических флуктуаций.

Важным свидетельством неслучайности формы гистограмм являются их закономерные изменения во времени. Эти закономерности проявляются в:

- Наличии суточных периодов изменения вероятности реализации гистограмм данной формы, рис. 29. Переход от часовых гистограмм, рис. 29 к гистограммам длительностью 1 мин привел к расщеплению суточного периода, заключающегося в существовании хорошо разрешимых “звездного” (1436 минут) и “солнечного” (1440 минут) суточных периодов, рис. 30. Наличие этих периодов означает зависимость формы гистограмм от вращения Земли вокруг своей оси. При этом экспозиция относительно картины звездного неба и экспозиция относительно Солнца определяют форму гистограмм независимо.
- Наличии около-27-и суточных периодов изменения вероятности реализации гистограмм данной формы. Эти периоды можно считать признаком зависимости формы гистограмм от экспозиции относительно ближайших небесных тел – Солнца, Луны и, возможно, планет. [10]
- Наличии годовых периодов, проявляющихся в высокой вероятности реализации сходных гистограмм через годы; в существовании хорошо разрешимых “календарного” (365 солнечных суток) и “звездного” (сидерического: 365 солнечных суток плюс 6 часов и 9 минут) годовых периодов. [18]

- Результат, представляющийся парадоксальным: в дни равноденствий наблюдается четкий период увеличения вероятности сходства гистограмм, равный 718 минутам, т.е. звездным полусуткам. В дни солнцестояний такого периода нет. Этот феномен свидетельствует о зависимости формы гистограмм от положения Солнца на Эклиптике. При верности этого вывода можно ожидать, что на экваторе период 718 минут будет наблюдаться круглогодично.

Все вышеперечисленные периоды означают зависимость формы гистограмм от 1) вращения Земли вокруг своей оси и от 2) движения Земли по околосолнечной орбите

2.5.4 Эффект местного времени.

Зависимость формы гистограмм от вращения Земли вокруг своей оси с большой четкостью проявляется также в эффекте местного времени - высокой вероятности появления пар сходных гистограмм в разных географических пунктах – от Арктики до Антарктики, в Западном и Восточном полушариях в одно и то же местное (долготное) время. Необходимо отметить, что синхронность по местному времени с точностью в 1 минуту наблюдается на предельно возможных на Земле расстояниях (около 15 тысяч километров) независимо от широты местности. В некоторых случаях наблюдается также синхронность по абсолютному времени.

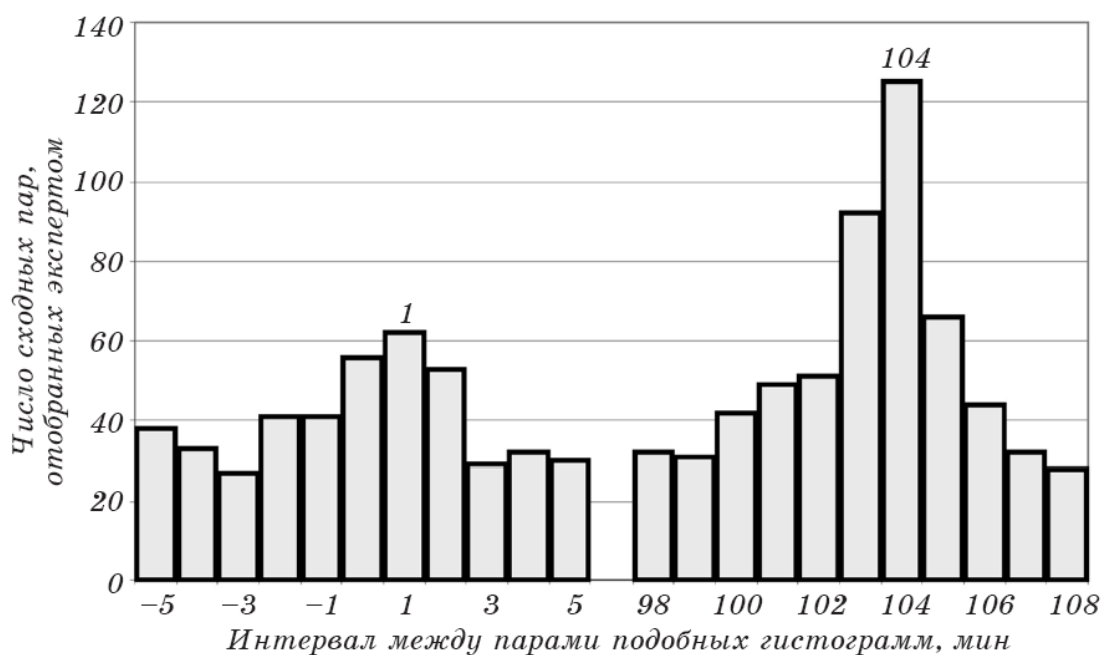


Рис. 31 Синхронность по абсолютному (левый рис.) и местному (правый рис.) времени. Ось X – интервал между парами подобных гистограмм, мин.; ось Y – число сходных пар, отобранных экспертом.

На рис. 31, в качестве примера, приведены два распределения интервалов, построенные на основе рядов флуктуаций скорости α -распада ^{239}Pu , измеренных 1.03.2003 в Пуццино ($54^\circ 50'$ с.ш., $37^\circ 38'$ в.д.) и на станции Новолазаревская в Антарктиде ($70^\circ 02'$ ю.ш., $11^\circ 35'$ в.д.). Расстояние между точками измерений около 14500 км и разность местного времени 103 мин. В левой части рис. 31 приведено распределение интервалов, показывающее

синхронность появления сходных пар гистограмм по абсолютному времени, в правой части рис. 31 – по местному времени. Видно, что синхронность по местному времени выражена значительно более ярко. [3, 6]

Интенсивные исследования эффекта местного времени, проводимые в 1985-2005 г.г., показали, что данный эффект существует для расстояний между местами проведения измерений от сотен километров до, практически, максимально возможных на Земле расстояний (ок. 15000 км). Эти результаты были получены, в основном, с использованием в качестве исходного экспериментального материала флуктуаций скорости α -распада. В период 2005-2007 г.г. были начаты исследования эффекта местного времени с использованием полупроводниковых источников флуктуаций, позволивших увеличить временное разрешение метода исследований [116-120]. Благодаря этому удалось показать существование эффекта местного времени вплоть до расстояний порядка одного метра.

Возможность наблюдения эффекта местного времени на расстояниях порядка одного метра позволила его лабораторные исследования, недоступные ранее, в ходе которых был обнаружен ряд новых свойств эффекта, к которым можно отнести следующие:

- фрактальная структура пика местного времени;
- зависимость выраженности эффекта местного времени от пространственных направлений;
- возможность наблюдения эффекта местного времени для случая движущихся источников флуктуаций.

Рассмотрим подробнее перечисленные свойства.

2.5.4.1. Фрактальная структура пика местного времени.

Пример того, что имеется в виду под фрактальной структурой пика местного времени приведен на рис. 32 [117, 120]. Здесь, на рис. 32), приведен пример эффекта местного времени, полученный на основе пространственно-разнесенных измерений в г. Пущино и пос. Большевик, долготная разность расстояний между которыми равна 15 км. Соответствующая этому расстоянию разность местного времени равна 62.7 сек. Представленное на рис. 32) распределение интервалов, получено в результате экспертного сравнения последовательностей 1-секундных гистограмм. Как можно видеть, распределению присущ четкий пик, соответствующий временному интервалу между гистограммами равному 63 сек. С учетом точности синхронизации момента начала измерений (0.1 – 0.2 сек), и длительности используемых гистограмм, можно считать, что данный пик с хорошей точностью соответствует разности местного времени $\Delta t = 62.7$ сек между г. Пущино и пос. Большевик.

Пик местного времени, получаемый на распределениях интервалов, как правило, всегда очень острый, шириной в 1-2 гистограммы [3, 6, 10], т.е., практически лишен какой-либо структуры. То же можно видеть и на рис. 32). Возможность построения 0.2 сек гистограмм, на основе тех же рядов экспериментальных данных, на основе которых было получено распределение рис. 32), позволило рассмотреть задачу детального исследования структуры полученного на рис. 32) пика местного времени, увеличив, таким образом, временное разрешение метода в пять раз, рис. 32).

Результаты экспертного сравнения для последовательностей 0.2 сек гистограмм, приведены на рис. 32 б). Как можно видеть из этого рисунка, максимальное подобие наблюдается

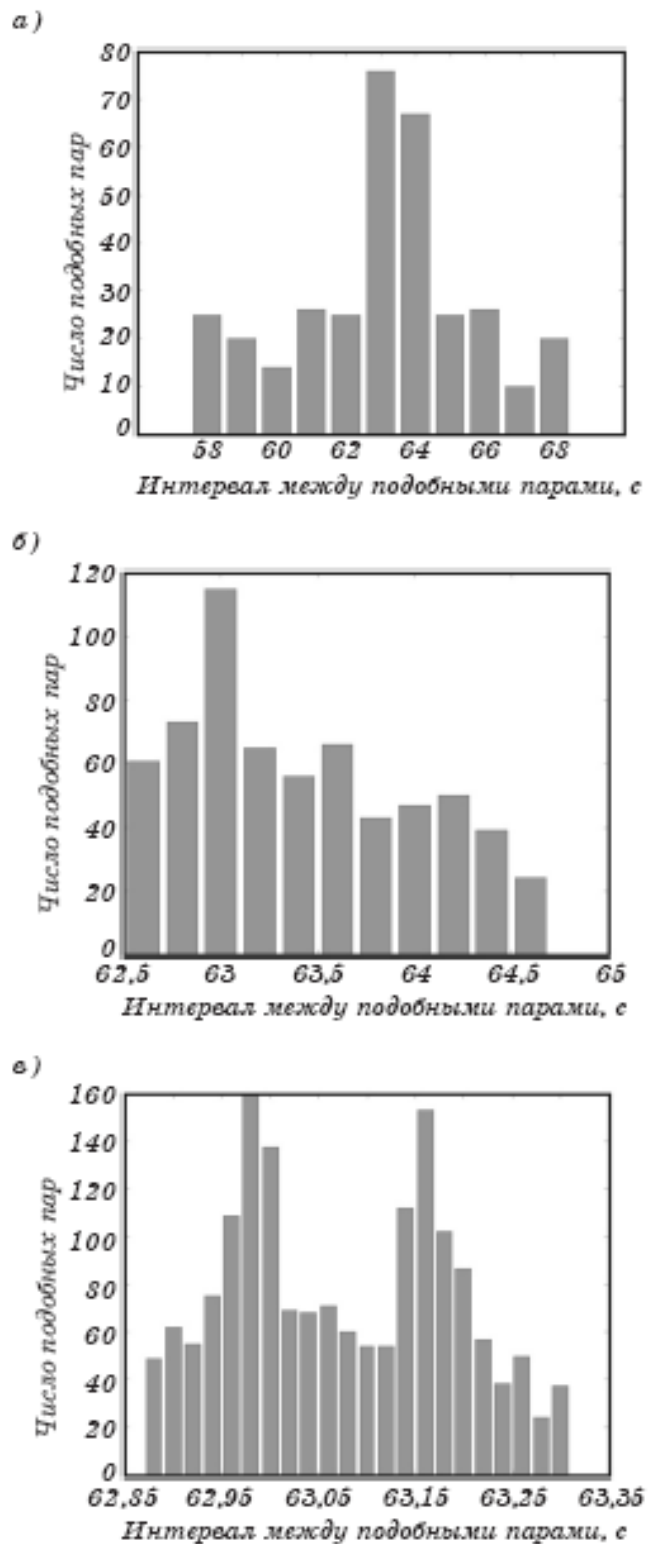


Рис. 32 Эффект местного времени, **а-б)** и его расщепление на два субпика, **в)**. Ось X – интервал между парами подобных гистограмм, сек.; ось Y – число сходных пар, отобранных экспертом.

для пар гистограмм, разделенных интервалом, соответствующим разности местного времени по прежнему равной 63 сек, но теперь уже эта величина определена с точностью в 0.2 сек.

Обращает на себя внимание, что в результате 5-кратного увеличения разрешения при переходе от 1-секундных к 0.2-секундным гистограммам мы получили снова один острый пик, соответствующий разности местного времени для мест проведения измерений. Т.е. изменение временного масштаба не привело к изменению вида распределения интервалов. Для того чтобы увеличить временное разрешение еще на порядок мы перешли к гистограммам, построенным по 0.02-секундным отрезкам временных рядов.

Распределение интервалов для этого случая представлено на рис. 32 в). В результате экспериментального сравнения было получено распределение интервалов, состоящее из двух пиков. Первый пик, с учетом принятого смещения, соответствует разности местного времени 62.98 сек, а второй – 63.16 сек. Соответственно, разность между пиками составляет 0.18 сек. Видно, что пик на рис. 32 б) расщепился на два субпика, рис. 32 в), которые, как и пики на рис. 32 а) и рис. 32 б), практически бесструктурные.

Эффект местного времени, как отмечено в [124], связан с вращательным движением Земли. Самое простое объяснение этого феномена могло бы быть следующим. Через время Δt равное разности долготных времен в местах проведения измерений, измерительная №2 попадает в ту же точку пространства, где до этого была система №1, что и обуславливает подобие формы наблюдаемых гистограмм. Однако, кроме вращательного Земле присуще еще и поступательное, орбитальное движение, скорость которого значительно превышает скорость вращательного движения. Поэтому утверждение о том, что измерительная система №2 попадет в ту же точку пространства, становится неверным. Здесь можно говорить только о направлениях. Действительно, если мы проведем две плоскости, проходящие через ось вращения Земли и точки проведения измерений, то через время Δt плоскость №2 будет параллельна плоскости №1. Отсюда можно сделать вывод о том, что подобие формы гистограмм каким-то образом связано с направлением в пространстве.

Расщепление суточного периода повторяемости формы гистограмм на солнечный и звездный говорит о том, что в пространстве существуют два выделенных направления: на Солнце и на “сферу неподвижных звезд”. Действительно через время равное 1436 мин Земля совершает один полный оборот вокруг своей оси и плоскость измерительной системы становится параллельной направлению, которое она имела одни звездные сутки назад. А еще через 4 мин в эту плоскость попадает Солнце, что совпадает с солнечными сутками равными 1440 мин.

Предположим, что расщепление (рис. 32) имеет ту же природу, что и расщепление суточного периода на солнечный и звездный. Тогда из разности между солнечным и звездным периодами ΔT равной $\Delta T = 4$ мин. вычисляем коэффициент пропорциональности k :

$$k = \frac{4\text{мин}}{24\text{ч}} = \frac{240\text{сек}}{86400\text{сек}} \approx 2.78 \cdot 10^{-3} \quad (6)$$

Разность долготного времени Пушино-Большевик травна $t \approx 62.7$ сек. Отсюда, разность звездного и солнечного периодов, соответствующая данному Δt :

$$\Delta t = kt = 62.7 \times 2.78 \cdot 10^{-3} \approx 0.17. \text{ мин.} \quad (7)$$

Как можно видеть из рис. 32, разность между пиками в распределении интервалов равна 0.18 сек, что с учетом точности определения пиков (0.02 сек) совпадает с расчетной величиной. На этом основании мы можем говорить о том, что обнаруженное нами расщепление пика

местного времени обусловлено теми же причинами, какие вызывают появление солнечного и звездного периодов в картине расщепления суточного периода.

Говоря о выделенных направлениях в пространстве, неявно предполагается, что измерительная система является “направленной” и в силу этого может чувствовать, разрешать эти направления. Подобные представления, вполне приемлемые для случая, когда мы говорим о расщеплении суточного периода, становятся проблематичными, в случае расщепления, обнаруженного нами для эффекта местного времени и описанного выше. Здесь угол, который должна разрешать измерительная система, уже настолько мал, что говорить о какой-либо присущей ей диаграмме направленности не приходится. В этом случае представляется более логичным предположить, что мы имеем дело с неоднородностью околоземного пространства, каким-то образом тесно связанной с выделенными направлениями, обсуждавшимися выше. Судя по обнаруженному расщеплению пика местного времени, эта неоднородность имеет определенную структуру, предположительно фрактальную, с коэффициентом подобия (6). Проверка этого предположения была впервые произведена в эксперименте, представленном ниже [120].

В этом эксперименте [120] в качестве исходных экспериментальных данных использовались синхронные пространственно-разнесенные измерения флуктуаций тока в полупроводниках, проведенные в г. Пущино и г. Ростов-на-Дону. Методика обработки экспериментальных данных и построения гистограмм в данном эксперименте такая же, как описано выше.

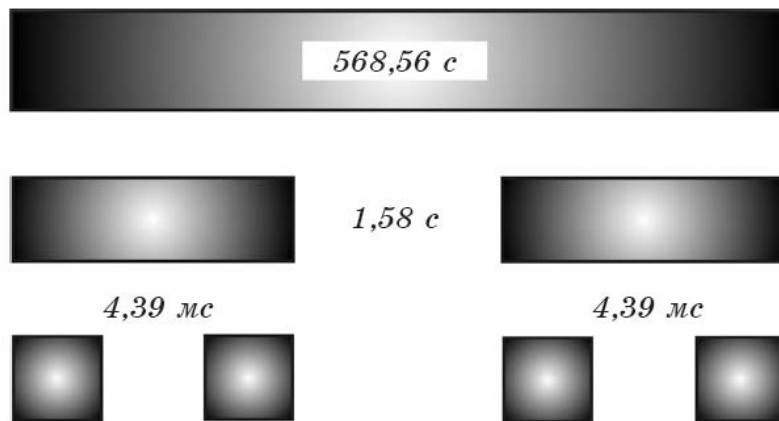


Рис. 33 Схема расщеплений первого и второго порядков для величины местного времени 568.56 сек.

На рис 33 приведена диаграмма, представляющая расчет ожидаемой схемы расщеплений первого и второго порядков для случая, когда величина местного времени равна 568.56 сек. Величина расщепления первого порядка вычислялась на основании (6) – (7). Для расщепления второго порядка использовалась формула:

$$\Delta t'' = k \Delta t'. \quad (8)$$

Для величины расщепления первого порядка получено значение 1.58 сек, а для расщепления второго порядка – 4.39 мсек. На диаграммах аналогичных показанной на рис. 33, в отличие от рис. 32, приводятся не распределения интервалов, а только относительные положения максимумов. Это позволяет отобразить на одном рисунке структуру расщеплений, содержащуюся в большом числе распределений интервалов.

Результаты экспертного сравнения приведены на диаграмме, рис. 34. Здесь, слева от каждой линии прямоугольников, указаны длительности используемых гистограмм. Между прямоугольниками, во второй и четверной строке указаны величины расщеплений, при этом расчетные величины для каждого случая приведены в скобках. Эти величины совпадают с приведенными на рис. 33. Внутри каждого из прямоугольников приведено положение пика на временной оси. За нуль оси принимается момент начала синхронной записи в г. Пущино и г. Ростов-на-Дону.

Как можно видеть, рис. 34, кроме расщепления первого порядка, здесь также присутствует расщепление второго порядка. Правда, в рассматриваемом эксперименте удалось получить расщепление только для “солнечного” пика. Нами было подмечено, что выраженность солнечного и звездного пиков связана с ориентацией измерительной системы. Но эта особенность пока не была объектом детального изучения.

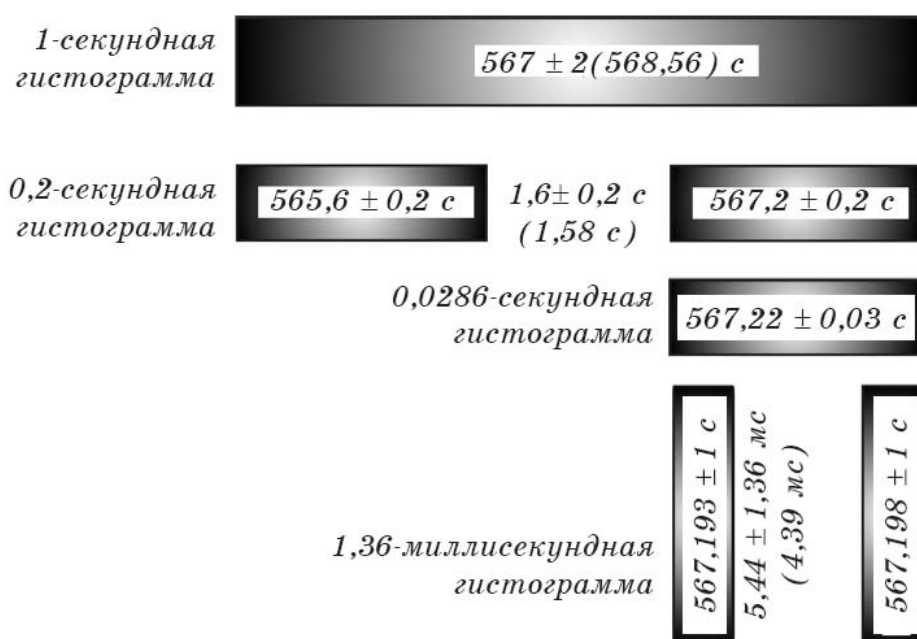


Рис. 34 Структура расщеплений пика местного времени, полученная для синхронных пространственно-разнесенных измерений в г. Пущино и г. Ростов-на-Дону [120]

Дальнейшие эксперименты, направленные на исследование расщеплений второго порядка в суточном периоде, показали, что наблюдаемая картина расщеплений симметрична, как это показано на рис. 33. На основе полученных результатов, можно предположить, что в общем случае возможны также расщепления третьего и высших порядков, величина которых может быть оценена на основании:

$$\Delta t^n = k^n \Delta t. \quad (9)$$

Проверка правильности этого предположения, а также нахождения максимальной величины n для которого справедливо выражение (9) - задача наших будущих исследований.

2.5.4.2. Азимутальная выраженность эффекта местного времени [125-128].

На рис. 35) представлена схема экспериментальной установки, использовавшейся для исследования выраженности эффекта местного времени в зависимости от направлений в про-

странстве. Установка состоит из пары источников флуктуаций, жестко фиксированных на деревянном основании на расстоянии 1.36 м друг от друга. Основание может вращаться в горизонтальной плоскости, вокруг своей оси и может быть зафиксировано в любом, требуемом направлении. Аналоговый сигнал с выхода каждого источника оцифровывается при помощи цифрового запоминающего осциллографа, объем памяти которого достаточен для получения записи необходимой длины [126].

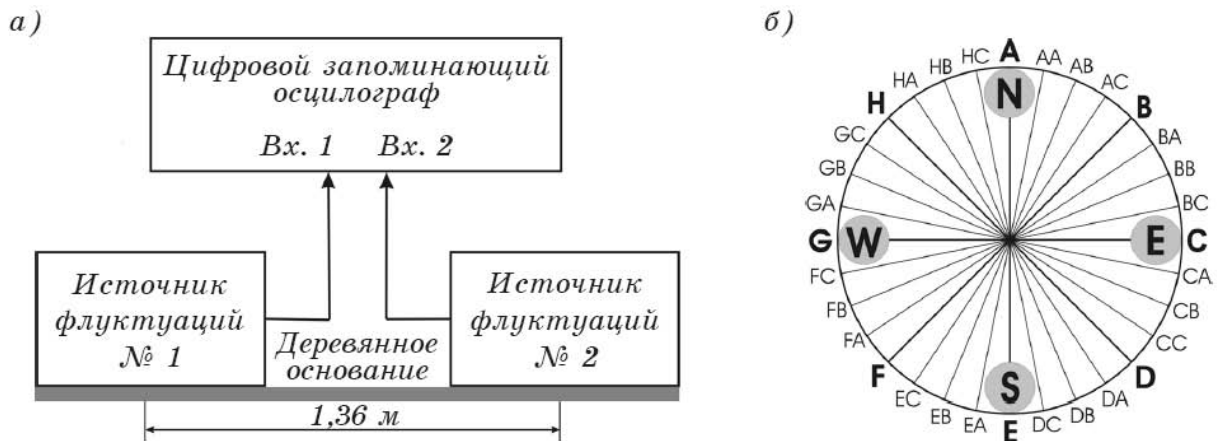


Рис. 35 Схема эксперимента, а) и диаграмма исследуемых пространственных направлений, б)

Под направлением в пространстве понимается линия, проходящая через источники флуктуаций. Диаграмма пространственных направлений, исследованных в эксперименте, показана на рис. 35). Буквами обозначены положения источника флуктуаций №1. Буквы N, S, W и E показывают ориентацию по сторонам света.

После того, как деревянное основание с источниками флуктуаций устанавливалось в необходимом пространственном направлении, производилась синхронная по обоим каналам регистрация сигналов флуктуаций. В каждом эксперименте, как следует из рис. 35), производилось 32 синхронных регистрации сигналов флуктуаций.

Обычно считается, что величина эффекта местного времени пропорциональна разности географических долгот мест измерений и не зависит от разности их широт. Это утверждение подтверждено всей совокупностью ранее проведенных экспериментов [1-18]. Отсюда возникло представление об аксиальной симметрии фактора, влияющего на форму гистограмм.

Независимость эффекта от широты мест измерений, может быть истолковано, как зависимость формы гистограмм от направления в пространстве, задаваемого осью вращения Земли и точкой проведения измерений. В таком случае важен только угол между двумя плоскостями в пространстве, определяемыми местами проведения измерений. Этот угол равен разности долгот и не зависит от широты мест измерений. Но, в случае измерений с фиксированной пространственной базой $\Delta L_0 = Const$, как это имеет место в настоящем исследовании, результат эксперимента становится зависимым от широты, θ . Действительно, время, через которое источник флуктуаций № 2 будет определять то же направление в пространстве, что и источник флуктуаций № 1 до этого, зависит от скорости v , с которой каждый из источников флуктуаций пройдет расстояние $\Delta L_0 = Const$. Используя в качестве модельных представлений о фигуре Земли эллипсоид Красовского [129] и учитывая высоту

места проведения измерений над уровнем моря, h , можем записать выражение для v :

$$v(\theta, h) = \frac{2\pi}{T} \left(\sqrt{\frac{R_p^2}{\frac{R_p^2}{R_e^2} + tg^2(\theta)}} + h \right), \quad (10)$$

где $R_p = 6356863$ м. и $R_e = 6378245$ м. - величины полярного и экваториального радиусов Земли, принятые для эллипсоида Красовского [129], $\theta = 54^\circ 50.037'$ - широта места проведения измерений (г. Пущино, Московской области), определенная при помощи GPS-приемника, $T = 86160$ сек- период вращения Земли и $h = 170$ м - высота места измерения над уровнем моря. Учет численных значений в (10) дает величину $v(\theta, h) = 268/$. В случае проведения измерений, например, в приэкваториальных областях, данная величина а, следовательно, и связанная с ней величина местного времени, может отличаться почти в два раза от полученного значения $v(\theta, h) = 268/$ из-за широтной зависимости $v(\theta, h)$.

Полученное значение $v(\theta, h)$ позволяет рассчитать величину местного времени для каждого из измеренных направлений, рис. 35). Т.к., предполагается, что величина местного времени зависит только от долготной разности расстояний между местами измерений, то ожидаемая величина местного времени, в этом случае будет равна:

$$\Delta t = \frac{\Delta L_0 \sin \alpha}{v(\theta, h)} \quad (11)$$

где $\Delta L_0 = 1.36$ м- расстояние между генераторами, $\alpha \in [0, 2\pi]$ - азимутальный угол, отсчитываемый от направления на север ("А") по часовой стрелке. Для используемого расстояния между генераторами величина Δt равна 5 мсек для случая ориентации измерительной системы в направлении восток-запад и равна нулю для направления север-юг. На рис. 38 сплошной линией показан результат расчета величины местного времени в зависимости от угла α , произведенного на основании (11). По оси абсцисс на рис. 38 показаны пространственные направления, соответствующие направлениям, показанным на рис. 35).

На рис. 36 представлены суммарные, по нескольким сериям экспериментов, распределения интервалов для исследованных пространственных направлений. Схема азимутальных направлений в центре рисунка совпадает со схемой на рис. 35). Внимательное изучение рис. 36 позволяет выделить две группы распределений интервалов. Первой группе, она выделена на рис. 36 серым цветом, соответствуют направлениям "А", "С", "Е" и "G", т.е., направления север-юг и восток-запад и их ближайшая (около $\pm 11.25^\circ$) окрестность. Суммарные распределения, соответствующие этим направлениям, показаны отдельно на рис. 37. Данным распределениям присуща четко выраженная структура с ярко выраженным пиком, величина которого с хорошей точностью соответствует полученной из (11) величине местного времени. В то же время, "диагональные" распределения интервалов в широкой окрестности направлений "В", "D", "F" и "H" полиэкстремальны и не имеют четко выраженного пика. В тех же случаях, когда на "диагональных" распределениях присутствует единственный пик, он выражен значительно хуже чем в случае направлений "А", "С", "Е" и "G". И, как правило, в этом случае его величина отличается от расчетной.

Примеры распределений интервалов для "недиагональных" пространственных направлений показаны на рис. 37. Для направлений А и Е, показанных в первом столбце на рис. 37 (направления вдоль земного меридиана, север-юг) характерна форма с пиком в нуле. В то

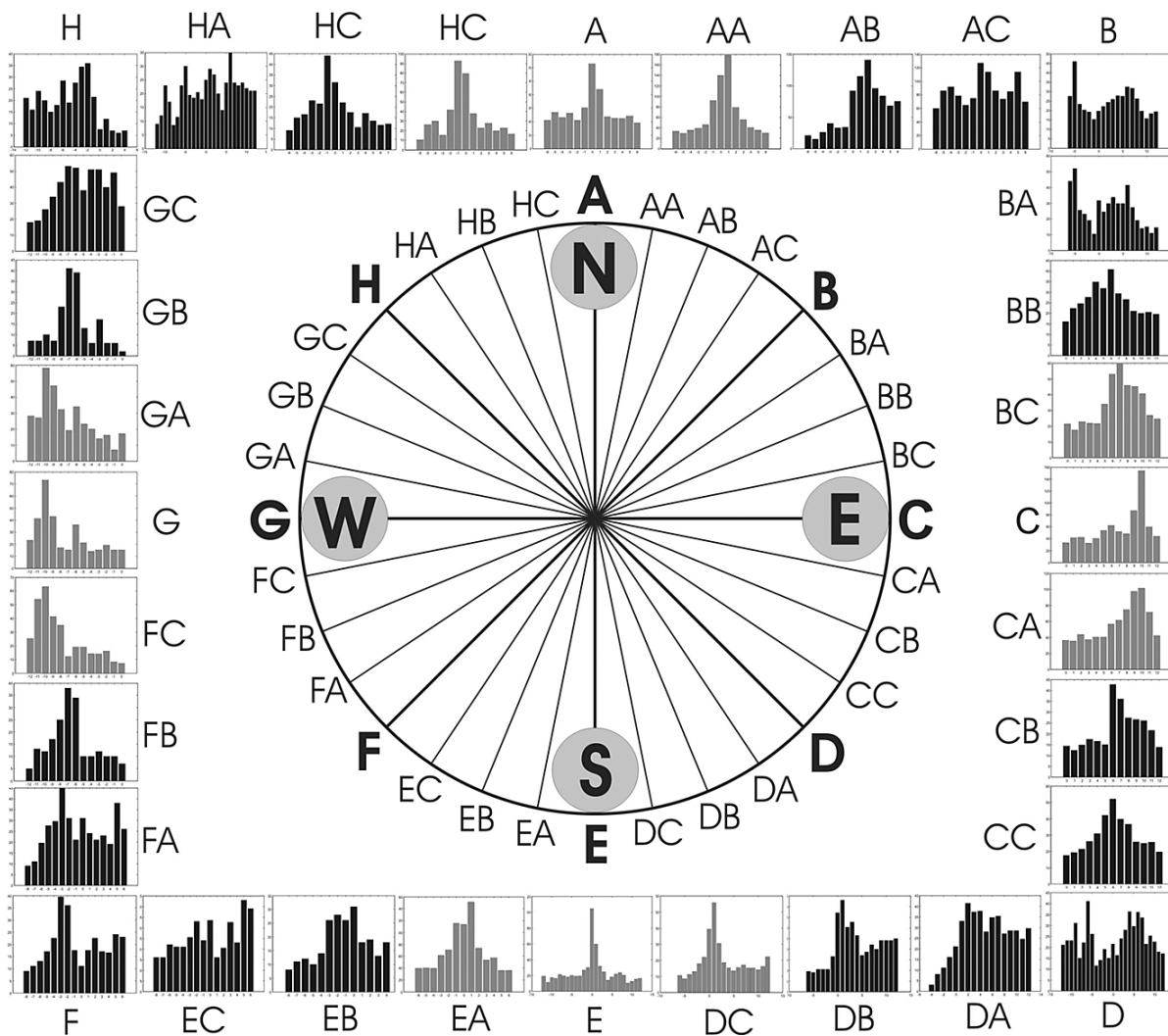


Рис. 36 Суммарные распределения интервалов для исследованных пространственных направлений. Схема азимутальных направлений в центре рисунка совпадает со схемой на рис. 35)

время как для направлений C и G – второй столбец на рис. 37, (направление вдоль земной параллели, Восток-Запад) пик на распределениях соответствует интервалу равному величине разности местного времени для расстояния между источниками флуктуаций 1.36 м. Распределения для направлений C и G – это, по сути, измерения вдоль одного и того же направления, при противоположной ориентации генераторов, поэтому, как можно видеть из правого столбца на рис. 37, эти распределения зеркально симметричны. Зеркальная симметрия распределений C и G является свидетельством выраженности эффекта для данного пространственного направления.

На рис. 38, положения максимумов на распределениях интервалов, для всех серий экспериментов показаны точками. Данный рисунок носит качественный характер, т.к. часть распределений интервалов имеет сложную полиэкстремальную структуру и, поэтому, не может быть графически представлено в виде одной точки на графике. Для таких, полиэкстремальных распределений бралось не более трех главных максимумов, которые, затем, учитывались при построении графика на рис. 38. Главная задача построения рис. 38 – дать суммарную

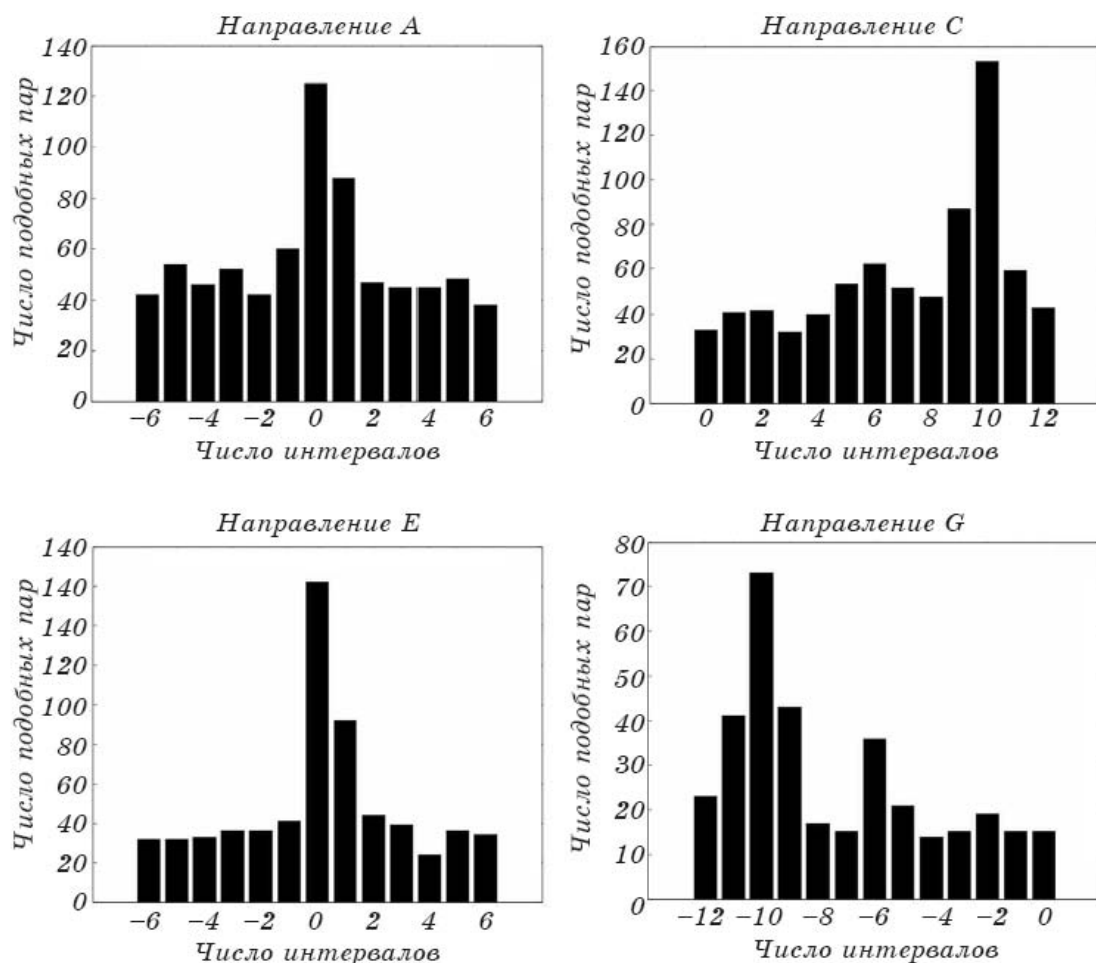


Рис. 37 Примеры распределений интервалов для недиагональных пространственных направлений. Ось X – число интервалов длительностью 0.5 мсек каждый между парами подобных гистограмм; ось Y – число сходных пар, отобранных экспертом.

картину совпадения полученных результатов с результатами модельного расчета.

Из рис. 38 хорошо видно, что для диагональных пространственных направлений экспериментальные результаты с хорошей точностью совпадают с теоретической кривой. В то время как для диагональных направлений такого согласия не наблюдается.

Суммируя можно отметить, что в результате исследования выраженности эффекта местного времени в зависимости от выделенных пространственных направлений было найдено, что данный эффект наиболее четко выражен в узкой окрестности направлений север-юг и восток-запад. В этих направлениях обнаруживаемая экспериментально величина местного времени с хорошей точностью совпадает с расчетной. В окрестности направлений "В", "D", "F" и "H" распределения интервалов, как правило, полиэкстремальны и не всегда имеют достаточно выраженный пик местного времени. Величина местного времени, связываемая с положением максимального пика на этих распределениях, обычно плохо согласуется с расчетной.

Полученные результаты имеют принципиальное значение, для понимания феномена местного времени, т.к. указывают на характер анизотропии околоземного пространства. Действительно, если предположить, что эффект местного времени обусловлен некоторой, неизменной

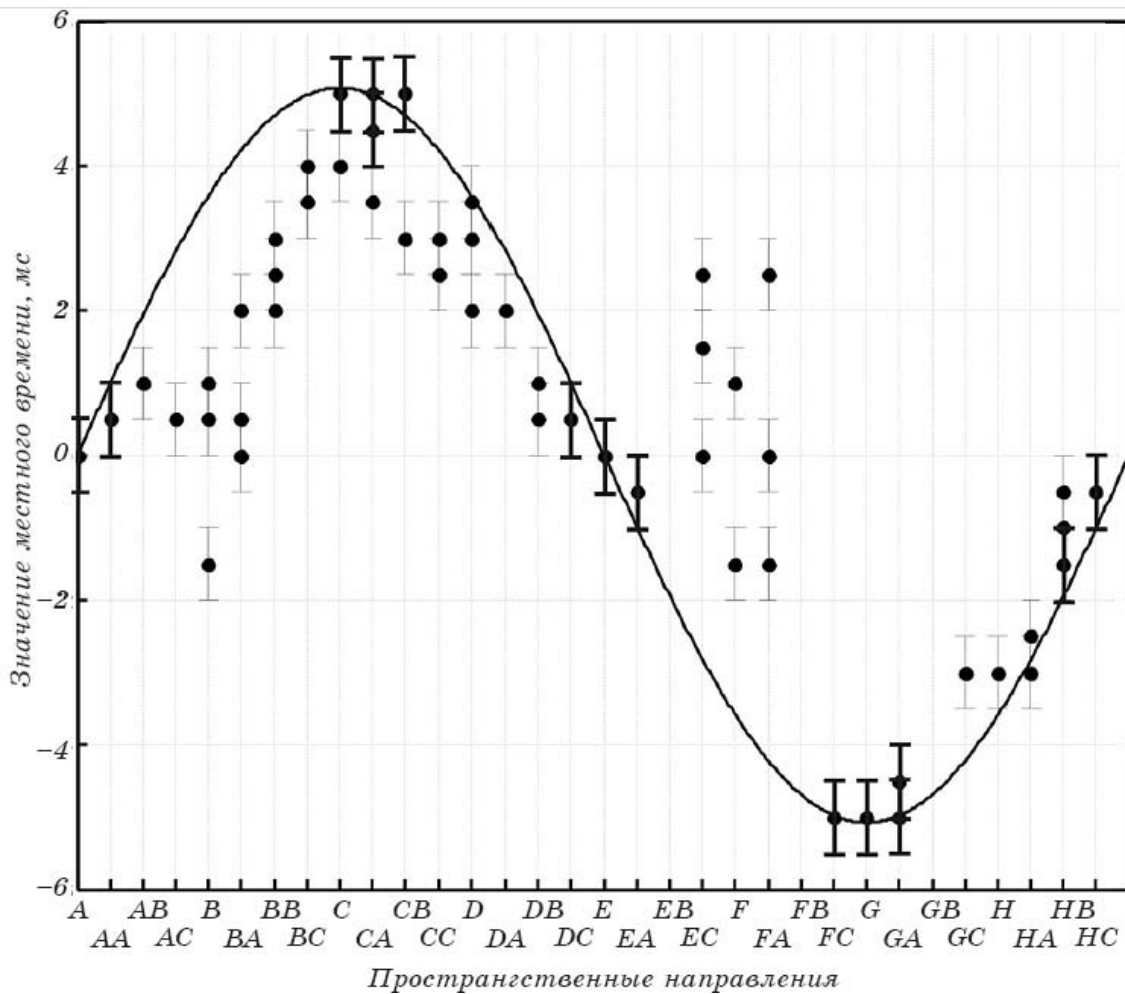


Рис. 38 Расчетные величины местного времени (сплошная линия) и величины полученные экспериментально. По оси абсцисс показаны азимутальные направления, соответствующие направлениям, показанным на рис. 35 б).

вдоль меридиана секторной структурой, то выраженность пика местного времени на “диагональных” распределениях интервалов должна быть такой же четкой, как и для направлений “А”, “С”, “Е” и “G”. Т.к. это не обнаружено, то мы вынуждены отказаться от представления о секторной структуре и заключить, что фактор, оказывающий влияние на форму тонкой структуры гистограмм связан с некоторыми выделенными направлениями, которые, вероятно, определяются как анизотропными свойствами использованных источников флуктуаций, так и анизотропией околоземного пространства. Узость областей, в которых наблюдается четкий пик местного времени, позволяет говорить как об острой анизотропии пространства, так и о том, что существует достаточно эффективный механизм, позволяющий источнику флуктуаций чувствовать эту анизотропию.

2.5.4.3. Эффект местного времени для движущихся источников флуктуаций [116, 130].

Как следует из вышеизложенного, эффект местного времени обусловлен перемещением измерительной системы за счет вращательного движения Земли. Отсюда следует предположение,

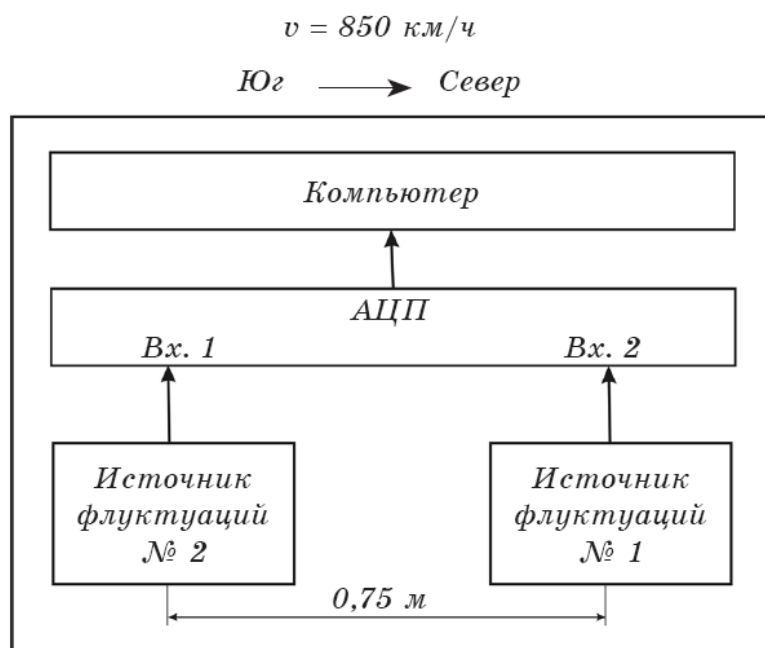


Рис. 39 Схема эксперимента для исследования эффекта местного времени для движущихся источников флуктуаций.

что должно быть верно и обратное: измерительная система, связанная с движущимся объектом, способна зарегистрировать это движение в форме эффекта местного времени. Для проверки этого предположения был поставлен эксперимент на борту самолета А319, движущегося со средней скоростью 850 км/ч в направлении юг-север, который рассмотрен ниже.

На рис. 39 представлена схема эксперимента. Экспериментальная установка, в данном случае, состояла из двух источников флуктуаций №1 и №2, фиксированных на расстоянии 0.75 м по ходу движения самолета, устройства аналогово-цифрового преобразования и персонального компьютера, используемого для управления процессом сбора и хранения экспериментальных данных. Как следует из рис. 39, в основных чертах, используемая в рассматриваемом эксперименте установка не отличается от измерительной системы, показанной на рис. 35). За одним существенным исключением – измерительная система, показанная на рис. 39 – является движущейся. Т.е. представленный на рис. 39 эксперимент является в некотором роде “искусственным” вариантом эффекта местного времени, т.к., в нем вместо вращения Земли, задающего величину долготной разности времен для используемой пространственной базы, используется “искусственное” движение, связанное со скоростью движения самолета. Расчетная разность местного времени равна величине временного интервала, необходимого для прохождения самолетом используемой пространственной базы – 0.75 м. В данном случае, расчетная величина местного времени Δt равна $\Delta t = 3.18 \pm 0.5$ мсек.

Необходимо отметить, что возможность наблюдения эффекта местного времени связана с неоднородностью пространства, сквозь которое движется, показанная на рис. 39, измерительная система: через время, равное времени прохождения самолетом расстояния 0.75 м источник флуктуаций №2 будет в точности в том же месте пространства, что источник №1 до этого. Одинаковость условий, должна приводить к подобию форм гистограмм, разделенных временным интервалом Δt , равным разности местного времени.

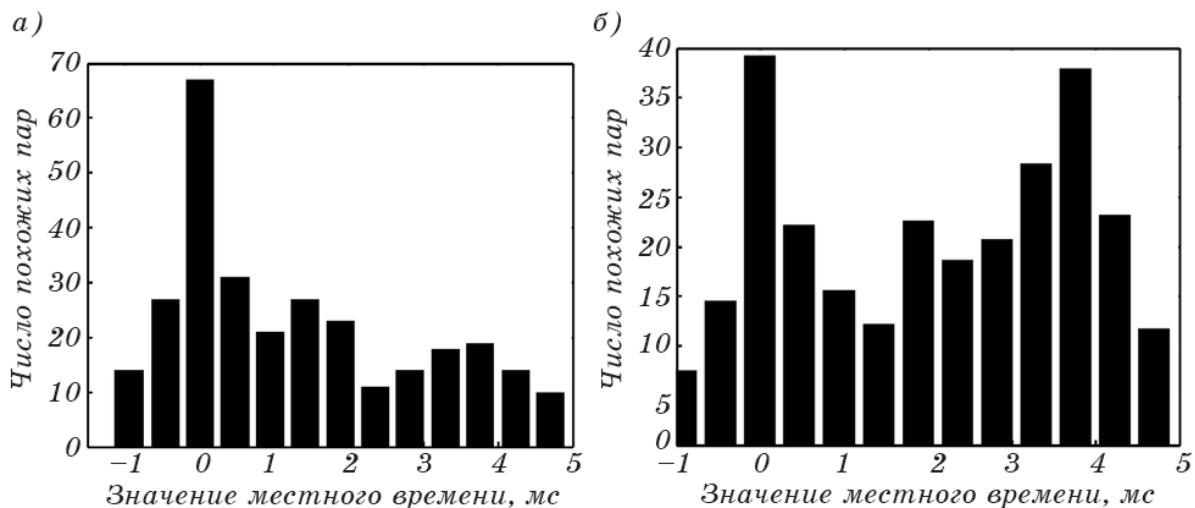


Рис. 40 Распределение интервалов для случая неподвижной, *а*) и движущейся измерительной системы, *б*). Измерения проводились в одно и то же время; обе системы ориентированы в направлении юг-север.

На рис. 40) приведены усредненные результаты экспертного сравнения для наземной неподвижной измерительной системы ориентированной в точности как и движущаяся. Измерения для неподвижной и движущейся измерительных систем проводились в одно и то же время суток. Как можно видеть из рис. 40) распределение интервалов в этом случае повторяет результаты, представленные на рис. 36 и рис. 37 для направлений вдоль земного меридиана.

Результат для случая движущейся измерительной системы представлен на рис. 40). Он совпадает с представленным на рис. 40) в части наличия нулевого пика, такого же как и на рис. 40), но отличается от рис. 40) наличием пика положение которого, с точностью до погрешностей эксперимента, совпадает с расчетной величиной разности местного времени Δt .

Полученные результаты подтверждают предположение о том, что эффект местного времени обусловлен движением измерительной системы в неоднородном пространстве. Справедливо и обратное: связанная с движущимся объектом измерительная система, состоящая из пары источников флуктуаций, способна зарегистрировать движение объекта. Нулевой пик, присутствующий на обоих распределениях, рис. 40, не связан с движением измерительной системы. Выяснение его природы требует дальнейших исследований.

2.5.5 Исчезновение суточных периодов при измерениях вблизи Северного полюса.

Зависимость формы гистограмм от вращения Земли вокруг своей оси проявляется в исчезновении околосуточных периодов при измерениях вблизи Северного полюса. Такие измерения были проведены в 2000 г под 82 град с.ш. Исчезли околосуточные периоды при построении гистограмм за 15 или 60 минут. Однако околосуточные периоды изменения вероятности повторного появления 1 - минутных гистограмм данной формы сохранялись. Сохранялась также синхронность по местному времени появления 1-минутных гистограмм сходной формы. [10]

Из этих результатов следовала необходимость проведения измерений возможно ближе к Полюсу. Ввиду невозможности реализации таких измерений были проведены измерения с использованием коллиматоров, вырезающих пучки α -частиц, вылетающих при радиоактивном распаде ^{239}Pu в разных направлениях. Результаты этих опытов привели к принципиальным изменениям наших представлений о природе МФ-эффекта. [9]

2.5.6 О зависимости результатов измерений от направления в пространстве

В 2001 – 2004 г.г. в нашей лаборатории были проведены исследования зависимости формы гистограмм, получаемых при обработке результатов измерений флуктуаций скорости α -распада образцов ^{239}Pu , от направления вылета α -частиц. Измерения осуществляли приборами, в которых полупроводниковые детекторы были установлены после коллиматоров, позволявших считать число альфа частиц, вылетающих при радиоактивном распаде в определенном направлении.

2.5.6.1. Неподвижный коллиматор направленный на Полярную звезду.

Для гистограмм, построенных по результатам подсчета α -частиц, вылетающих при радиоактивном распаде на Север (в направлении на Полярную звезду), околосуточные периоды не наблюдаются и эффекта ближней зоны нет. Эти измерения осуществлялись в Пушино под 54° град. с.ш., а эффект аналогичен, ожидавшемуся под 90° град. с.ш. – на Северном полюсе. Это означает зависимость формы гистограмм от направления в пространстве. Такая зависимость, в свою очередь, влечет за собой вывод об анизотропии пространства. [9, 10]

2.5.6.2. Опыты с неподвижными коллиматорами, направленными на восток и на запад.

Вывод об анизотропии пространства был подкреплён одновременными измерениями с двумя коллиматорами – одним, направленным на восток и другим, направленным на запад. В этих опытах было получено два важных эффекта:

1. Гистограммы, регистрируемые при измерениях с “восточным” коллиматором, с высокой вероятностью реализуются при измерениях с “западным” коллиматором на 718 минут позже, т.е. через половину “звездных” суток.
2. При измерениях с “восточным” и “западным” коллиматорами одновременно, сходные гистограммы не наблюдаются. В опытах без коллиматоров, при измерениях в одном и том же месте, сходные гистограммы с высокой вероятностью наблюдаются в одно и то же время. Этой синхронности нет при одновременных измерениях числа α -частиц, вылетающих через коллиматоры в разных направлениях.

Эти результаты согласуются с представлениями об анизотропии пространства, т.е. о зависимости формы гистограмм от направления вылета α -частиц. [98]

2.5.6.3. Опыты с вращаемыми коллиматорами.

Естественным развитием отмеченных выше исследований были опыты с вращением коллиматоров [99].

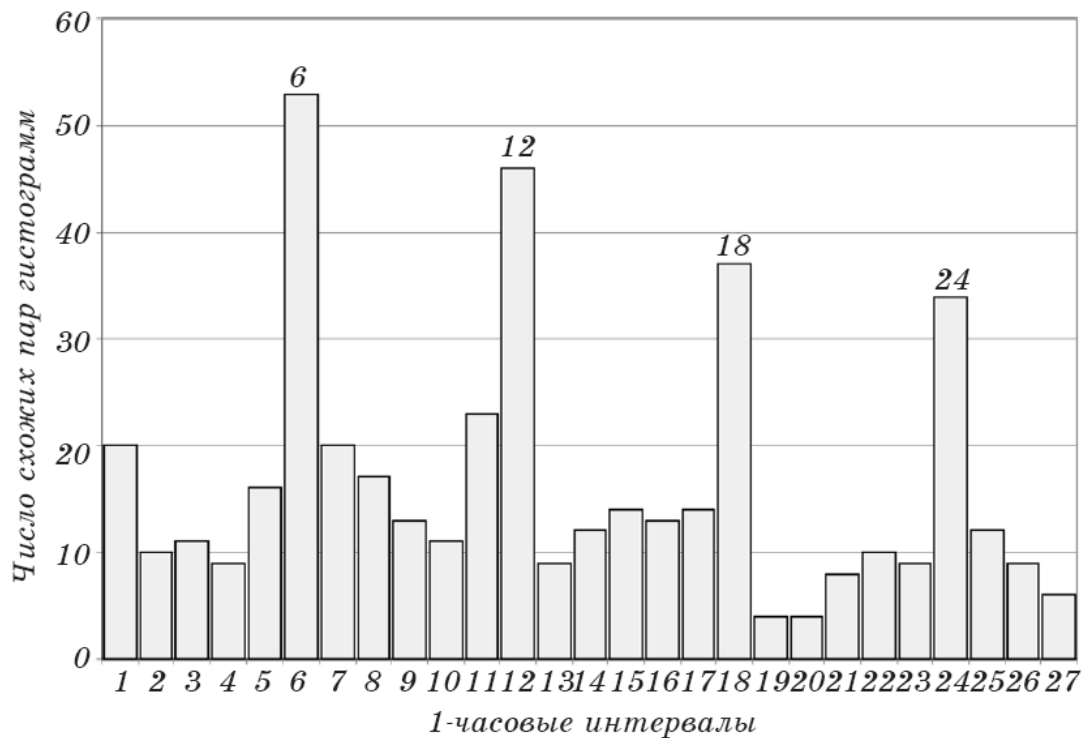


Рис. 41 Распределение интервалов, между парами 60-мин гистограмм, построенное на основе измерений флуктуаций скорости α распада с использованием коллимированного источника ^{239}Pu , совершающего 3 оборота против часовой стрелки, плюс один оборот самой Земли – всего 4 оборота в сутки, т.е. “обзор неба” за 6 часов. Видны четкие 6-часовые периоды. Ось X – 1-часовые интервалы, Y – число схожих пар гистограмм [99].

1. При вращении коллиматора против часовой стрелки (в направлении вращения Земли) сканирование небесной сферы производится с периодами, кратными числу оборотов коллиматора в сутки плюс 1 оборот, совершаемый самой Землей. Было проведено исследование зависимости вероятности появления сходных гистограмм от числа оборотов в сутки вращаемого коллиматора. Было получено, что вероятность появления гистограмм данной формы резко возрастает с периодами, равными 1440 минутам, деленным на число оборотов коллиматора плюс 1. Измеряли периоды при 1, 2, 3, 4, 5, 6, 7, 11 и 23 оборотах коллиматора в сутки. И нашли, соответственно, периоды равные 12, 8, 6 и т.д. часам. При 1-минутном разрешении эти периоды расщепились на два экстремума каждый – “звездный” и “солнечный”. Эти результаты означали, что форма гистограмм, в самом деле, определяется направлением вылета α -частиц [99]. Пример 6-часовых периодов, полученных при трех оборотах против часовой стрелки для последовательности 1-часовых гистограмм, показан на рис. 41.
2. При вращении коллиматора, совершающего 1 оборот в сутки по часовой стрелке, произошла компенсация вращения Земли – α -частицы все время вылетали в направлении к одному и тому же участку небесной сферы – и, соответственно, исчезли суточные периоды. Этот результат оказался полностью аналогичным результатам измерений вблизи Северного полюса и измерений с неподвижным коллиматором, направленным на Полярную звезду [98].

3. При вращении 1 оборот в сутки по часовой стрелке коллиматора, расположенного в плоскости Эклиптики и направленного на Солнце, α -частицы все время вылетают в направлении на Солнце. При этом, как и ожидалось, исчезли околосуточные периоды – и солнечный и звездный.

2.6 Выделенные формы гистограмм: новолуния и солнечные затмения.

Все представленные выше результаты имеют вероятностный характер и были получены в результате десятков тысяч попарных сравнений гистограмм. Совершенно иначе обстоит дело с поиском характерных форм гистограмм в период новолуний и солнечных затмений. В этих случаях смотрится форма гистограмм в заранее рассчитанный момент новолуния или максимума солнечного затмения. При этом было обнаружено, что для этих моментов, практически одновременно, на разных долготах и широтах, по всей Земле, появляются гистограммы характерной, специфической формы. Эти характерные гистограммы соответствуют отрезку временного ряда длиной всего 0.5 – 1 мин. [100, 131] В момент максимума солнечного затмения (как правило, не совпадающего с моментом наступления новолуния) появляется специфическая гистограмма, отличающаяся от гистограммы, появляющейся в моменты новолуний. Естественно, что формы подобные характерным для новолуний или затмений много раз встречаются во временном ряду. Но вероятность их случайного появления именно в эти моменты в разных местах, в разные даты (месяцы, годы) очень мала. Эти специфические формы не связаны с приливными эффектами. Не зависят они и от места на земной поверхности, на которое попадает тень от Луны при затмении или новолунии.

2.7 Выделенные формы гистограмм: восходы и заходы Солнца и Луны.

Форма гистограмм определяется сложным комплексом космо-физических факторов. Как следует из наличия около 27-суточных периодов, среди этих факторов может быть пространственное взаиморасположение и состояние Солнца, Луны и Земли. Мы неоднократно наблюдали сходство гистограмм во времена восходов или заходов Солнца или Луны. Выполнен очень большой объем работы. Однако найти форму гистограмм характерную только для этих моментов не удалось.

2.8 Зеркальная симметрия гистограмм.

Весьма часто (до 30% случаев) форма последовательных гистограмм зеркально симметрична – есть правые и левые формы. При этом, речь может идти об очень сложных формах. [3, 6, 10]

Заключение.

Экспериментальный материал, накопленный за более чем полувековую историю исследования феномена макроскопических флуктуаций, небольшая часть которого представлена выше, дает основания считать, что его феноменология может быть обусловлена только такой общей причиной, как гравитационное взаимодействие, которое, согласно современным воззрениям, определяет структуру пространства-времени. Вследствие этого, эти работы могут рассматриваться как шаг на пути исследования подобной структуры.

Второй момент, который хотелось бы отметить в этой связи – чрезвычайно малая, согласно существующим воззрениям, величина гравитационного воздействия на системы, флукту-

ации которых исследуются. При этом, говоря о воздействии, нужно отметить, что заметная часть рассмотренной феноменологии указывает на то, что данному воздействию должен быть присущ волновой характер. Здесь, в первую очередь, необходимо отметить феномены пространственной направленности и фрактальную структуру пика местного времени, которые находят естественную трактовку в предположении гравитационно-волновой природы действующего фактора. Но, в таком случае, согласно бытующим представлениям, отклик используемых в рассмотренных экспериментах систем должен быть пренебрежимо малым, что входит в кажущееся противоречие с представленным экспериментальным материалом. Выход из подобного противоречия, на наш взгляд, нужно искать в дальнейшем развитии теоретических воззрений на природу гравитационных волн и методов их экспериментального обнаружения, в особенности, с учетом коллективных эффектов в ансамблях частиц, составляющих тело детектора. В этой связи, можно отметить, например, работы [132-134], в которых развит новый подход к детектированию гравитационных волн, а также проанализирована причина неудач существующих попыток их прямого экспериментального наблюдения.

С другой стороны, как было отмечено во Введении, возможен “несиловой” механизм возникновения рассмотренных выше феноменов – за счет движения системы в неоднородном околоземном пространстве. Отмеченная в разделе 2.5.2 универсальность феномена макроскопических флуктуаций, его независимость от природы флуктуирующего процесса, хорошо согласуется с этим подходом. Гравитационное излучение в этом подходе можно, следуя Вейнбергу [135], образно представить, как пространственно-волновую рябь, покрывающую Вселенную.

Благодарности

Автор выражает признательность С.Э.Шнолю за многолетнее сотрудничество и помощь в работе над этим текстом.

Список литературы

- [1] Шноль С.Э., Удальцова Н.В., Агулова Л.П. Корреляции макроскопических флуктуаций в биологических и физико-химических процессах с космогеофизическими факторами // Сб. Электромагнитные поля в биосфере. М., Изд. Наука, 1984, том 1, с.220-224.
- [2] Удальцова Н.В., Коломбет В.А., Шноль С.Э. Возможная космофизическая обусловленность макроскопических флуктуаций в процессах разной природы. Пушино, Изд. НЦБИ АН СССР, 1987 – 96с.
- [3] Шноль С.Э., Коломбет В.А., Пожарский Э.В. Зенченко Т.А., Зверева И.М., Конрадов А.А. О реализации дискретных состояний в ходе флуктуаций в макроскопических процессах // Успехи Физических Наук, 1998, 168(10) с.1129-1140.
- [4] Shnoll S.E., Pozharski E.V., Zenchenko T.A., Kolombet V.A., Zvereva I.M., Konradov A.A. Fine Structure of Distributions in Measurements of Different Processes as Affected by Geophysical and Cosmophysical Factors // Phys. Chem. Earth (A), 1999, Vol.24, No.8, pp.711-714.

- [5] Зенченко Т.А., Пожарский Э.В., Зверева И.М., Коломбет В.А., Конрадов А.А., Шноль С.Э. Тонкая структура распределений результатов измерений процессов разной природы как проявление космофизических влияний // Росс. Хим. Журнал, 1999, т.43, №2, с.3-6.
- [6] Шноль С.Э., Зенченко Т.А., Зенченко К.И., Пожарский Э.В., Коломбет В.А., Конрадов А.А. Закономерное изменение тонкой структуры статистических распределений как следствие космофизических причин // Успехи Физических Наук, 2000, 170(2), с.214-218.
- [7] Шноль С.Э. Макроскопические флуктуации как следствие арифметических и космофизических причин (Факторы, определяющие тонкую структуру гистограмм, возможно, находятся за пределами Солнечной системы) // Биофизика, 2001, т.46, №5, с.775-782.
- [8] Шноль С.Э., Рубинштейн И.А., Зенченко К.И., Зенченко Т.А., Удальцова Н.В., Конрадов А.А., Шаповалов С.Н., Макаревич А.В., Горшков Э.С., Тропичев О.А. Зависимость “макроскопических флуктуаций” от географических координат. (по материалам Арктической (2000г) и антарктической (2001г) экспедиций) // Биофизика, 2003, т.48, №6, с.1123-1131.
- [9] Шноль С.Э., Зенченко К.И., Берулис И.И., Удальцова Н.В., Жирков С.С., Рубинштейн И.А. Зависимость “макроскопических флуктуаций” от космофизических факторов. Пространственная анизотропия // Биофизика, 2004, т.49, №1, с.132-139.
- [10] Шноль С.Э. Закономерные изменения тонкой структуры статистических распределений в случайных процессах, как следствие арифметических и космофизических причин. Труды семинара “Время, хаос и математические проблемы” выпуск 3, М., Книжный Дом “Университет”, 2004, с.121-154.
- [11] Shnoll S.E., Zenchenko K.I., Udaltsova N.V. Cosmophysical Effects in the Structure of Daily and Yearly Periods of Changes in the Shape of Histograms Constructed from the Measurements of ^{239}Pu alpha-Activity // Biophysics, 2004, Vol.49, Suppl.1, pp.155-164.
- [12] Shnoll S.E. Changes in the fine structure of stochastic distributions as consequence of space-time fluctuations // Progress in Physics, 2006, v.6, pp.39-45. (<http://arxiv.org/abs/physics/0602017>)
- [13] Перевертун Т.В., Удальцова Н.В., Коломбет В.А., Иванова Н.П., Брицина Т.Я., Шноль С.Э. Макроскопические флуктуации в водных растворах белков и других веществ как возможное следствие космогеофизических факторов // Биофизика, 1981, т. 26, вып. 4, с.604-614.
- [14] Лазарев П.П. Исследования по адаптации. М.-Л., Изд-во АН СССР, 1947, 260с.
- [15] Жвирблис В. Что нарушает симметрию? // Химия и жизнь, 1977, №12, с.42.
- [16] Жвирблис В.Е. О дрейфе нулевой точки визуального поляриметра // Изв. АН СССР, сер. биологическая, 1982, №3, с.465-468.
- [17] Бурдин К.С., Сизов А.Д. О регистрации синхронных изменений положения фотометрического равновесия в одновременных опытах двух независимых операторов // Биофизика, 1983, т.28, вып. 3., с. 523. (Деп. в ВИНТИ от 30 декабря 1982г. №6514-82)
- [18] Сизов А.Д. О биологической и физической детекции проникающего излучения от звездных источников // Биофизика, 1995, т.40, вып. 4, с.897-904.
- [19] Сизов А.Д. О феномене квантования результатов измерений времени в опытах по исследованию порога чувствительности глаза человека // Биофизика, 1992, Т.37, вып.2 с.395. (Деп. в ВИНТИ от 26 декабря 1991г. №4802-B91)
- [20] Сизов А.Д. К вопросу о космофизической обусловленности результатов измерений дифференциального зрительного порога человека. // Биофизика, 1993, Т.38, вып.3, с.557. (Деп. в ВИНТИ от 23 февраля 1993г. №432-B93)
- [21] Сизов А.Д. О корреляции некоторых биологических и физических характеристик // Биофизика, 1992, Т.37, вып.2 с.394. (Деп. в ВИНТИ от 26 декабря 1991г. №4803-B91)

- [22] Сизов А.Д. О некоторых результатах измерений порога чувствительности глаза человека // Биофизика, 1992, Т.37, вып.2 с.396-396. (Деп. в ВИНТИ от 26 декабря 1991г. №4801-B91)
- [23] Сизов А.Д. О гипотезе биений в механизме возникновения экстремальных значений дифференциального зрительного порога человека // Биофизика, 1992, Т.37, вып.6, с.1116. (Деп. в ВИНТИ от 7 апреля 1992г. №1171-B92)
- [24] Сизов А.Д. Флуктуации значений дифференциального зрительного порога человека в связи с изучением некоторых космофизических характеристик // Биофизика, 1992, Т.37, вып.6, с.1115. (Деп. в ВИНТИ от 7 апреля 1992г. №1170-B92)
- [25] Сизов А.Д. О возможной причине флуктуаций порога чувствительности глаза // Биофизика, 1992, Т.37, вып.6, с.1117. (Деп. в ВИНТИ от 7 апреля 1992г. №1175-B92)
- [26] Сизов А.Д. О сопоставлении длительности некоторых откликов сенсорной системы человека в связи с предположением об их космическом происхождении // Биофизика, 1993, Т.38, вып.3, с.556. (Деп. в ВИНТИ от 23 февраля 1993г. №430-B93)
- [27] Сизов А.Д. О явлении структурированности биофизических данных при использовании галактических координат // Биофизика, 1993, Т.38, вып.3, с.558. (Деп. в ВИНТИ от 23 февраля 1993г. №431-B93)
- [28] Сизов А.Д. О выделенности пространственных направлений как аргументе в пользу космофизической обусловленности некоторых физиологических откликов сенсорной системы человека // Биофизика, 1993, Т.38, вып.3, с.558. (Деп. в ВИНТИ от 23 февраля 1993г. №426-B93)
- [29] Сизов А.Д. К обоснованию способа оценки физиологического отклика при изучении некоторых реакций человека // Биофизика, 1993, Т.38, вып.3, с.556. (Деп. в ВИНТИ от 23 февраля 1993г. №428-B93)
- [30] Сизов А.Д. О модели физического устройства для демонстрации эффектов, связанных с появлением экстремальных значений дифференциального зрительного порога человека // Биофизика, 1992, Т.37, вып.6, с.1116. (Деп. в ВИНТИ от 7 апреля 1992г. №1172-B92)
- [31] Сизов А.Д. Флуктуации тока в мостике Уинстона. Возможные космофизические корреляции // Биофизика, 1998, т.43, вып.4, с.726-729.
- [32] Сизов А.Д. Несколько предположений по поводу биоиндикации неидентифицированного излучения звезд // Биофизика, 1993, Т.38, вып.3, с.558. (Деп. в ВИНТИ от 23 февраля 1993г. №427-B93)
- [33] Сизов А.Д. Возможная связь “аномальных” сигналов в мостике Уитстона с положением планет // Биофизика, 2001, Т.46, вып.5, с.811-815.
- [34] Сизов А.Д. Об информационной роли некоторых биофизических показателей в связи с возможностью получения астрономической информации // Биофизика, 1992, Т.37, вып.6, с.1117. (Деп. в ВИНТИ от 7 апреля 1992г. №1173-B92)
- [35] Сизов А.Д. К вопросу определения межзвездных расстояний безтелескопным методом // Биофизика, 1993, Т.38, вып.3, с.557. (Деп. в ВИНТИ от 23 февраля 1993г. №429-B93)
- [36] Сизов А.Д. К вопросу о существовании биофизического канала получения астрономической информации // Биофизика, 1992, Т.37, вып.6, с.1117. (Деп. в ВИНТИ от 7 апреля 1992г. №1174-B92)
- [37] Козырев Н.А. Астрономические наблюдения посредством физических свойств времени // Вспыхивающие звезды: Труды симпозиума, приуроченного к открытию 2.6-м телескопа Бюраканской астрофизической обсерватории, Бюракан, 5-8 октября 1976 года, Ереван, 1977, с.209-227; Избранные труды. Л., Изд-во Ленинградского ун-та, 1991, с.363-383.

- [38] Козырев Н.А., Насонов В.В. Новый метод определения тригонометрических параллаксов на основе измерения разности между истинным и видимым положением звезды // Астрометрия и небесная механика. Сер.: Проблемы исследования Вселенной. Вып.7, М., Л., 1978, с.168-179.
- [39] Лаврентьев М.М., Гусев В.А., Еганова И.А., Луцет М.К., Фоминых С.Ф. О регистрации истинного положения Солнца // Доклады АН СССР, 1990, т.315, №2, с.368-370.
- [40] Лаврентьев М.М., Еганова И.А., Луцет М.К., Фоминых С.Ф. О дистантном воздействии звезд на резистор // Доклады АН СССР, 1990, т.314, №2, с.352-355.
- [41] Лаврентьев М.М., Еганова И.А., Луцет М.К., Фоминых С.Ф. О регистрации реакции вещества на внешний необратимый процесс // Доклады АН СССР, 1991, т.317, №3, с.635-639.
- [42] Акимов А.Е., Ковальчук Г.У., Медведев В.Г., Олейник В.К., Пугач А.Ф. Предварительные результаты астрономических наблюдений по методике Н.А.Козырева // Препринт ГАО АН Украины, Киев, 1992, №92-5Р, 16с.
- [43] Dubrovskiy V.A. A correlation of a microseism background with space object distributions // Moscow University Mechanics Bulletin, 2004, Vol.59, No.5, pp.9-12.
- [44] Дубровский В.А. Упругая модель физического вакуума // Докл. АН СССР, 1985, 282, 83-88.
- [45] Dubrovskiy V.A. Measurements of the gravity waves velocity // astro-ph/0106350, 2001, 6р.
- [46] Личная переписка.
- [47] Аллен К.У. Астрофизические величины. М., Мир, 1977.
- [48] Куликовский П.Г. Справочник астронома-любителя. М., Физматгиз, 1961.
- [49] Соколовский В.В. О биохимическом механизме реакций живых организмов на изменения солнечной активности // Проблемы космической биологии, 1982, т.43, с.180-193.
- [50] Горшков Э.С., Шаповалов С.Н., Соколовский В.В., Трошичев О.А. О гравитационной обусловленности флуктуаций скорости реакции окисления унитиола нитритным ионом // Биофизика, 2000, т.45, вып.4, с.631-635.
- [51] Шаповалов С.Н., Горшков Э.С., Борисова Т.Д., Соколовский В.В., Трошичев О.А. Случайные флуктуации в показаниях измерительных приборов: эффекты космофизического влияния? // Биофизика, 2001, т.46, вып.5, с.819-822.
- [52] Удальцова Н.В., Коломбет В.А., Шноль С.Э. Возможная космофизическая обусловленность макроскопических флуктуаций в процессах разной природы. Пущино, НЦБИ АН СССР, 1987, 96с.
- [53] Космические лучи и солнечный ветер. Под. ред. Г.В.Шафера. Новосибирск, Наука, 1981, 221с.
- [54] Ресселл Г.Н., Дэган Р.С., Стюарт Д.К. Астрономия. т.1. Солнечная система. М., Л., ОНТИ, 1934, 383с.
- [55] Ivanov V.V., Sokolovsky V.V., Gorshkov E.S., Shapovalov S.N., Troshichev O.A. Periodical Changes in the Human Urinary Thiol Content as Related with Cosmophysical Factors // Biophysics, 2004, Vol.49, Suppl.1, pp.S115-S118.
- [56] Гедерим В.В., Соколовский В.В., Горшков Э.С., Шаповалов С.Н., Трошичев О.А. Периодические изменения некоторых гематологических показателей, характеризующих процесс адаптации в организме человека и вариации гравитационного поля // Биофизика, 2001, т.46, вып.5, с.833-834.
- [57] Sokolovsky V.V., Gorshkov E.S., Ivanov V.V., Shapovalov S.N., Troshichev O.A. Relation of the Regular Gravitational Field Variations to Biochemical Processes Observed In Vitro and In Vivo // Biophysics, 2004, Vol.49, Suppl.1, pp.S85-S91.

- [58] Troshichev O.A., Gorshkov E.S., Shapovalov S.N., Sokolovskii V.V., Ivanov V.V., Vorobeitchikov V.M. Variation of the gravitational field as a motive power for rhythmicity of biochemical processes // *Advances in Space Research* 34 (2004) 1619-1624.
- [59] Shapovalov S.N., Gorshkov E.S., Troshichev O.A., Borisova T.D., Frank-Kamenetsky A.V. Effects of Non-electromagnetic Disturbances from the Sun in “Computer Time” Instability // *Biophysics*, 2004, Vol.49, Suppl.1, pp.S79-S84.
- [60] Shapovalov S.N., Gorshkov E.S., Troshichev O.A. Cosmophysical Effects Observed in Impulses of the Microphotocolorimeter Current // *Biophysics*, 2004, Vol.49, Suppl.1, pp.S119-S122.
- [61] Горшков Э.С., Шаповалов С.Н., Соколовский В.В., Трошичев О.А. О детектировании импульсного космофизического излучения // *Биофизика*, 2000, т.45, вып.5, с.947-949.
- [62] Горшков Э.С., Бондаренко Е.Г., Шаповалов С.Н., Соколовский В.В., Трошичев О.А. Использование детектора сверхслабых излучений в качестве вариометра гравитационного поля // *Биофизика*, 2001, т.46, вып.5, с.816-818.
- [63] Владимирский Б.М., Темурьянц Н.А. Влияние солнечной активности на биосферу-ноосферу (Гелиобиология от А.Л.Чижевского до наших дней). М., Изд-во МНЭПУ, 2000, 374с.
- [64] Клочек Н.В., Паламарчук Л.Э., Плюснина Л.А., Никонова М.В. К вопросу о космическом воздействии неизвестной природы // *Биофизика*, 1992, т.37, вып.4, с.656-660.
- [65] Клочек Н.В., Паламарчук Л.Э., Никонова М.В. Предварительные результаты исследований воздействия космофизического излучения неэлектромагнитной природы на физические и биологические системы // *Биофизика*, 1995, 40(4), 889-896.
- [66] Башков В.И., Балакин А.Б., Гусев А.В. Экспериментальные тесты теории гравитации. Под ред. Брагинского В.Б. и Денисова В.И. М., МГУ, 1989, с.56.
- [67] Долгих Г.И., Копвиллем У.Х., Хаврошкин О.Б., Цыплаков В.В. Об одном физическом механизме возмущения атомных стандартов частоты. Владивосток, 1979, 37с. (Депонировано в ВИНТИ, №3070-79)
- [68] Морозов А.Н. Пространственная анизотропия флуктуаций подвижности ионов // Выступление на семинаре Фонда содействия развитию финслеровой геометрии, 25.11.2006.
- [69] Козырев Н.А. Избранные труды. Л., Изд-во Ленинградского ун-та, 1991, 448 с.
- [70] Богданович Б.Ю., Щедрин И.С., Смирнов В.Н., Егоров Н.В. Особый способ вращения массы – инструмент для астрофизических исследований // Научная сессия МИФИ, 2003, т.7, 45-46.
- [71] Богданович Б.Ю., Щедрин И.С., Смирнов В.Н., Егоров Н.В. Предварительные аналитические оценки изменения кинетической энергии вращающейся массы от координатно-временного положения Солнца и Луны // Научная сессия МИФИ, 2003, т.7, 47-48.
- [72] Богданович Б.Ю., Егоров Н.В., Кулага А.П., Смирнов В.Н. Регистрация детектором гравитационных взаимодействий различной орбитальной конфигурации планет солнечной системы // Научная сессия МИФИ, 2006, т.7, 101-104.
- [73] Смирнов В.Н., Егоров Н.В. Моделирование приемника информации для исследования волновых процессов // *Стабильность и процессы управления*, 2005, pp.226-235.
- [74] Смирнов В.Н. Патент на изобретение №2172501 “Способ и устройство для измерения изменения состояния вращающегося волчка (Пространственно временной геометризатор)”. Приоритет от 03.06.1999. Зарегистрирован в Гос. реестре изобретений РФ, г.Москва 20.08.2001.
- [75] Физическая энциклопедия. т.1. М., “Советская энциклопедия”, 1998, 704с.
- [76] Aharonov Y., Bohm D. Significance of electromagnetic potentials in the quantum theory // *Phys. Rev.*, 1959, v.115, p.485-491.

- [77] Фейнберг Е.Л. Об “особой роли” электромагнитных потенциалов в квантовой механике // Успехи Физических Наук, 1962, т.78, вып.9, с.53-64
- [78] Спасский Б.И.,Московский А.В. О нелокальности в квантовой физике // Успехи Физических Наук, 1984, т.142, вып.4, с.599-617.
- [79] Труханов К.А. Векторный потенциал электромагнитного поля // В кн.: Электромагнитные поля в биосфере. Т.1. Электромагнитные поля в атмосфере Земли и их биологическое значение. М., Наука, 1984, с.331-350.
- [80] Труханов К.А. О возможной роли эффекта Ааронова-Бома в биологическом действии магнитного поля // В кн.: Физико-математические и биологические основы действия ЭМП и ионизация воздуха. М., Наука, 1975, Т.1, с.151-152.
- [81] Труханов К.А. О возможной роли эффекта Ааронова-Бома в биологическом действии магнитного поля // Космическая биология и медицина, 1978, Т.12, №3, с.82-83.
- [82] Бауров Ю.А.,Труханов К.А. Возможная роль космологического векторного потенциала как фактора космо- и гелиофизических связей // Биофизика, 1998, т.43, вып.5, с.928-934.
- [83] Бауров Ю.А. Структура физического пространства и новый способ получения энергии. М., Изд-во “Кречет”, 1998 – 240с.
- [84] Бауров Ю.А.,Яковенко В.А.,Комиссаров А.В.,Вержиковский В.Г., Конрадов А.А. Экспериментальное исследование нового информационного канала в природе, обусловленного квантовыми свойствами физического пространства (вакуума) с помощью кварцевого резонатора // Биофизика, 2001, т.46, вып.5, с.823-828.
- [85] Baurov Yu.A.,Yakovenko V.A.,Komissarov A.V.,Verzhikovski V.G., Baurov A.Yu.,Konradov A.A.,Zenchenko T.A. Preliminary results of an experimental investigation of a new information channel in nature with the aid of quartz resonators’ system // International Journal of Scientific Research, 2006, Vol.16, pp.469-473.
- [86] Манчестер Р.,Тейлор Дж. Пульсары. М., Мир, 1980 – 292с.
- [87] Пархомов А.Г.,Макляев Е.Ф. Исследование ритмов и флуктуаций при длительных измерениях радиоактивности, частоты кварцевых резонаторов, шума полупроводников, температуры и атмосферного давления // Физическая мысль России, №1, 2005.
- [88] Пархомов А.Г. Экспериментальные исследования инфранизкочастотных флуктуаций в полупроводниках. Закономерности. Космические ритмы. М.: МНТЦ ВЕНТ, 1991, 24 с.
- [89] Пархомов А.Г. Вариации интенсивности низкочастотных флуктуаций в проводниках //Атлас временных вариаций природных, антропогенных и социальных процессов, т. 2, М., Научный мир, 1998, с. 310-312.
- [90] Эйби Дж. Землетрясения. М., Недра, 1982 - 264 с.
- [91] Пархомов А.Г. Распределение и движение частиц скрытой материи. М., МНТЦ ВЕНТ, 1993 - 76с.
- [92] Пархомов А.Г. Необычное космическое излучение. Обнаружение, гипотезы, проверочные эксперименты. М., МНТЦ ВЕНТ, 1995 – 51с.
- [93] Пархомов А.Г. Астрономические наблюдения по методике Козырева и проблема мгновенной передачи сигнала // Физическая мысль России, 2000, №1, с.18-25.
- [94] Шноль С.Э.,Удальцова Н.В.,Коломбет В.А.,Намиот В.А.,Бодрова Н.Б. Закономерности в дискретных распределениях результатов измерений (космофизические аспекты) // Биофизика, 1992, 37(3), 467-488.

- [95] Шноль С.Э., Намиот В.А., Жвирблис В.Е., Морозов В.Н., Темнов А.В., Морозова Т.Я. Возможная общность макроскопических флуктуаций скоростей биохимических и химических реакций, электрофоретической подвижности клеток и флуктуаций при измерениях радиоактивности, оптической активности и фликкерных шумов // Биофизика, 1983, 28(1), 153-157.
- [96] Шноль С.Э. Форма спектров состояний, реализуемых в ходе макроскопических флуктуаций, зависит от вращения Земли вокруг своей оси // Биофизика, 1995, 40(4), 865 - 875.
- [97] Shnoll S.E., Zenchenko K.I., Udaltsova N.V. Cosmo-physical effects in structure of the daily and yearly periods of change in the shape of the histograms constructed by results of measurements of alpha-activity Pu²³⁹. <http://arxiv.org/abs/physics/0504092>
- [98] Shnoll Simon E., Zenchenko Konstantin I., Berulis Iosias I., Udaltsova Natalia V., Rubinstein Iia A. Fine structure of histograms of alpha-activity measurements depends on direction of alpha particles flow and the Earth rotation: experiments with collimators. <http://arxiv.org/abs/physics/0412007>
- [99] Shnoll S.E., Rubinshtejn I.A., Zenchenko K.I., Shlekhtarev V.A., Kaminsky A.V., Konradov A.A., Udaltsova N.V. Experiments with Rotating Collimators Cutting out Pencil of α -Particles at Radioactive Decay of ²³⁹Pu Evidence Sharp Anisotropy of Space // Progress in Physics, V.1, 2005, pp.81-84. <http://arxiv.org/abs/physics/0501004>
- [100] Shnoll S.E., Zenchenko K.I., Shapovalov S.N., Gorshkov S.N., Makarevich A.V. and Troshichev O.A. The specific form of histograms presenting the distribution of data of α -decay measurements appears simultaneously in the moment of New Moon in different points from Arctic to Antarctic. <http://arxiv.org/abs/physics/0412152>
- [101] Шноль С.Э. О самопроизвольных синхронных переходах молекул актомиозина в растворе из одного состояния в другое // Вопр. мед. химии, 1958, т.4, вып.6, с.443-454.
- [102] Шноль С.Э. Синхронные обратимые изменения (“конформационные колебания”) препаратов мышечных белков // Дисс. на соискание уч. степени д.б.н., Пущино, 1969 –307с.
- [103] Шноль С.Э. Синхронные в макрообъеме колебания АТФ-азной активности в концентрированных препаратах актомиозина // В кн.: Колебательные процессы в биологических и химических системах. Пущино, 1971, с.20-24.
- [104] Белоусов Б.П. Периодически действующая реакция и ее механизм // В сб. рефератов по радиационной медицине за 1958 год. М., Медгиз, 1959, с.145-147.
- [105] Жаботинский А.М. Периодический ход окисления малоновой кислоты в растворе (исследование кинетики реакции Белоусова) // Биофизика, 1964, т.9, с.306-311.
- [106] Шноль С.Э., Смирнова Н.А. Колебания концентрации SH-групп в растворах актомиозина, актина и миозина // Биофизика, 1964, т.9, вып.4, с.532–534.
- [107] Шноль С.Э. “Синхронные конформационные колебания молекул актина, миозина и актомиозина в растворах” в сб. Молекулярная биофизика, М., Изд. “Наука”, 1965, с. 56 – 81
- [108] Shnoll S.E., Chetverikova E.P. Synchronous reversible alterations in enzymatic activity (conformational fluctuations) in actomyosin and creatine kinase preparations // Biochem. Biophys. Acta, 1975, v.403, p.89-97
- [109] Шноль С.Э., Четверикова Е.П., Рыбина В.В. Синхронные в макрообъеме конформационные колебания в препаратах белков актомиозинового комплекса и в растворах креатинкиназы // В сб. “Молекулярная и клеточная биофизика”. М., Изд. Наука, 1977 с.79-92.
- [110] Шноль С.Э., Рыбина В.В. Однотипность конформационных колебаний макромолекул белков актомиозинового комплекса и креатинкиназы // В сб. “Биофизические основы и регуляция процессов мышечного сокращения”. М., Изд. “Наука”, 1977.

- [111] Шноль С.Э., Рыбина В.В. Синхронные конформационные колебания титра сульфгидрильных групп в растворах белков. Обратимое окисление как возможная причина этого явления // Биофизика, 1979, т.24, вып.6, с.970-976.
- [112] Шноль С.Э., Намиот В.А., Жвирблис В.Е., Морозов В.Н., Темнов А.В., Морозова Т.Я. Возможная общность макроскопических флуктуаций скоростей биохимических и химических реакций, электрофоретической подвижности клеток и флуктуаций при измерениях радиоактивности, оптической активности и фликкерных шумов // Биофизика, 1983, т.28, вып.1, с.153-157.
- [113] Шноль С.Э. Макроскопические флуктуации с дискретным распределением амплитуд в процессах различной природы // В кн.: Итоги Науки и Техники. Молекулярная биология. Ред. В.П.Скулачев, М., ВИНТИ, 1985, т.5, с.130-200.
- [114] Зенченко К.И., Зенченко Т.А., Кужевский Б.М., Вилкен Б., Аксфорд Я., Шноль С.Э. Форма гистограмм, построенных по измерениям α -распада ^{228}Ra в Линдау (Германия) и потока нейтронов в Москве, изменяется синхронно по местному времени // Биофизика, 2001, т.46, №5, с.783-785.
- [115] Fedorov M.V., Belousov L.V., Voeikov V.L., Zenchenko T.A., Zenchenko K.I., Pozharskii E.V., Konradov A.A., Shnoll S.E. Synchronous Changes in Dark Current Fluctuations in Two Separate Photomultipliers in Relation to Earth Rotation // Astrophysics & Space Science, 2003, №1 pp.105-112.
- [116] Панчелюга В.А., Шноль С.Э. Феноменология эффекта местного времени на малых пространственно-временных масштабах и в случае движущихся источников флуктуаций // Метафизика. Век XXI. Вып.2, М., БИНОМ. Лаборатория знаний, 2007, с.320-326.
- [117] Панчелюга В.А., Коломбет В.А., Каминский А.В., Панчелюга М.С., Шноль С.Э. Эффект местного времени в шумовых процессах. // Вестник Калужского университета. 2006, №2, с.3-8.
- [118] Панчелюга В.А., Коломбет В.А., Панчелюга М.С., Шноль С.Э. Исследование эффекта местного времени на малых пространственно-временных масштабах // Гиперкомплексные числа в геометрии и физике, 1(5), Vol.3, 2006, с.116-121.
- [119] Panchelyuga V.A., Kolombet V.A., Panchelyuga M.S., Shnoll S.E. Local-time effect on small space-time scale // Space-Time Structure. Collected papers. Editor-in-chif Pavlov D.G. M.: TETRU, 2006. – pp.344–350.
- [120] Panchelyuga Victor A., Kolombet Valery A., Panchelyuga Maria S., Shnoll Simon E. Experimental Investigations of the Existence of Local-Time effect on the Laboratory Scale and the Heterogeneity of Space-Time // Progress in Physics, V.1, January, 2007, pp.64-69.
- [121] Шноль С.Э., Панчелюга В.А. Феномен макроскопических флуктуаций. Методика измерений и обработки экспериментальных данных // Мир измерений, 2007, №6, с.49-55.
- [122] Панчелюга В.А., Шноль С.Э. Экспериментальное исследование влияния быстро вращающегося массивного тела на форму функций распределения амплитуд флуктуаций скорости α -распада // Гиперкомплексные числа в геометрии и физике, 1(5), Vol.3, 2006, с.102-115.
- [123] Panchelyuga V.A., Shnoll S.E. Experimental investigation of spinning massive body influence on fine structure of distribution functions of alpha-decay rate fluctuations // Space-Time Structure. Collected papers. Editor-in-chif Pavlov D.G. M.: TETRU, 2006, pp.328–343.
- [124] Шноль С.Э. Космофизические корреляции в процессах разной природы // Вестник Калужского университета. 2006, №1, с.33-40.
- [125] Панчелюга В.А., Шноль С.Э. Результаты исследований эффекта местного времени на малых пространственно-временных масштабах // VII Международная крымская конференция “Космос и биосфера”. Тезисы докладов. Судак, Крым, Украина, 1–6 октября, 2007, с.55-56.

- [126] Панчелюга В.А.,Шноль С.Э. О пространственной анизотропии выявляемой при исследовании “эффекта местного времени” // Гиперкомплексные числа в геометрии и физике. 2006, №2 (6), Vol.3 с.188-193.
- [127] Panchelyuga Victor A., Shnoll Simon E. Space-Time Anisotropy revealed during investigation of various processes fluctuations // ADVANCES IN MODERN NATURAL SCIENCES. Proceedings of 3rd International Conference INTERNAS’2007, Kaluga, Russia, May 22-25, 2007, pp.23-27.
- [128] Panchelyuga Victor A., Shnoll Simon E. On the Dependence of a Local-Time Effect on Spatial Direction // Progress in Physics, July 2007, V.3, pp.51-54.
- [129] Физические величины. Справочник. // Под ред. И.С.Григорьева, Е.З.Мейлихова. - М., Энергоатомиздат, 1991 – 1232с.
- [130] Panchelyuga Victor A., Shnoll Simon E. A Study of a Local Time Effect on Moving Sources of Fluctuations // Progress in Physics, July, 2007 V.3, pp.55-56.
- [131] Shnoll S.E.,Panchelyuga V.A. On the characteristic form of histograms appearing at the culmination of solar eclipse // physics/0603029, 2006, 11p.
- [132] Borissova L. Gravitation Waves and Gravitational Inertial Waves in the General Theory of Relativity: A Theory and Experiments // Progress in Physics, V.2, 2005, pp.30-62.
- [133] Borissova L.B., Rabounski D.D. Fields, Vacuum, and the Mirror Universe. Editorial URSS, Moscow, 2001, 272 p. (the 2nd revised ed.: CERN, EXT-2003-025)
- [134] Rabounski D., Borissova L. Exact Theory to a Gravitational Wave Detector. New Experiments Proposed // Progress in Physics, V.2, 2006, pp.31-38.
- [135] Вейнберг С. Гравитация и космология. Принципы и приложения общей теории относительности. М., Мир, 1975 – 696с.

Научное издание

Основы финслеровой геометрии и ее приложения в физике

Материалы Международной школы-семинара для
старшекурсников, аспирантов физико-математических
факультетов и молодых ученых

Москва, 12 июля – 14 августа, 2010 г.

Подписано в печать 23.06.10 Формат 70x108/16

Усл.печ.л. 36.05

Тираж 300 экз. Заказ № 367