

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

В.А. Авдюшев

**ЧИСЛЕННОЕ
МОДЕЛИРОВАНИЕ
ОРБИТ НЕБЕСНЫХ ТЕЛ**

Томск
Издательский Дом Томского государственного университета
2015

УДК 521
ББК 22.62
A186

Авдюшев В.А.

A186 Численное моделирование орбит небесных тел. —
Томск : Издательский Дом Томского государственного университета, 2015. — 336 с.

ISBN 978-5-94621-519-0

В монографии представлен вычислительно-математический инструментарий, предназначенный для численного моделирования орбит небесных тел. Рассматриваются широко используемые на практике численные методы для решения прямой и обратной задач орбитальной динамики. Исследуется их эффективность применительно к численному моделированию орбит малых тел Солнечной системы.

Для специалистов в области небесной механики, а также студентов и аспирантов астрономических специальностей.

УДК 521
ББК 22.62

Рецензент
доктор физ.-мат. наук, проф. **Т.В. Бордовицына**

ISBN 978-5-94621-519-0 © Авдюшев В.А., 2015
© Томский государственный университет, 2015

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	8
СПИСОК ОБОЗНАЧЕНИЙ	13
1. ЧИСЛЕННАЯ МОДЕЛЬ ОРБИТЫ	15
1.1. Терминология	15
1.2. Ошибки численного моделирования	19
1.3. Быстродействие численного моделирования	22
1.4. Эффективность численного моделирования	23
2. МОДЕЛИ СИЛ	25
2.1. Гравитация протяженного тела	26
2.2. Гравитация материальной точки	29
2.3. Радиационные силы	30
2.4. Реактивные силы	32
2.5. Атмосферное торможение	33
2.6. Релятивистские возмущения	34
2.7. Сравнительный анализ силовых факторов	35
3. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ОРБИТ	38
3.1. Линейные и регулярные уравнения	39
3.1.1. Уравнения в SB-переменных	43
3.1.2. Уравнения в KS-переменных	44
3.2. Сглаженные уравнения	48
3.3. Стабилизированные уравнения	53
3.3.1. Неустойчивость кеплеровского движения	57
3.3.2. Диссипативная стабилизация Баумгарте	59
3.3.3. Консервативная стабилизация Баумгарте	64
3.3.4. Стабилизация по времени	65
3.3.5. Стабилизация Накози	66

3.4.	Уравнения в орбитальных элементах	66
3.4.1.	Уравнения Роя	68
3.4.2.	Уравнения в SB-элементах	69
3.4.3.	Уравнения в KS-элементах	69
3.4.4.	Уравнения в элементах Пайнса	70
3.5.	Уравнения Энке	71
3.5.1.	Классические уравнения Энке	73
3.5.2.	Уравнения Энке в KS-переменных	75
3.5.3.	Улучшение опорной орбиты	76
3.6.	Гравицентрические системы координат и уравнения синхронного слежения	79
3.6.1.	Ошибки округления в кеплеровских членах. Задача двух тел	80
3.6.2.	Ограниченная задача трех тел	84
3.6.3.	Временные ошибки	87
3.6.4.	Метод синхронного слежения	91
3.6.5.	Эффективность гравитационных координатных систем	93
3.6.6.	Влияние ошибок в положении планеты на точность моделирования астероидной орбиты	97
3.7.	Проблема короткопериодических возмущений и усредненные уравнения	100
3.8.	Сравнительный анализ эффективности уравнений	110
3.8.1.	Численный эксперимент	110
3.8.2.	Выбранные объекты	113
3.8.3.	Численные результаты	115
4.	МЕТОДЫ ИНТЕГРИРОВАНИЯ	124
4.1.	Терминология	126
4.2.	Метод разложения в ряд Тейлора	127
4.3.	Методы Рунге–Кутты	130
4.3.1.	Явные методы Рунге–Кутты	130
4.3.2.	Оценка погрешности и выбор шага	134
4.3.3.	Вложенные методы Рунге–Кутты	142
4.3.4.	Неявные методы Рунге–Кутты	144

4.3.5.	Коллокационные методы	146
4.3.6.	Методы Гаусса	150
4.3.7.	Интегратор Эверхарта	152
4.3.8.	Порядок и шаг интегрирования при компьютерной реализации метода	161
4.3.9.	Орбитальная устойчивость методов Рунге–Кутты	164
4.4.	Экстраполяционные методы	166
4.4.1.	Общий подход	166
4.4.2.	Алгоритм Эйткена–Невилла	168
4.4.3.	Метод Грэгга	169
4.4.4.	Выбор шага	170
4.5.	Многошаговые методы	170
4.5.1.	Методы Адамса	171
4.5.2.	Формулы дифференцирования	173
4.5.3.	Предиктор–корректор	174
4.5.4.	Линейные многошаговые методы	175
4.5.5.	Порядок многошаговых методов	176
4.5.6.	Устойчивость многошаговых методов	178
4.5.7.	Наивысший достижимый порядок для устойчивых методов	179
4.5.8.	Практическая оценка локальной погрешности	180
4.5.9.	Выбор шага	180
4.5.10.	Порядок и шаг интегрирования при компьютерной реализации методов Адамса	183
4.6.	Геометрические методы	185
4.6.1.	Уравнения гармонического осциллятора	185
4.6.2.	Методы Эйлера	186
4.6.3.	Проекционный метод	189
4.6.4.	Симплектические и симметричные методы	191
4.7.	Сравнительный анализ эффективности методов	203
4.7.1.	Численный эксперимент	204
4.7.2.	Выбранные методы	205
4.7.3.	Численные результаты	206

5.	ОПРЕДЕЛЕНИЕ ОРБИТ ИЗ НАБЛЮДЕНИЙ	213
5.1.	Обратная задача орбитальной динамики	213
5.2.	Модельные представления наблюдений	216
5.3.	Предварительное определение орбиты	217
5.4.	Изохронные производные	220
5.5.	Методы решения обратной задачи	223
5.5.1.	Метод градиентного спуска	224
5.5.2.	Метод Ньютона	224
5.5.3.	Метод Гаусса–Ньютона	225
5.5.4.	Демпфированный метод Гаусса–Ньютона	226
5.5.5.	Метод Левенберга–Марквардта	226
5.5.6.	Овражные методы	227
5.6.	Сравнительный анализ эффективности методов	232
5.7.	Проблема неоднозначного определения орбит	237
5.7.1.	Наблюдения на длительном временном интервале	238
5.7.2.	Наблюдения на короткой орбитальной дуге	253
6.	ОЦЕНИВАНИЕ ПАРАМЕТРИЧЕСКОЙ ТОЧНОСТИ	258
6.1.	Задача наименьших квадратов, ее геометрия и доверительное оценивание	259
6.2.	Моделирование доверительной области	264
6.3.	Моделирование возможных значений параметров	267
6.3.1.	Линейные методы	267
6.3.2.	Нелинейные методы	269
6.4.	Взвешенная задача наименьших квадратов	272
6.5.	Бутстрэп-метод	274
6.6.	Показатели нелинейности	276
6.7.	Сравнительный анализ эффективности методов	278
7.	ЛИНЕАРИЗАЦИЯ ОРБИТАЛЬНОЙ МОДЕЛИ	294
7.1.	Временной перенос доверительной области	295
7.2.	Вычисление ляпуновских показателей	296
7.3.	Оценивание вероятности попадания объекта в малый объем	298

7.4. Оценивание вероятности столкновения астероида с планетой	300
ПРИЛОЖЕНИЕ	303
ЛИТЕРАТУРА	314
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ	328

Числа показывают, как управляется мир.
И. Гете

Используйте только то, что работает.
Брюс Ли

ПРЕДИСЛОВИЕ

Численное моделирование является мощным средством для исследования орбит небесных тел. Оно стало уже почти традиционным в современной практике. Значение численного моделирования орбит трудно переоценить, поскольку его приложение очень широко. Орбиты в компьютерной миниатюре позволяют прогнозировать движение небесных тел в будущем и выявлять обстоятельства их происхождения в прошлом, проектировать космические миссии и оценивать вероятности столкновения астероидов с планетами, предсказывать с высокой точностью затменные явления и определять местоположения наземных объектов по навигационным спутниковым системам. Причем эти возможности открываются фактически для каждого, у кого просто есть компьютер, независимо, супер он или нет¹. Страстно жаждущему макрокосмических исследований в микрокосмосе цифр остается только освоить соответствующий вычислительно-математический инструментарий и научиться им пользоваться на практике. Так вот, данная монография как раз и представляет этот инструментарий.

Стоит отметить, что трактат фактически является результатом значительной переработки предыдущей монографии автора (Авдюшев, 2010а). До нее последняя фундаментальная работа на ту же тему (Бордовицына, 1984) вышла в свет еще в прошлом веке, в начале восьмидесятых, почти четверть века назад. К слову, она до сих пор пользуется большим спросом среди специа-

¹К настоящему времени вполне приличный компьютер уже не роскошь. Кстати, почти все численные результаты, представленные в монографии, были получены на малобюджетном ноутбуке. Между тем, что касается программного обеспечения, то среди необходимых для моделирования приложений сейчас достаточно много бесплатных.

листов. Впоследствии все же предпринималась попытка переиздать основную ее часть в рамках учебного пособия (Бордовицына, Авдюшев, 2007). Однако соответствующий материал относительно численного моделирования не был подвергнут серьезной редакции и, кроме того, в прикладном плане ограничивался только рассматриваемыми в пособии прямыми задачами орбитальной динамики искусственных спутников Земли. Собственно, эти обстоятельства и побудили автора написать свое предшествующее произведение (Авдюшев, 2010а).

Настоящая монография, как и предыдущая, не претендует на всеобъемлющий обзор требуемого для моделирования математического аппарата. Тем не менее, по мнению автора, изложенного материала здесь вполне достаточно для освоения, по крайней мере, основ численного моделирования орбит. В любом случае при необходимости в получении более детальной информации читатель может обратиться к цитируемой литературе, приведенной в конце монографии.

Образно говоря, любая численная модель должна быть столь же эффективной, сколь эффектна и грациозна топ-модель на подиуме: иные модели просто не востребованы ни в науке, ни в модельном шоу-бизнесе. Поэтому автор предпочел представить в изложении именно тот «концентрат», который действительно работает эффективно с точки зрения численного моделирования и уже одобрен широким кругом специалистов.

Конечно, внимательный читатель наверняка заметит особые предпочтения автора к задачам динамики тел Солнечной системы, что вовсе не удивительно, учитывая, что большая часть его работ была выполнена в рамках этой тематики. Впрочем, излагаемый инструментарий вполне может быть применен также и к решению иных задач динамической астрономии, например звездной или галактической динамики (Aarseth, 2003; Орлов, Рубинов, 2008), которые не рассматриваются в монографии.

Безусловно, для адекватного восприятия материала от читателя требуется специальная подготовка, которую, впрочем, можно получить за государственный счет в вузе по прошествии двух

общеобразовательных курсов на каком-либо естественнонаучном факультете, не игнорирующем, правда, астрономические дисциплины, либо выработать самостоятельно, проштудировав, по меньшей мере, несколько фундаментальных работ по небесной механике и вычислительной математике.

Содержательная часть книги состоит из семи глав. Первая глава вводная и поэтому занимает всего несколько страниц. В ней дается общее определение численной модели орбитального движения, а также обязательные атрибуты и характеристики эффективности численного моделирования.

Во второй главе описываются математические модели основных сил, обуславливающих орбитальное движение. Это — гравитационные силы протяженного тела и материальной точки; радиационные и реактивные силы, вызванные солнечным светом; силы атмосферного торможения, а также некоторые возмущающие силы общей теории относительности в постньютоновском приближении. На примере внутреннего спутника Юпитера, Амальтеи, проводится анализ структуры основных силовых факторов в спутниковом движении.

Третья глава посвящена теоретическому и практическому исследованию огромного арсенала различных дифференциальных уравнений, представляющих математические модели орбитального движения. В частности, рассматриваются линейные и регулярные уравнения; уравнения Энке и Лагранжа; стабилизированные уравнения Баумгарта и т.д. В главе также дается обоснование их применения к численному решению задач небесной механики. Проводится сравнительный анализ эффективности уравнений орбитального движения применительно к численному моделированию спутниковых, астероидных и планетных орбит, для того чтобы дать четкие рекомендации по их использованию.

В четвертой главе рассматриваются широко используемые на практике методы численного интегрирования дифференциальных уравнений орбитального движения. Излагаются основные принципы их построения. В частности, рассматриваются методы Рунге–Кутты; экстраполяционные и многошаговые методы, а также

так называемые геометрические методы. Описываются результаты сравнительного анализа эффективности методов применительно к невозмущенным, слабовозмущенным и сильновозмущенным задачам небесной механики.

Следующие две главы посвящены методам численного решения обратных задач орбитальной динамики, иначе говоря, методам определения орбит из астрометрических наблюдений и оценивания точности определяемых орбитальных параметров. В пятой главе описываются итерационные методы минимизации целевой функции, характеризующей близость наблюдаемых и моделируемых величин. В частности, рассматриваются известные методы численной оптимизации, а именно: метод (наискорейшего) градиентного спуска, методы Гаусса–Ньютона и Левенберга–Марквардта, а также овражный составной метод, предлагаемый автором для решения обратных задач динамики близких спутников планет. Проводится сравнительный анализ эффективности методов на примере определения орбиты близкого спутника Юпитера Адрастеи.

Шестая глава — это фактически обзор стохастических линейных и нелинейных методов типа Монте-Карло для численного оценивания точности определяемых из наблюдений орбитальных параметров. Все эти методы основаны на концепциях классического регрессионного анализа. В конце главы описываются результаты стохастического моделирования при использовании исследуемых методов для оценивания точности орбиты нового внешнего спутника Юпитера S/2003 J04, параметры движения которого плохо определены вследствие скудного состава имеющихся на данный момент наблюдений. Вводится показатель нелинейности обратной задачи, необходимый главным образом для принятия решения в выборе между линейными и нелинейными методами.

Последняя глава — факультативная. В ней рассматриваются основанные на линеаризации орбитальной модели преобразования и некоторые их приложения к численному исследованию орбит, а именно для временного переноса доверительных областей и

вычисления ляпуновских показателей, а также для оперативного оценивания вероятности столкновения астероида с планетой.

В конце монографии приводится список цитируемых научных работ и предметный указатель используемых терминов, а в приложении даются основные формулы (интегралы и соотношения) задачи двух тел, которые могут быть полезными при численном моделировании орбит.

Автор признателен своим коллегам, проф. Т.В. Бордовицкой и проф. А.М. Черницову, за конструктивные обсуждения результатов, а также за ценные замечания и предложения. Особые благодарности проф. Э. Хайреру за плодотворную дискуссию и советы по некоторым ключевым вопросам четвертой главы.

СПИСОК ОБОЗНАЧЕНИЙ

a	скалярная величина
\mathbf{a}	вектор (столбец)
$\ \mathbf{a}\ $	норма вектора \mathbf{a}
$ \mathbf{a} $ или a	модуль трехмерного вектора \mathbf{a}
\mathbf{A}	матрица или вектор
\mathbf{A}^T	транспонированная матрица или вектор \mathbf{A}
\mathbf{A}^{-1}	обратная к \mathbf{A} матрица
$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$	скалярное произведение векторов \mathbf{a} и \mathbf{b}
$\mathbf{a}\mathbf{b}^T$	тензорное произведение векторов \mathbf{a} и \mathbf{b}
$\mathbf{a} \times \mathbf{b}$	векторное произведение векторов \mathbf{a} и \mathbf{b}
$\Delta \mathbf{a}$	ошибка, отклонение, возмущение вектора \mathbf{a} или поправка к нему
$\delta \mathbf{a}$	отклонение, возмущение (разд. 3.5) или случайная ошибка вектора \mathbf{a} (гл. 6)
$d\mathbf{a}/dt$ или $\dot{\mathbf{a}}$	первая производная от вектора \mathbf{a} по времени
$d^2\mathbf{a}/dt^2$ или $\ddot{\mathbf{a}}$	вторая производная от вектора \mathbf{a} по времени
\mathbf{a}' и \mathbf{a}''	первая и вторая производные от вектора \mathbf{a} по какой-либо независимой переменной
$\partial \mathbf{a} / \partial b$ или \mathbf{a}'_b	первая частная производная от вектора \mathbf{a} по переменной b
$\partial \mathbf{a} / \partial \mathbf{b}$ или $\mathbf{a}'_{\mathbf{b}}$	вектор (строка; столбец в разд. 4.6) частных производных от величины a по компонентам вектора \mathbf{b}
$\partial \mathbf{a} / \partial \mathbf{b}$ или $\mathbf{a}'_{\mathbf{b}}$	матрица частных производных от вектора \mathbf{a} по компонентам вектора \mathbf{b}
$\partial^2 \mathbf{a} / \partial b \partial c$ или \mathbf{a}''_{bc}	вторая частная производная от вектора \mathbf{a} по переменным b и c
$\partial^2 a / \partial \mathbf{b}^2$	квадратная симметричная матрица частных производных от величины a по компонентам вектора \mathbf{b}

x	вектор положения в физическом или фазовом (гл. 4) пространстве
t	физическое время
μ	гравитационный параметр
p	вектор положения в наблюдаемом пространстве
q	вектор параметров модели
F	центральная сила (ускорение)
P	возмущающая сила (ускорение)
\mathcal{P}	правая часть дифференциального уравнения орбитального движения

Термин бывает удачным или неудачным,
но не правильным или неправильным.

К.В. Холшевников

Глава 1. ЧИСЛЕННАЯ МОДЕЛЬ ОРБИТЫ

1.1. Терминология

В привычном понимании орбита — это траектория небесного тела в физическом пространстве². Небесная колея³ дается нам в наблюдениях как временной ряд положений космического объекта на небесной сфере⁴. Встает задача исследовать орбиту и движение по ней, используя наблюдательный материал.

Одним из методов исследования орбит является математическое моделирование. Математическая модель орбитального движения представляет собой совокупность уравнений, связывающих между собой параметры, которые непосредственно характеризуют динамическое состояние объекта на орбите. Таким образом, исследовательский процесс сводится к решению конкретной математической задачи, а именно к решению соответствующих уравнений. В современной постановке почти все такие задачи не имеют

²Пожалуй, официально ни один из небесных механиков не подпишется под таким определением, но в приватной беседе каждый как раз именно это будет иметь в виду, говоря об орбите конкретного космического объекта. Дело в том, что в небесной механике орбита — это математическая абстракция. Ряд определений орбиты можно найти в работе (Холшевников, Титов, 2007). Мы же выходим за теоретические рамки, с тем чтобы обозначить ассоциируемый с реальностью предмет, который исследуется посредством моделирования, вполне понимая, что сама реальность предмета исследования возможна только лишь в воображении.

³Орбита от лат. *orbita* — путь, колея.

⁴Впрочем, в качестве наблюдений могут выступать не только положения объекта, но и другие наблюдаемые величины, имеющие непосредственное отношение к орбитальному движению.

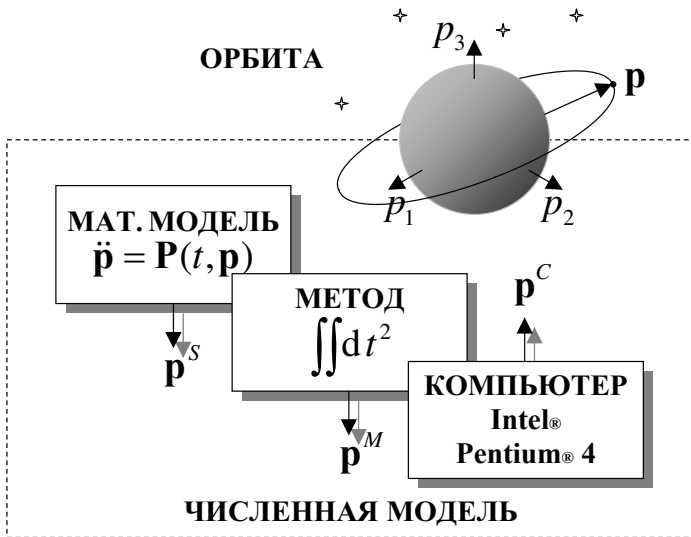


Рис. 1.1. Численное моделирование орбиты

точных решений, поэтому при моделировании используются либо приближенные аналитические, либо численные методы.

В последнее время главным образом за счет интенсивного развития компьютерных технологий для исследования орбит все чаще прибегают к численным методам. Процесс реализации математической модели численным методом на компьютере называется численным моделированием. Следовательно, численная модель представляет собой связку, по меньшей мере, трех компонент: математической модели, численного метода и компьютера (рис. 1.1). Современная математическая модель орбиты основывается на достаточно сложных дифференциальных уравнениях небесной механики, которые не интегрируются аналитически. Для их решения прибегают к приближенным численным методам интегрирования, и весь этот процесс реализуется на подручном компьютере.

Формально численную модель орбиты можно представить как

$$\mathbf{p}^C = \mathbf{p}^C(t, \mathbf{q}) = \mathbf{T}(t, \mathbf{x}(t, \mathbf{q}^D), \mathbf{q}^T). \quad (1.1)$$

Здесь t — время; \mathbf{x} — положение небесного тела в физическом пространстве относительно некоторой выбранной системы прямоугольных координат; $\mathbf{q} = (\mathbf{q}^D, \mathbf{q}^T)$ — вектор всех параметров модели; \mathbf{T} — преобразование перехода от физического пространства \mathbf{x} в пространство \mathbf{p} , где осуществляется мониторинг движения объекта (например, в пространство угловых координат на небесной сфере, отнесенной к топоцентру); $\mathbf{q}^D = (\mathbf{x}_0, \dot{\mathbf{x}}_0, t_0, q_8, \dots)^T$ и \mathbf{q}^T — параметрические векторы, связанные соответственно с движением объекта и с координатным преобразованием; \mathbf{x}_0 и $\dot{\mathbf{x}}_0$ — векторы динамического состояния небесного тела в начальный момент времени t_0 (начальная эпоха). Запись (1.1) следует понимать как представление состава входных и выходных параметров численной модели: на входе — время t и параметры \mathbf{q} , на выходе — положение на орбите \mathbf{p}^C как преобразованное положение в физическом пространстве \mathbf{x} , вычисляемое посредством численного интегрирования дифференциальных уравнений движения.

Поскольку в (1.1) только координаты \mathbf{x} непосредственно имеют отношение к численному интегрированию, преобразование \mathbf{T} , вообще говоря, можно опустить и считать численной моделью

$$\mathbf{p}^C = \mathbf{p}^C(t, \mathbf{q}^D) = \mathbf{x}(t, \mathbf{q}^D), \quad (1.2)$$

что фактически представляет численное решение дифференциальных уравнений движения

$$\frac{d^2\mathbf{x}}{dt^2} = \mathcal{P}(t, \mathbf{x}, \dot{\mathbf{x}}, \mathbf{q}^D), \quad \mathbf{x}_0 = \mathbf{x}(t_0), \quad \dot{\mathbf{x}}_0 = \dot{\mathbf{x}}(t_0). \quad (1.3)$$

Здесь \mathcal{P} — равнодействующая всех учитываемых в модели сил⁵. Уравнения (1.3) образуют математическую модель орбиты (или орбитального движения). Впрочем, для получения координат \mathbf{x} (1.2) вместо (1.3) может использоваться и другая формализация.

⁵Здесь, конечно же, имеются в виду силы за единицу массы.

Задача Коши (1.3) при заданных параметрах \mathbf{q}^D составляет прямую задачу орбитальной динамики в математической постановке. Между тем общее решение дифференциальных уравнений (1.3) (при произвольных параметрах) описывает многопараметрическое семейство орбит. Наблюдения позволяют выделить из него ту единственную модельную орбиту, которая с точки зрения интерпретации наблюдательных данных наилучшим образом соответствует орбите небесного тела. Задача определения динамических параметров \mathbf{q}^D из наблюдений называется обратной задачей орбитальной динамики. Как правило, в качестве определяемых параметров выступают лишь компоненты шестимерного вектора начального динамического состояния, т.е. \mathbf{x}_0 и $\dot{\mathbf{x}}_0$, либо их аналоги. Решение обратной задачи является неотъемлемым этапом в численном моделировании орбит реальных небесных тел.

Классические дифференциальные уравнения орбиты вида (1.3) являются следствием двух законов механики: второго закона Ньютона о взаимосвязи ускорения тела и приложенной к нему силы, а также ньютоновского закона всемирного тяготения. Поэтому любая численная модель орбиты, основанная на пошаговом интегрировании дифференциальных уравнений (1.3), всегда является динамической, в отличие от кинематических моделей (например, прецессирующих эллипсов), в которых действие силовых факторов на исследуемое тело не рассматривается.

Наконец, стоит также определиться в вопросе о принципиальном различии аналитических и численных моделей⁶. При использовании приближенных аналитических методов для решения уравнений (1.3) построенная аналитическая модель (как конечный формульный результат решения) будет иметь вид (1.2), и, таким образом, она становится и математической моделью. Можно ли считать ее компьютерную реализацию численной моделью? В нашем понимании она не будет таковой, поскольку необходимый атрибут численного моделирования — это пошаговость, т.е. последовательное определение динамических состояний объекта

⁶См. также гл. 4 на с. 124.

на заданные моменты времени непосредственно по предыдущим динамическим состояниям. В аналитических моделях этот атрибут, вообще говоря, отсутствует. С другой стороны, весьма интересно, что численную модель в пределах одного шага формально можно рассматривать как аналитическую. Действительно, ведь в соответствии с используемой схемой численного интегрирования приближенное решение также можно представить в формульном виде (1.2). Однако, как только выполняется второй шаг, аналитическая модель становится численной.

Численное моделирование орбиты — это последовательное вычисление положений (а также скоростей) объекта через малые промежутки времени. Орбита буквально восстанавливается (собирается) по кусочкам, и в этом смысле используемый нами в дальнейшем термин «интегрирование орбиты» вполне оправдан⁷. Интерполяция вычисленных положений позволяет приближенно представить орбиту как параметризованную временем кривую. Эта же интерполяция фактически будет определять и закон движения. Поэтому моделирование орбиты и моделирование орбитального движения суть один и тот же численный процесс, предполагающий прежде всего вычисление эфемериды небесного тела, по которой уже после можно исследовать как траекторию движения, так и движение по траектории.

1.2. Ошибки численного моделирования

Очевидно, моделирование на основе численного интегрирования сопряжено с ошибками. Допустим, \mathbf{p}^S (Simulation) — представление движения математической моделью; \mathbf{p}^M (Method) — численная реализация математической модели, иначе говоря, решение уравнений движения численным методом, а \mathbf{p}^C (Calculation) — фактически получаемый результат моделирования на компьютере (рис. 1.1). Следовательно, можно выделить три группы ошибок численного моделирования: $\Delta\mathbf{p}^S$, $\Delta\mathbf{p}^M$ и $\Delta\mathbf{p}^C$, которые обусловлены: 1) неточностью представления связей между парамет-

⁷Интегрирование от лат. *integratio* означает восстановление.

рами \mathbf{p} в математической модели и ошибками исходных параметров \mathbf{q} ($\Delta\mathbf{p}^S = \mathbf{p} - \mathbf{p}^S$); 2) неточностью численного метода ($\Delta\mathbf{p}^M = \mathbf{p}^S - \mathbf{p}^M$); 3) округлением чисел в компьютерной арифметике ($\Delta\mathbf{p}^C = \mathbf{p}^M - \mathbf{p}^C$). Ошибки $\Delta\mathbf{p}^S$, $\Delta\mathbf{p}^M$ и $\Delta\mathbf{p}^C$ называются модельными, методическими и вычислительными соответственно. Из них складывается полная ошибка численной модели:

$$\Delta\mathbf{p} = \Delta\mathbf{p}^S + \Delta\mathbf{p}^M + \Delta\mathbf{p}^C = \mathbf{p} - \mathbf{p}^C. \quad (1.4)$$

Поскольку методические и вычислительные ошибки не должны исказить динамическую картину, описываемую математической моделью, а методическая точность не должна быть выше вычислительной, правильная реализация численной модели предполагает выполнение условия

$$\|\Delta\mathbf{p}^S\| > \|\Delta\mathbf{p}^M\| > \|\Delta\mathbf{p}^C\|.$$

Практически оценить полную ошибку (1.4) можно лишь вследствие сопоставления результатов моделирования \mathbf{p}^C с наблюдениями небесного тела \mathbf{p}^O (Observation) как невязку численной модели $\mathbf{p}^O - \mathbf{p}^C$. Впрочем, из невязки в чистом виде полную ошибку моделирования выделить нельзя, поскольку сами наблюдения обременены ошибками различного происхождения $\Delta\mathbf{p}^O$. С другой стороны, предельная величина ошибок наблюдений обычно приблизительно известна, и если значение $\|\mathbf{p}^O - \mathbf{p}^C\|$ сопоставимо с этой величиной, то говорят, что точность численного моделирования хорошо согласуется с точностью наблюдений и вполне пригодна для их представления. Если же величина невязки значительно превосходит ожидаемую величину ошибок наблюдений, то это должно означать, что в невязке доминирует ошибка численного моделирования, и тогда первую можно рассматривать как оценку последней.

Между тем в современных численных моделях ошибки численного интегрирования $\Delta\mathbf{p}^M + \Delta\mathbf{p}^C$ существенно меньше ошибок математической модели $\Delta\mathbf{p}^S$, на несколько порядков и более. Поэтому в действительности большая невязка $\mathbf{p}^O - \mathbf{p}^C$ представляет

собой, прежде всего, оценку ошибки математической модели $\Delta \mathbf{p}^S$. В то же время выбор силовых факторов при составлении математической модели (дифференциальных уравнений) фактически не ограничен, и даже порой спектр моделируемых сил при численном исследовании орбит бывает избыточен. Поэтому основной причиной низкой точности математической модели (при грамотной формульной сборке, разумеется) являются достаточно грубые значения ее параметров \mathbf{q} , которые, в свою очередь, приводят к неудовлетворительным невязкам.

Используемые в модели параметры определяются из наблюдений и, таким образом, параметрические ошибки являются следствием прежде всего ошибок наблюдений. Методы статистического анализа позволяют получить приближенные оценки параметрических ошибок $\Delta \mathbf{q}$. Между тем численная модель дает возможность оценить ошибку $\Delta \mathbf{p}^S$, вызванную параметрическими ошибками, на любой момент времени без сопоставления результатов моделирования с наблюдениями.

Игнорирование того или иного силового фактора в математической модели также приводит к ошибке $\Delta \mathbf{p}^S$. Впрочем, этот факт сам по себе еще не является основанием для включения силового фактора в математическую модель: модельная ошибка должна быть достаточно большой в сравнении с ошибками наблюдений либо исследуемыми эффектами в орбитальной динамике. Чтобы оценить величину такой модельной ошибки, на практике сопоставляются результаты численного моделирования с полной структурой возмущающих сил и после вычленения соответствующего силового фактора. Разность между численными результатами дает оценку $\Delta \mathbf{p}^S$, которую также можно рассматривать как оценку возмущения, вызываемого силовым фактором.

Подобным образом оцениваются ошибки численного интегрирования $\Delta \mathbf{p}^M + \Delta \mathbf{p}^C$. Однако в данном случае результаты моделирования сопоставляются с результатами, имеющими более высокую методическую точность. Последние получаются либо увеличением порядка метода интегрирования, либо повышением задаваемой точности на шаге (иначе говоря, уменьшением шага ин-

тегрирования). Если же при сопоставлении результатов $\|\Delta\mathbf{p}^M\| > \|\Delta\mathbf{p}^C\|$, получим оценку методической ошибки $\Delta\mathbf{p}^M$, в противном случае — приближенную оценку вычислительной $\Delta\mathbf{p}^C$.

Иной способ оценивания ошибок численного интегрирования — это сопоставление результатов прямого и обратного моделирования. Сначала выполняется прямое моделирование от начального t_0 до заданного момента времени t_1 : $(\mathbf{x}_0, \dot{\mathbf{x}}_0) \rightarrow (\mathbf{x}_1, \dot{\mathbf{x}}_1)$, затем — обратное моделирование от t_1 до t_0 : $(\mathbf{x}_1, \dot{\mathbf{x}}_1) \rightarrow (\mathbf{x}_0, \dot{\mathbf{x}}_0)$, где в качестве начальных условий выбираются результаты прямого моделирования $(\mathbf{x}_1, \dot{\mathbf{x}}_1)$. Разность векторов положения \mathbf{x}_0 (или \mathbf{p}^C на начальный момент времени) до моделирования и после даст оценку ошибки $\Delta\mathbf{x}^M + \Delta\mathbf{x}^C$ (или $\Delta\mathbf{p}^M + \Delta\mathbf{p}^C$).

Впрочем, такой способ оценивания точности основан прежде всего на том предположении, что ошибки прямого и обратного моделирования не компенсируют друг друга. Между тем имеются численные методы интегрирования, для которых это предположение по отношению к методическим ошибкам не выполняется (см. симметричные методы в подразд. 4.6.4), и в таком случае удастся оценить лишь вычислительные ошибки.

1.3. Быстродействие численного моделирования

Характерной особенностью численного моделирования, как мы уже отметили, является прежде всего то, что векторы положения \mathbf{x} в физическом пространстве вычисляются путем пошагового интегрирования, на что могут потребоваться значительные затраты компьютерного времени, которые увеличиваются пропорционально временному интервалу прогнозирования орбитального движения от начального момента t_0 . Объективной характеристикой быстродействия численного моделирования может быть процессорное время, затраченное на решение задачи. Однако недостатком такой характеристики является то, что она зависит от архитектуры компьютера, а также от оптимизации программного кода. Поэтому даже если две численные модели, разработанные двумя авторами, дают одинаковые результаты, процессорное вре-

мя, требуемое на их вычисление, на разных компьютерах может существенно отличаться.

Между тем, для высокоточных численных моделей с громоздкими дифференциальными уравнениями (1.3) подавляющее количество компьютерного времени тратится на многократное вычисление правой части \mathcal{P} (см. гл. 4). Следовательно, в качестве альтернативной характеристики быстродействия, удобной в особенности для исследования эффективности методов интегрирования, может выступать число перевычислений правой части NF (Number of Functions) за весь процесс численного моделирования либо, если оценивается быстродействие при использовании одного метода, число выполненных шагов NS (Number of Steps).

1.4. Эффективность численного моделирования

Таким образом, как и любой спортсмен, в особенности игрового вида спорта, эффективная численная модель должна быть не только быстрой, но и высокоточной⁸.

Центральными результатами моделирования орбитального движения являются векторы динамического состояния исследуемого небесного тела: в прямых задачах — это прогнозируемые векторы на заданные моменты времени, получаемые на основе начальных, тогда как в обратных — это начальные векторы динамического состояния, получаемые на основе сопоставления результатов моделирования прямой задачи с наблюдениями. В этом смысле построение численной модели по наблюдениям можно рассматривать как процесс совместного решения прямой и обратной задач, что предусматривает использование самых разнообразных математических моделей, численных методов и компьютерного оборудования для повышения точности и быстродействия вычислений.

Говоря об эффективных математических моделях, методах численного моделирования и компьютерных средствах вычисления,

⁸ Впрочем, от численной модели как компьютерной программы можно потребовать наличие еще одной спортивной характеристики, а именно устойчивости.

в дальнейшем будем для определенности иметь в виду такие, которые повышают либо точность при сохранении быстродействия на несколько порядков и выше, либо быстродействие при сохранении точности в несколько раз и более, либо, в лучшем случае, обе характеристики эффективности одновременно.

Сила правит миром.
Б. Паскаль

Глава 2. МОДЕЛИ СИЛ

Состав сил \mathbf{P} в (1.2) предопределяет модельную точность. Чем богаче спектр учитываемых силовых факторов, действующих на небесное тело, тем точнее представление его орбиты математической моделью. Очевидно, повышение точности сопряжено с усложнением математической модели (1.2) и, как следствие, с низким быстродействием вычислительного процесса, основанного на пошаговом интегрировании. С другой стороны, адекватность модели должна соотноситься с точностью современных наблюдений либо с величинами исследуемых эффектов в орбитальном движении. Поэтому тенденция к повышению точности и усложнению модели порой может быть неоправданной. Это как раз тот случай, когда лучшее есть враг хорошего. К примеру, совершенно нецелесообразно учитывать гравитационное влияние Плутона на движение искусственного спутника Земли, модель которого и без того, как правило, чрезвычайно сложна; либо учитывать влияние галактик на астероидные орбиты.

Набор сил определяется в рамках физической постановки задачи с учетом характерных особенностей исследуемого орбитального движения, а также физических свойств рассматриваемого небесного тела. Если для моделирования движения астероидов достаточно ограничиться гравитационными силами тел Солнечной системы, то уже орбитальные модели комет обязательно должны дополняться негравитационными силами, вызванными испарением льда на их поверхностях под воздействием солнечного света. В то же время более сложные модели⁹ требуются для представ-

⁹Пожалуй, сейчас это самые сложные орбитальные модели, используемые на практике в динамической астрономии.

ления орбитального движения искусственных спутников Земли, в особенности низколетящих, где помимо гравитационного влияния планеты необходимо также учитывать многочисленные силовые факторы как гравитационной, так и негравитационной природы.

Силы обычно моделируются в физическом пространстве с использованием привычных, так скажем, осязаемых величин типа прямоугольных координат \mathbf{x} и скоростей $\dot{\mathbf{x}}$. При необходимости, разумеется, их можно преобразовать в другое пространство, более удобное для моделирования орбиты. Ниже приводятся наиболее используемые на практике математические модели силовых факторов, а именно: гравитационных сил протяженного тела \mathbf{P}_V (Voluminous) и материальной точки \mathbf{P}_P (Point); радиационных \mathbf{P}_L (Light) и реактивных \mathbf{P}_J (Jet) сил, вызванных солнечным светом; сил атмосферного торможения \mathbf{P}_D (Drag), а также некоторых возмущающих сил общей теории относительности \mathbf{P}_R (Relativistic) в постньютоновском приближении.

2.1. Гравитация протяженного тела

Гравитационный потенциал протяженного тела (вне его объема) можно представить в виде ряда Лапласа (Антонов и др., 1988)

$$V = \frac{\mu}{b} \sum_{n=0}^{\infty} \sum_{k=0}^n \left(\frac{b}{|\mathbf{x}|} \right)^{n+1} P_n^k(\sin \varphi) (C_n^k \cos k\lambda + S_n^k \sin k\lambda), \quad (2.1)$$

где μ — гравитационный параметр тела; b — масштабный множитель (если тело — планета, то это, как правило, ее экваториальный радиус); $P_n^k(z)$ — присоединенные многочлены Лежандра степени n и индекса k ; C_n^k и S_n^k — гармонические коэффициенты ($C_0^0 = 1$), характеризующие гравитационную структуру тела; φ и λ — соответственно широта и долгота точки, в которой вычисляется потенциал. Прямоугольные и сферические координаты связаны как

$$x_1 = |\mathbf{x}| \cos \varphi \cos \lambda, \quad x_2 = |\mathbf{x}| \cos \varphi \sin \lambda, \quad x_3 = |\mathbf{x}| \sin \varphi.$$

Заметим, что здесь вектор положения \mathbf{x} задается в системе координат, жестко связанной с гравитирующим телом, с началом в центре масс, поэтому для определения гравитационного воздействия тела в любой другой системе необходимо знать его относительную динамику.

Сферические гармоники (функции, зависящие от φ и λ) в (2.1) условно делят на три группы: зональные ($k = 0$), тессеральные ($n > k$) и секториальные ($n = k$). Если представить область значений $-90^\circ \geq \varphi \geq 90^\circ$ и $0^\circ \geq \lambda \geq 360^\circ$ на сфере, то участки знакопостоянных значений сферических гармоник будут иметь типичные для каждой группы конфигурации: широтные полосы для зональной, сферические трапеции для тессеральной и меридианные лепестки для секториальной. Член нулевого порядка $\mu/|\mathbf{x}|$ в (2.1), не зависящий от φ и λ , соответствует гравитации однородного шара или материальной точки.

Введем шаровые функции

$$c_n^k = P_n^k(\sin \varphi) \frac{\cos k\lambda}{|\mathbf{x}|^{n+1}}, \quad s_n^k = P_n^k(\sin \varphi) \frac{\sin k\lambda}{|\mathbf{x}|^{n+1}} \quad (2.2)$$

и перепишем потенциал (2.1) как (Brumberg, 1995)

$$V = \mu \sum_{n=0}^{\infty} \sum_{k=0}^n b^n (C_n^k c_n^k + S_n^k s_n^k). \quad (2.3)$$

Для вычисления шаровых функций (2.2) используют рекуррентные соотношения. Сначала вычисляют функции при $n = k$:

$$\begin{pmatrix} c_{n+1}^{n+1} \\ s_{n+1}^{n+1} \end{pmatrix} = \frac{2n+1}{|\mathbf{x}|^2} \begin{bmatrix} x_1 \begin{pmatrix} c_n^n \\ s_n^n \end{pmatrix} + x_2 \begin{pmatrix} -s_n^n \\ c_n^n \end{pmatrix} \end{bmatrix},$$

где

$$c_0^0 = \frac{1}{|\mathbf{x}|}, \quad s_0^0 = 0;$$

затем оставшиеся функции при $n > k$:

$$(n-k+1) \begin{pmatrix} c_{n+1}^k \\ s_{n+1}^k \end{pmatrix} = x_3 \frac{2n+1}{|\mathbf{x}|^2} \begin{pmatrix} c_n^k \\ s_n^k \end{pmatrix} - \frac{n+k}{|\mathbf{x}|^2} \begin{pmatrix} c_{n-1}^k \\ s_{n-1}^k \end{pmatrix}.$$

Сила гравитации определяется как градиент потенциальной функции $\partial V/\partial \mathbf{x}$, для чего требуются производные от шаровых функций по координатам. Они вычисляются по уже известным значениям шаровых функций (Drożnyer, 1977; Brumberg, 1995):

$$\begin{aligned}\frac{\partial}{\partial x_1} \begin{pmatrix} c_n^k \\ s_n^k \end{pmatrix} &= -\frac{1}{2} \begin{pmatrix} c_{n+1}^{k+1} \\ s_{n+1}^{k+1} \end{pmatrix} + \frac{1}{2} (n-k+2)(n-k+1) \begin{pmatrix} c_{n+1}^{k-1} \\ s_{n+1}^{k-1} \end{pmatrix}, \\ \frac{\partial}{\partial x_2} \begin{pmatrix} c_n^k \\ s_n^k \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} -s_{n+1}^{k+1} \\ c_{n+1}^{k+1} \end{pmatrix} + \frac{1}{2} (n-k+2)(n-k+1) \begin{pmatrix} -s_{n+1}^{k-1} \\ c_{n+1}^{k-1} \end{pmatrix}, \\ \frac{\partial}{\partial x_3} \begin{pmatrix} c_n^k \\ s_n^k \end{pmatrix} &= -(n-k+1) \begin{pmatrix} c_{n+1}^k \\ s_{n+1}^k \end{pmatrix},\end{aligned}$$

где при $k = 0$

$$(n+2)(n+1) \begin{pmatrix} c_{n+1}^{-1} \\ s_{n+1}^{-1} \end{pmatrix} = \begin{pmatrix} -c_{n+1}^1 \\ s_{n+1}^1 \end{pmatrix}.$$

Тогда, согласно (2.3), сила гравитации будет

$$\mathbf{P}_V^T = \frac{\partial V}{\partial \mathbf{x}} = \mu \sum_{n=0}^{\infty} \sum_{k=0}^n b^n \left(C_n^k \frac{\partial c_n^k}{\partial \mathbf{x}} + S_n^k \frac{\partial s_n^k}{\partial \mathbf{x}} \right). \quad (2.4)$$

Нельзя забывать, что сила (2.4) определяется в координатной системе, жестко связанной с телом, поэтому требуется еще выполнить ее преобразование к исходной системе координат, относительно которой моделируется орбита.

Если известно только зональное поле ($k = 0$) с точностью до нескольких гармоник (как, например, у газового гиганта Юпитера), использовать рекуррентные формулы для вычисления гравитационной силы нецелесообразно и лучше прибегнуть к явным формулам

$$\mathbf{P}_V^T = \frac{\partial V}{\partial \mathbf{x}} = \mu \sum_n C_n^0 b^n \frac{\partial}{\partial \mathbf{x}} \frac{P_n(\sin \varphi)}{|\mathbf{x}|^{n+1}}, \quad (2.5)$$

$$\frac{\partial}{\partial \mathbf{x}} \frac{P_n(\sin \varphi)}{|\mathbf{x}|^{n+1}} = \frac{1}{|\mathbf{x}|^{n+1}} \left[P_n'(\sin \varphi) \frac{\partial}{\partial \mathbf{x}} \sin \varphi - P_n(\sin \varphi) (n+1) \frac{\mathbf{x}^T}{|\mathbf{x}|^2} \right],$$

$$\frac{\partial}{\partial \mathbf{x}} \sin \varphi = \frac{1}{|\mathbf{x}|} \left(\mathbf{a} - \sin \varphi \frac{\mathbf{x}}{|\mathbf{x}|} \right)^T,$$

где $\mathbf{a} = (0, 0, 1)^T$ — орт оси аппликат.

2.2. Гравитация материальной точки

Поскольку $V \sim \mu/|\mathbf{x}|$ при $|\mathbf{x}| \rightarrow \infty$, то достаточно отдаленное гравитирующее тело можно рассматривать как материальную точку, сила гравитации которой в связанной с ней системе координат задается как

$$\mathbf{P}_P = -\mu \frac{\mathbf{x}}{|\mathbf{x}|^3} \equiv \mathbf{F}. \quad (2.6)$$

Когда тело (с гравитационным параметром μ_P) не является центральным, а его положение в некоторой абсолютной координатной системе задается вектором \mathbf{x}_P , будем иметь гравитационную силу

$$\mathbf{P}_P = -\mu_P \frac{\mathbf{x} - \mathbf{x}_P}{|\mathbf{x} - \mathbf{x}_P|^3}. \quad (2.7)$$

В относительной же системе координат, начало которой связано с центром масс какого-либо другого материального тела, следует учитывать инерциальную силу, вызываемую влиянием гравитирующего тела на центральное. Тогда формула (2.7) переписется как

$$\mathbf{P}_P = -\mu_P \frac{\mathbf{x} - \mathbf{x}_P}{|\mathbf{x} - \mathbf{x}_P|^3} - \mu_P \frac{\mathbf{x}_P}{|\mathbf{x}_P|^3}. \quad (2.8)$$

Центральная сила (2.6) при отсутствии других силовых факторов определяет кеплеровское движение, описываемое задачей двух тел. Если исследуемое тело массивное с гравитационным параметром, скажем, μ_P , оно будет вызывать соответствующую инерциальную (заметим, центральную) силу, и тогда результирующая сила останется кеплеровской, но с видоизмененным гравитационным параметром:

$$\mathbf{F} = -\mu \frac{\mathbf{x}}{|\mathbf{x}|^3} - \mu_P \frac{\mathbf{x}}{|\mathbf{x}|^3} = -(\mu + \mu_P) \frac{\mathbf{x}}{|\mathbf{x}|^3}.$$

Однако в большинстве случаев мы все же будем пользоваться формулой (2.6), полагая, что гравитационный параметр исследуемого тела либо пренебрежимо мал, либо он неявно входит в μ .

2.3. Радиационные силы

Значительное влияние на небесные тела, в особенности на мелкие частицы, могут оказывать радиационные силы: световое давление и эффект Пойнтинга–Робертсона. Предположение о существовании светового давления высказывалось еще И. Кеплером в XVII в. для объяснения образования хвостов комет при их прохождении вблизи Солнца, но только в 1873 г. Дж. Максвелл (Maxwell, 1873) объяснил это явление в рамках классической электродинамики. В 1903 г. Д.Г. Пойнтинг выявил эффект торможения в движении малых метеоров и космической пыли при их сближении с Солнцем (Poynting, 1903), и уже в 1937 г. Г.П. Робертсон интерпретировал его с точки зрения релятивистской динамики (Robertson, 1937). Впоследствии торможение частиц под действием света стали называть эффектом Пойнтинга–Робертсона.

В системе координат с началом в центре масс Солнца радиационные силы можно представить в виде

$$\mathbf{P}_L = L \frac{\mathbf{x}}{|\mathbf{x}|} - L \left(\frac{\dot{\mathbf{x}} \cdot \mathbf{x}}{c|\mathbf{x}|} \frac{\mathbf{x}}{|\mathbf{x}|} + \frac{\dot{\mathbf{x}}}{c} \right), \quad L = \kappa \theta \frac{a_E^2}{|\mathbf{x}|^2} \frac{\sigma}{m}, \quad (2.9)$$

где c — скорость света; $\kappa = 4.56 \cdot 10^{-6}$ Н/м² — солнечная постоянная; θ — постоянная, характеризующая отражающие свойства объекта ($\theta = 1$ соответствует зеркальному отражению); a_E — астрономическая единица (среднее расстояние от Земли до Солнца); σ и m — площадь миделева сечения, отнесенного к плоскости, перпендикулярной вектору положения, и масса исследуемого объекта соответственно. Первая составляющая в (2.9) отвечает за световое давление, вторая — за эффект Пойнтинга–Робертсона.

В зависимости от отношения миделева сечения σ к массе m небесного тела преобладающими будут либо силы гравитационного притяжения, либо светового давления, либо радиационного

торможения, обусловленного эффектом Пойнтинга–Робертсона. Так, на массивные тела Солнечной системы основное влияние оказывает сила притяжения Солнца. Если же масса небесного тела очень мала, действие светового давления значительно, но при этом его влияние может оказаться недостаточным, чтобы вытолкнуть частицу за границы Солнечной системы, между тем эффект Пойнтинга–Робертсона, действуя на длительном интервале времени, будет приводить к медленному падению частицы на Солнце по спирали. Наконец, для еще более мелких частиц (например, космической пыли) световое отталкивание уже существенно преобладает как над тяготением к Солнцу, так и над эффектом Пойнтинга–Робертсона, и солнечные лучи будут гнать частицы по гиперболическим орбитам за пределы Солнечной системы.

При моделировании радиационных сил, очевидно, нужно также учитывать затменные явления, когда исследуемый объект попадает в тень, отбрасываемую другим небесным телом. Для учета затменных явлений вводится так называемая функция тени Φ , формально умножаемая на \mathbf{P}_L . Она фактически определяет площадь видимой с объекта части диска светила, перекрываемого объемным телом. Предполагая, что светило и тело имеют шарообразную форму, из геометрических соображений нетрудно вывести формулы для функции тени.

Пусть r — угловое расстояние относительно исследуемого объекта между центрами светила и заслоняющего его тела с угловыми радиусами их дисков r_S и r_P соответственно. Радиусы дисков и угловое расстояние определяются как

$$r_S = \arcsin \frac{b_S}{|\mathbf{x}|}, \quad r_P = \arcsin \frac{b_P}{|\mathbf{x} - \mathbf{x}_P|}, \quad r = \arccos \frac{\mathbf{x} \cdot (\mathbf{x} - \mathbf{x}_P)}{|\mathbf{x}| |\mathbf{x} - \mathbf{x}_P|},$$

где b_S и b_P — соответствующие радиусы объектов в пространстве.

Функция тени вводится при необходимом условии затмения:

$$r < r_S + r_P \quad \text{и} \quad |\mathbf{x}| > |\mathbf{x} - \mathbf{x}_P|.$$

Ее можно представить в виде

$$\Phi = 1 - \frac{s}{s_S},$$

где s — площадь ущерба:

$$s = \begin{cases} s_{SP}, & \text{если } r > |r_S - r_P|, \\ s_P, & \text{если } r \leq |r_S - r_P| \text{ и } r_P < r_S, \\ s_S, & \text{если } r \leq |r_S - r_P| \text{ и } r_P \geq r_S; \end{cases}$$

s_S , s_P и s_{SP} — площади дисков светила и тела, а также их пересечения соответственно:

$$s_S = \pi r_S^2, \quad s_P = \pi r_P^2, \quad s_{SP} = \frac{r_S^2}{2}(\varphi_S - \sin \varphi_S) + \frac{r_P^2}{2}(\varphi_P - \sin \varphi_P).$$

Величины φ_S и φ_P — это углы между радиальными направлениями из центров дисков в точки пересечения их границ. Они определяются по формулам

$$\varphi_S = 2 \arccos \left(\frac{r^2 + r_S^2 - r_P^2}{2rr_S} \right), \quad \varphi_P = 2 \arccos \left(\frac{r^2 + r_P^2 - r_S^2}{2rr_P} \right).$$

2.4. Реактивные силы

Сублимация (испарение) льда под воздействием солнечного света на поверхности кометы вызывает реактивную силу, заметно влияющую на орбитальное движение объекта при его подлете к Солнцу. Для моделирования негравитационных эффектов, обусловленных сублимацией, как правило, прибегают к эмпирической модели Марсдена–Секанины (Marsden et al., 1973). Согласно этой модели, величина реактивной силы определяется как

$$|\mathbf{P}_J| = A \left(\frac{|\mathbf{x}|}{|\mathbf{x}|_0} \right)^{-m} \left[1 + \left(\frac{|\mathbf{x}|}{|\mathbf{x}|_0} \right)^n \right]^{-k},$$

где \mathbf{x} — гелиоцентрический вектор положения кометы; $m = 2.15$; $n = 5.093$; $k = 4.6142$; $A = 0.111262$; $|\mathbf{x}|_0 = 2.808$ а.е.

Направление реактивной силы удобно задавать в орбитальной системе координат. Таким образом, ее можно представить в виде

$$\mathbf{P}_J = |\mathbf{P}_J|(\alpha_1 \mathbf{S} + \alpha_2 \mathbf{T} + \alpha_3 \mathbf{W}), \quad (2.10)$$

где \mathbf{S} , \mathbf{T} и \mathbf{W} — соответственно радиальная, трансверсальная и нормальная орты координатной системы, а константы α_1 , α_2 и α_3 определяются из наблюдений совместно с другими орбитальными параметрами кометы. Напомним,

$$\mathbf{S} = \frac{\mathbf{x}}{|\mathbf{x}|}, \quad \mathbf{T} = \mathbf{W} \times \mathbf{S}, \quad \mathbf{W} = \frac{\mathbf{x} \times \dot{\mathbf{x}}}{|\mathbf{x} \times \dot{\mathbf{x}}|}.$$

Модель (2.10) допускает изменение параметров от появления к появлению кометы, что предполагает введение для каждой составляющей экспоненциального множителя:

$$\mathbf{P}_J = |\mathbf{P}_J|(\alpha_1 e^{-\beta_1 \tau} \mathbf{S} + \alpha_2 e^{-\beta_2 \tau} \mathbf{T} + \alpha_3 e^{-\beta_3 \tau} \mathbf{W}),$$

где β_1 , β_2 и β_3 — также определяемые из наблюдений константы, а τ — время относительно начальной эпохи в сутках, деленное обычно на 10000.

Поскольку параметр α_3 определяется очень неуверенно, на практике обычно используют усеченную модель

$$\mathbf{P}_J = |\mathbf{P}_J|(\alpha_1 e^{-\beta_1 \tau} \mathbf{S} + \alpha_2 e^{-\beta_2 \tau} \mathbf{T}). \quad (2.11)$$

Заметим, что трансверсальная составляющая в (2.11) обусловлена главным образом комбинированным эффектом сублимации и вращения кометы. Таким образом, по значениям определяемых констант α_2 и β_2 можно судить о скорости и ускорении вращательного движения объекта. Очевидно, значение $\alpha_2 = 0$ свидетельствует об отсутствии вращения кометы.

2.5. Атмосферное торможение

Если моделирование орбитального движения выполняется вблизи поверхности массивного тела, высланной газовой оболочкой (атмосферой), то необходимо учитывать силу сопротивления газа, который приводит к торможению исследуемого объекта.

Сила сопротивления газа вычисляется по формуле

$$\mathbf{P}_D = -\frac{1}{2} c_D \frac{\sigma}{m} \rho |\dot{\mathbf{x}} - \dot{\mathbf{x}}_A| (\dot{\mathbf{x}} - \dot{\mathbf{x}}_A),$$

где c_D — безразмерный коэффициент сопротивления (для сильно разреженного газа, например, $c_D = 2$); σ и m — площадь миделева сечения, отнесенного к плоскости, перпендикулярной относительному вектору скорости в газовой среде, и масса объекта соответственно; ρ — плотность газа; $\dot{\mathbf{x}}_A$ — локальная скорость газового потока.

Главной трудностью при моделировании атмосферного торможения является вычисление плотности ρ и скорости $\dot{\mathbf{x}}_A$, которые зависят от множества факторов. Впрочем, при моделировании околопланетной орбитальной динамики на больших высотах можно полагать, что атмосфера изотермическая и стационарная. Тогда $\dot{\mathbf{x}}_A = \dot{\mathbf{x}}_P + \boldsymbol{\omega} \times (\mathbf{x} - \mathbf{x}_P)$, а плотность будет распределена экспоненциально по высоте $h = |\mathbf{x} - \mathbf{x}_P| - b_P$:

$$\rho = \rho_0 e^{-(h-h_0)/H},$$

где $\boldsymbol{\omega}$ — угловая скорость вращения планеты; ρ_0 — плотность атмосферы на высоте h_0 ; H — так называемая шкала высот.

2.6. Релятивистские возмущения

Природу гравитации объясняет общая теория относительности, поэтому небесная механика по сути должна быть релятивистской, поскольку исследуемое орбитальное движение есть результат главным образом гравитационных взаимодействий. В большинстве же случаев эти гравитационные взаимодействия достаточно слабые, и с точки зрения адекватного представления современных наблюдений точность весьма сложных уравнений общей теории относительности оказывается избыточной. Тем не менее релятивистские эффекты наблюдаются, и поэтому они должны каким-то образом учитываться при численном моделировании орбит.

Достаточно простые уравнения с релятивистскими поправками удается получить в результате постньютоновского приближения общей теории относительности (т.е. с точностью до членов порядка $|\dot{\mathbf{x}}|^2/c^2$) в гармонических координатах и в динамически

невращающейся координатной системе (Брумберг, 1972). Представим наиболее значимые релятивистские поправки, а именно шварцшильдовские, которые обусловлены сферически-симметричным гравитационным полем одного массивного тела. В координатной системе, связанной с этим телом, они будут иметь вид

$$\mathbf{P}_R = 4 \frac{\mu^2}{c^2 |\mathbf{x}|^4} \mathbf{x} + \frac{\mu}{c^2 |\mathbf{x}|^3} [4(\mathbf{x} \cdot \dot{\mathbf{x}}) \dot{\mathbf{x}} - (\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}) \mathbf{x}].$$

2.7. Сравнительный анализ силовых факторов

Как уже отмечалось выше, набор моделируемых сил определяется в рамках физической постановки задачи с учетом характерных особенностей исследуемого орбитального движения, а также физических свойств рассматриваемого небесного тела. В качестве примера проведем анализ степени влияния некоторых основных силовых факторов на движение близкого спутника Юпитера Амальтеи.

Спутник Амальтея — малый объект Солнечной системы с размерами порядка $250 \times 150 \times 150$ км. Он движется внутри орбит галилеевых спутников, самых крупных в Солнечной системе, по почти круговой йовиэкваториальной орбите на расстоянии от Юпитера 2.5 его радиуса. Ввиду чрезвычайной близости спутника к Юпитеру, его движение главным образом подчинено мощному гравитационному влиянию массивной планеты, вследствие чего частота обращения спутника очень высока и соответствующий ей период составляет почти 0.5 сут.

Физическая постановка задачи в данном случае формулируется следующим образом. Будем рассматривать движение спутника Амальтеи, объекта с пренебрежимо малой массой, в поле тяготения Юпитера (с точностью до 6-й зональной гармоники) под действием гравитационных сил галилеевых спутников, Солнца и сторонних планет, а также релятивистских возмущающих сил в рамках задачи Шварцшильда. Ввиду малых размеров гравитирующих тел: галилеевых спутников, Солнца и планет (за исключением Юпитера), по отношению к их расстояниям до Амальтеи

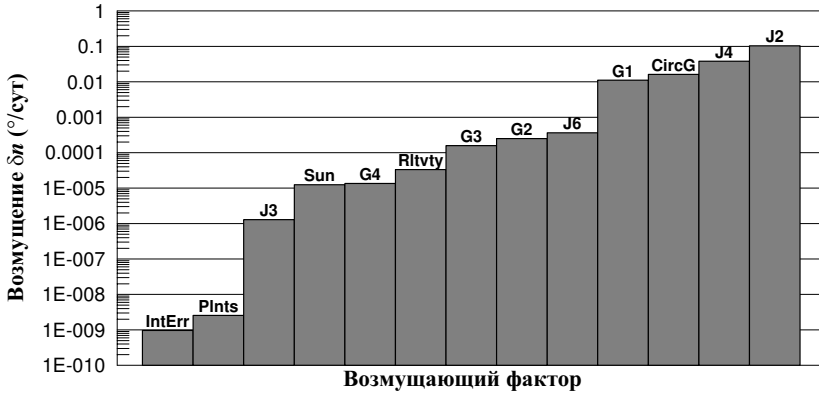


Рис. 2.1. Возмущения в среднем движении Амальтеи

они рассматриваются как материальные точки. Координаты галилеевых спутников вычисляются из теории Лейни L1 (Laineu et al., 2004), а координаты Солнца и планет — из эфемериды DE405 (Standish, 1998). Константы гравитационного поля Юпитера, а также массы галилеевых спутников берутся из теории Лиске E5 (Lieske, 1998), тогда как массы сторонних планет (относительно массы Солнца) — из DE405.

Степень влияния силовых факторов на орбитальное движение Амальтеи представлена на рис. 2.1. Результаты на рисунке получены путем численной оценки скорости вековых отклонений (возмущений) в средней долготе вследствие игнорирования тех или иных сил. Таким образом, эти результаты можно также рассматривать как модельные ошибки, вызванные неучетом силовых факторов. Разумеется, чтобы получить оценки отклонений в координатах, нужно умножить соответствующие отклонения в долготе на радиус-вектор спутника. Заметим, что результаты для нулевой гармоники йовипотенциала не приводятся ввиду очевидной необходимости ее учета при моделировании.

Из рис. 2.1 видно, что наиболее весомыми являются влияния от первых двух четных гармоник гравитационного поля Юпитера.

ра (**J2**, **J4**), а также от притяжения наиболее близкого галилеева спутника Ио (**G1**). В частности, уже через 10 лет возмущение в быстрой переменной от сжатия Юпитера становится порядка одного оборота. Возмущения от других галилеевых спутников (**G2**, **G3**, **G4**), очевидно, уменьшаются в соответствии с удаленностью от планеты. Степень влияния притяжений Солнца (**Sun**) и наиболее далекого галилеева спутника Каллисто (**G4**), а также релятивистских эффектов (**Rltvty**) приблизительно одного порядка. Наименьший вклад в движение Амальтеи вносят возмущения от сторонних планет (**Plnts**).

Кроме того, на рисунке приведены оценки возмущений, обусловленных ошибками интегрирования (**IntErr**) и упрощением орбитальной модели галилеевых спутников до круговой (**CircG**). Видно, что влияние ошибок интегрирования, как и полагается, существенно меньше учитываемых возмущений. В то же время оценка **CircG** показывает, что ошибка вследствие упрощения орбитальной модели спутников-гигантов настолько велика, что ее использование для итоговой обработки наблюдений близких спутников совершенно недопустимо.

Законы природы выражаются
дифференциальными уравнениями.

И. Ньютон

Глава 3. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ОРБИТ

При численном моделировании исследуемые орбиты небесных тел формализуются на основе систем дифференциальных уравнений, которые интегрируются численно. В связи с этим для высокоэффективного моделирования очень важно, насколько удачно выбраны уравнения, поскольку от этого непосредственно зависят точность и быстродействие их интегрирования.

Классические уравнения орбитального движения в прямоугольных координатах относительно массивного центрального тела можно представить в виде

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{\mu}{|\mathbf{x}|^3}\mathbf{x} + \mathbf{P} \equiv \mathbf{F} + \mathbf{P}. \quad (3.1)$$

Здесь \mathbf{x} — вектор положения; t — время; μ — гравитационный параметр центрального тела; \mathbf{F} и \mathbf{P} — центральная и возмущающая силы, причем обычно $|\mathbf{P}| \ll |\mathbf{F}|$.

Численное интегрирование уравнений (3.1) связано со следующими трудностями. Правые части уравнений представляют собой быстроизменяющиеся функции, которые необходимо интегрировать с малым шагом. Это приводит к увеличению объема вычислений, что, в свою очередь, сопряжено с быстрым накоплением ошибок округления. Данная трудность усугубляется наличием в уравнениях движения особенности при $\mathbf{x} = \mathbf{0}$, что в случае высокоэксцентричных орбит становится причиной сильно неравномерного поведения величины правой части (3.1). Кроме того, уравнения неустойчивы по Ляпунову. Хорошо известно, что ляпуновская неустойчивость дифференциальных уравнений при числен-

ном интегрировании создает благоприятные условия для усиления всевозможных ошибок, неизбежно сопровождающих любой численный процесс.

В данной главе мы сфокусируем свое внимание на известных методах формализованного представления орбитального движения, которые устраняют названные выше трудности. Будем называть их методами теории специальных возмущений¹⁰ (Авдюшев, 2006b; Авдюшев, 2007). Получение с их помощью качественно новых уравнений, как правило, основывается на некоторой априорной информации об исследуемом орбитальном движении. Большинство из тех методов, которые здесь излагаются, основано на решениях задачи двух тел в предположении, что исследуемая орбита близка к кеплеровской. В частности, рассмотрим методы лиnearизации и регуляризации, сглаживающие преобразования, методы вариации координат (Энке) и постоянных (Лагранжа), численную стабилизацию Баумгарте и Накози, а также дадим обоснование их применения к решению задач небесной механики. Далее проведем сравнительный анализ эффективности получаемых уравнений применительно к численному моделированию спутниковых, астероидных и планетных орбит, для того чтобы дать четкие рекомендации по их использованию.

3.1. Линейные и регулярные уравнения

Проблема регуляризации двойных соударений в задаче трех тел в пространственном случае впервые была рассмотрена К. Сундманом (Sundman, 1912). В частности, было выведено регуляризующее преобразование независимой переменной, которое нашло применение во многих методах регуляризации двойных соударений, предложенных после Сундмана. Т. Леви-Чивита (Levi-Civita, 1903) удалось получить преобразование координат для плоского случая, позволяющее вместе с временным преобразованием Сунд-

¹⁰В зарубежной литературе под специальными возмущениями понимают численные методы (Херрик, 1977; Рой, 1981).

мана привести уравнение движения невозмущенной задачи двух тел к линейному виду.

В дальнейшем появилось много работ, посвященных построению различных регуляризирующих преобразований (Stiefel et al., 1967; Себехей, 1982; Baumgarte, Stiefel, 1974). Детальный обзор работ на эту тему представлен В. Себехеем (Себехей, 1982), который исследовал регуляризацию в ограниченной задаче трех тел. Разработан ряд методов (Baumgarte, Stiefel, 1974; Шефер, 1986; Brumberg, Brumberg, 1999), понижающих порядок сингулярности дифференциальных уравнений и регуляризирующих их решение путем использования различных преобразований независимой переменной.

П. Кустаанхеймо и Е. Штифелю (Kustaanheimo, Stiefel, 1965) удалось продолжить координатное преобразование Леви-Чивиты на четырехмерное пространство. Это позволило адаптировать линейную теорию задачи двух тел для пространственного случая, что является, пожалуй, едва ли не самым известным результатом в теории регуляризации.

Д. Хегги (Heggie, 1974) предложены более общие регуляризирующие алгоритмы, устраняющие особенности дифференциальных уравнений движения в окрестностях двойных соударений в общей задаче N тел. Алгоритмы Хегги нашли широкое применение в задачах небесной механики (Aarseth, Zare, 1974; Шефер, 1980; Mikkola, 1985). Одновременно с Хегги применительно к задачам звездной динамики С. Аарсетом и К. Заре (Aarseth, Zare, 1974) были разработаны эффективные преобразования для регуляризации тройных соударений. Позже С. Микколой и С. Аарсетом (Mikkola, Aarseth, 1990) предложена так называемая цепочная регуляризация для исследования кратных сближений многих тел.

Заметный вклад в развитие KS-алгоритмов внес коллектив томских небесных механиков. Их работы, посвященные теоретическим и экспериментальным исследованиям регуляризирующих преобразований Кустаанхеймо–Штифеля на предмет эффективности их использования в численных моделях движения различных небесных объектов, позволили значительно расширить об-

ласть их применения во многих прикладных задачах небесной механики (Шефер, 1980; Бордовицына, Шарковский, 1994; Бордовицына и др., 1998b; Авдюшев, 1999; Bordovitsyna et al., 2001; Авдюшев, 2004).

Цель методов линеаризации и регуляризации состоит в том, чтобы представить уравнения движения в линейном и в регулярном виде. Рассмотрим сначала уравнения невозмущенного орбитального движения

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}(\mathbf{x}). \quad (3.2)$$

Предположим, что уравнения (3.2) имеют интегралы

$$\begin{aligned} \mathbf{G}_i(t, \mathbf{x}, \dot{\mathbf{x}}) - \mathbf{g}_i &\equiv \mathbf{0} \quad (i = 1, \dots, n), \\ H_j(t, \mathbf{x}, \dot{\mathbf{x}}) - h_j &\equiv 0 \quad (j = 1, \dots, m). \end{aligned} \quad (3.3)$$

Здесь \mathbf{G}_i и H_j — векторные и скалярные интегральные функции соответственно, а \mathbf{g}_i и h_j — интегральные параметры, которые в невозмущенном движении постоянны. Далее введем координатное и временное преобразования

$$\mathbf{x} = \mathbf{A}\mathbf{u}, \quad dt = f(\mathbf{x})ds, \quad (3.4)$$

которые позволяют перейти к новым переменным \mathbf{u} и s .

Главная идея линеаризации и регуляризации состоит во введении в уравнения, записанные в новых переменных, интегральных соотношений (3.3) (Szebehely, 1976). В результате уравнения принимают вид (Шефер, 1991)

$$\begin{aligned} \mathbf{u}'' &= \mathbf{A}^{-1}(f^2\mathbf{F} + f^{-1}f'(\mathbf{A}\mathbf{u}' + \mathbf{A}'\mathbf{u}) - 2\mathbf{A}'\mathbf{u}' - \mathbf{A}'\mathbf{u}) + \\ &+ \mathbf{A}^{-1} \sum_{i=1}^n N_i(\mathbf{G}_i - \mathbf{g}_i) + \mathbf{A}^{-1} \sum_{j=1}^m \mathbf{M}_j(H_j - h_j), \end{aligned} \quad (3.5)$$

где штрих означает производную по s , а N_i и \mathbf{M}_j — неопределенные коэффициенты, которые задаются таким образом, чтобы

уравнения принимали линейный и регулярный вид (Шефер, 1991)

$$\mathbf{u}'' = \underbrace{k_1(\mathbf{g}_1, \dots, \mathbf{g}_n, h_1, \dots, h_m)}_{\text{const}} \mathbf{u} + \underbrace{k_2(\mathbf{g}_1, \dots, \mathbf{g}_n, h_1, \dots, h_m)}_{\text{const}}. \quad (3.6)$$

В возмущенном случае, применяя вышеизложенные преобразования, будем иметь слабонелинейные уравнения вида

$$\mathbf{u}'' = k_1(\mathbf{g}_1, \dots, \mathbf{g}_n, h_1, \dots, h_m) \mathbf{u} + \mathbf{k}_2(\mathbf{g}_1, \dots, \mathbf{g}_n, h_1, \dots, h_m) + \mathbf{P}(\mathbf{P}). \quad (3.7)$$

Поскольку здесь интегральные параметры уже не являются постоянными и, кроме того, вследствие появления возмущающей силы $\mathbf{P} = \mathbf{P}(t)$ правая часть становится функцией времени, систему (3.7) необходимо дополнить уравнениями

$$\begin{aligned} \mathbf{g}'_i &= f \frac{\partial \mathbf{G}_i}{\partial \dot{\mathbf{x}}} \mathbf{P} \quad (i = 1, \dots, n), \\ h'_j &= f \frac{\partial H_j}{\partial \dot{\mathbf{x}}} \mathbf{P} \quad (j = 1, \dots, m), \quad t' = f. \end{aligned}$$

Очевидное неудобство в использовании регулярных уравнений состоит в том, что физическое время является интегрируемой переменной. В этой связи возникает проблема выхода на заданный момент времени t , так как невозможно заранее, до численного интегрирования, точно установить соответствующий момент фиктивного времени s как независимой переменной. Проблема обычно разрешается путем реорганизации процесса численного интегрирования. Выход из интегратора на заданный момент физического времени осуществляется следующим образом.

Вместо фиктивного времени на каждом шаге интегрирования ведется контроль за физическим временем. Численное интегрирование выполняется в обычном режиме (с постоянной или переменной величиной шага Δs) до тех пор, пока заданный момент физического времени t не окажется внутри шага интегрирования, т.е. между моментами в начале и в конце шага. Далее величина шага Δs , согласно дифференциальному соотношению (3.4), выбирается по приближенной формуле $\Delta s = \Delta t / f$, где Δt — разность

между заданным моментом физического времени и моментом на конце шага. Поскольку выбор Δs приближенный, его требуется повторить еще несколько раз, пока величина $|\Delta t|$ не станет пренебрежимо малой.

В результате подбора преобразований (3.4), а также соответствующих коэффициентов N_i и \mathbf{M}_j можно получить многочисленное семейство систем дифференциальных уравнений вида (3.7). Среди таких систем широко используются на практике системы уравнений в переменных Шперлинга–Бодэ (SB) (Burdet, 1968; Sperling, 1969; Silver, 1975)¹¹ и Кустаанхеймо–Штифеля (KS) (Kustaansheimo, Stiefel, 1965; Stiefel, Scheifele, 1971).

3.1.1. Уравнения в переменных Шперлинга–Бодэ

В системе SB-уравнений $\mathbf{x} = \mathbf{u}$, $dt = |\mathbf{x}|ds$ (преобразование Сундмана), а в качестве интегралов выступают интегралы энергии h и Лапласа \mathbf{g} (см. Приложение)

$$h = \frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|}, \quad \mathbf{g} = \dot{\mathbf{x}} \times (\mathbf{x} \times \dot{\mathbf{x}}) - \mu \frac{\mathbf{x}}{|\mathbf{x}|},$$

тогда как сами уравнения имеют вид

$$\begin{aligned} \mathbf{x}'' - 2h\mathbf{x} + \mathbf{g} &= |\mathbf{x}|^2 \mathbf{P}, \\ \mathbf{g}' &= 2\mathbf{x}(\mathbf{x}' \cdot \mathbf{P}) - \mathbf{x}'(\mathbf{x} \cdot \mathbf{P}) - \mathbf{P}(\mathbf{x} \cdot \mathbf{x}'), \quad h' = (\mathbf{x}' \cdot \mathbf{P}), \\ \tau' &= -\frac{1}{2h} \left[\mu + |\mathbf{x}|(\mathbf{x} \cdot \mathbf{P}) - \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|} \frac{h'}{h} \right], \quad t = \tau + \frac{1}{2h} \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|}. \end{aligned} \quad (3.8)$$

Здесь для определения времени t вводится временной элемент τ , который в слабозмущенном движении ведет себя почти линейно, а потому интегрирование его уравнения выполняется точнее. Скорости $\dot{\mathbf{x}}$ и \mathbf{x}' относительно физического времени и фиктивного, очевидно, связаны как $\dot{\mathbf{x}} = \mathbf{x}'/|\mathbf{x}|$.

¹¹Хотя подобные уравнения можно найти в более ранней работе (Дубошин, 1964).

При моделировании эллиптического движения ($h < 0$) удобнее использовать SB-уравнения, записанные относительно так называемой обобщенной эксцентрической аномалии E , которая связана с фиктивным временем s посредством дифференциального соотношения $dE = 2\omega ds$, где $\omega = \sqrt{-h/2}$ (Stiefel, Scheifele, 1971). После перехода к новой независимой переменной дифференциальные уравнения преобразуются к виду

$$\mathbf{x}'' + \mathbf{x} + \frac{\mathbf{g}}{4\omega^2} = \frac{1}{4\omega^2} [|\mathbf{x}|^2 \mathbf{P} + (\mathbf{x}' \cdot \mathbf{P}) \mathbf{x}'],$$

$$\mathbf{g}' = 2\mathbf{x}(\mathbf{x}' \cdot \mathbf{P}) - \mathbf{x}'(\mathbf{x} \cdot \mathbf{P}) - \mathbf{P}(\mathbf{x} \cdot \mathbf{x}'), \quad \omega' = -\frac{1}{4\omega}(\mathbf{x}' \cdot \mathbf{P}), \quad (3.9)$$

$$\tau' = \frac{1}{8\omega^3} \left[\mu + |\mathbf{x}|(\mathbf{x} \cdot \mathbf{P}) + 2 \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|} (\mathbf{x}' \cdot \mathbf{P}) \right], \quad t = \tau - \frac{1}{2\omega} \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|}.$$

Здесь штрих означает производную уже по обобщенной эксцентрической аномалии E , и соответственно скорости $\dot{\mathbf{x}}$ и \mathbf{x}' будут связаны как $\dot{\mathbf{x}} = 2\omega \mathbf{x}' / |\mathbf{x}|$.

3.1.2. Уравнения в переменных Кустаанхеймо–Штифеля

Система KS-уравнений получается в результате введения преобразований $\mathbf{x} = \mathbf{L}(\mathbf{u})\mathbf{u}$, $dt = |\mathbf{x}|ds$, где $\mathbf{u} = (u_1, u_2, u_3, u_4)^T$ — 4-мерный позиционный вектор; $\mathbf{L}(\mathbf{u})$ — матрица Кустаанхеймо–Штифеля (обобщенная матрица Леви-Чивиты) (Stiefel, Scheifele, 1971; Полещиков, Холопов, 1999):

$$\mathbf{L}(\mathbf{u}) = \begin{pmatrix} u_1 & -u_2 & -u_3 & u_4 \\ u_2 & u_1 & -u_4 & -u_3 \\ u_3 & u_4 & u_1 & u_2 \\ u_4 & -u_3 & u_2 & -u_1 \end{pmatrix}.$$

Заметим, что здесь вектор положения \mathbf{x} — 4-мерный, получаемый из 3-мерного путем добавления четвертой нулевой компоненты: $\mathbf{x} = (x_1, x_2, x_3, 0)^T$. Следовательно, каждой орбите фи-

зического пространства \mathbf{x} будет соответствовать однопараметрическое семейство орбит пространства \mathbf{u} . Чтобы в начале интегрирования выделить из них единственную, можно, согласно KS-преобразованию, использовать следующую сводку формул (Stiefel, Scheifele, 1971): если $x_1 \geq 0$,

$$u_1^2 = \frac{|\mathbf{x}| + x_1}{2}, \quad u_2 = \frac{x_2}{2u_1}, \quad u_3 = \frac{x_3}{2u_1}, \quad u_4 = 0;$$

в противном случае

$$u_1 = \frac{x_2}{2u_2}, \quad u_2^2 = \frac{|\mathbf{x}| - x_1}{2}, \quad u_3 = 0, \quad u_4 = \frac{x_3}{2u_2}.$$

Между тем скорости $\dot{\mathbf{x}}$ и \mathbf{u}' связаны соотношениями

$$\dot{\mathbf{x}} = \frac{2}{|\mathbf{u}|^2} \mathbf{L}(\mathbf{u})\mathbf{u}', \quad \mathbf{u}' = \frac{1}{2} \mathbf{L}^T(\mathbf{u})\dot{\mathbf{x}}.$$

Приведение уравнений к линейному и регулярному виду в KS-переменных с независимой переменной s оказывается возможным при использовании лишь интеграла энергии, и тогда система уравнений представляется как

$$\mathbf{u}'' - \frac{h}{2} \mathbf{u} = \frac{|\mathbf{u}|^2}{2} \mathbf{L}^T \mathbf{P}, \quad h' = 2(\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}), \quad (3.10)$$

$$\tau' = -\frac{1}{2h} \left[\mu + |\mathbf{u}|^2 (\mathbf{u} \cdot \mathbf{L}^T \mathbf{P}) - 2(\mathbf{u} \cdot \mathbf{u}') \frac{h'}{h} \right], \quad t = \tau + \frac{1}{h} (\mathbf{u} \cdot \mathbf{u}').$$

Как и для SB-уравнений в случае эллиптического движения, после введения обобщенной эксцентрической аномалии KS-уравнения могут быть приведены к виду

$$\mathbf{u}'' + \frac{1}{4} \mathbf{u} = \frac{1}{4\omega^2} \left[\frac{|\mathbf{u}|^2}{2} \mathbf{L}^T \mathbf{P} + 2(\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P})\mathbf{u}' \right], \quad \omega' = -\frac{1}{2\omega} (\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}), \quad (3.11)$$

$$\tau' = \frac{1}{8\omega^3} \left[\mu + |\mathbf{u}|^2 (\mathbf{u} \cdot \mathbf{L}^T \mathbf{P}) + 8(\mathbf{u} \cdot \mathbf{u}') (\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}) \right],$$

$$t = \tau - \frac{1}{\omega}(\mathbf{u} \cdot \mathbf{u}').$$

Здесь штрих означает производную по аномалии E , и соответственно скорости $\dot{\mathbf{x}}$ и \mathbf{u}' будут связаны соотношениями

$$\dot{\mathbf{x}} = \frac{4\omega}{|\mathbf{u}|^2} \mathbf{L}(\mathbf{u})\mathbf{u}', \quad \mathbf{u}' = \frac{1}{4\omega} \mathbf{L}^T(\mathbf{u})\dot{\mathbf{x}}.$$

Следует заметить, что уравнения (3.9) и (3.11) обладают стабилизирующими свойствами. Действительно, в невозмущенном случае эти уравнения движения сводятся к гармоническим осцилляторам, решения которых, как известно, устойчивы по Ляпунову. В возмущенном случае уравнения движения неустойчивы, однако эта неустойчивость будет гораздо менее опасной для численного интегрирования, нежели неустойчивость классических уравнений (3.1) (Avdyushev, 2003).

На рис. 3.1 наглядно демонстрируется колоссальная эффективность линейных и регулярных уравнений Шперлинга–Бодде и Кустаанхеймо–Штифеля, применяемых к моделированию эллиптических орбит. Как видно из рисунка, линеаризирующие и регуляризирующие преобразования позволяют существенно повысить точность интегрирования, в особенности для высокоэллиптических орбит, когда периодически происходят очень тесные сближения с центральным телом. Интересно, что даже для круговых орбит регулярные уравнения обеспечивают более высокую точность, нежели классические. В данном случае повышение точности как раз связано со стабилизирующим эффектом, о котором говорилось выше, и он становится тем значительнее, чем больше интервал интегрирования¹². Нельзя, конечно, не заметить, что точность интегрирования для KS-уравнений заметно выше, нежели для SB-уравнений, более чем на один порядок. По этой причине использование первых предпочтительнее. Кроме того, порядок KS-системы на единицу меньше.

Обратим также внимание на то, что результаты, представленные на рис. 3.1 для регулярных уравнений, получены при сов-

¹²См. также разд. 3.3

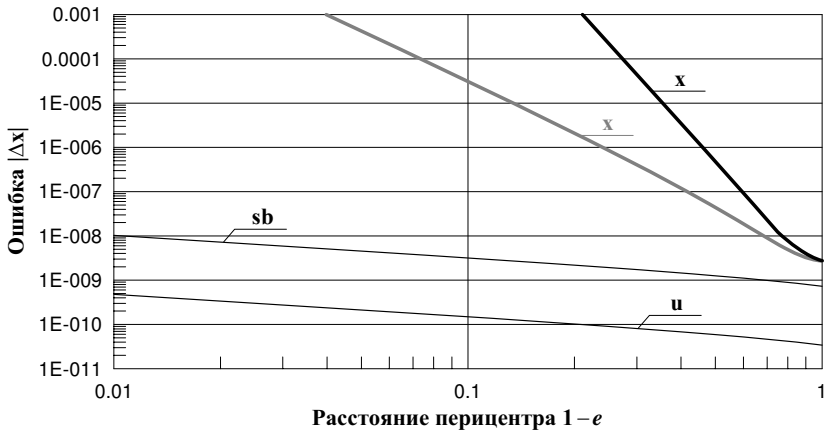


Рис. 3.1. Точность численного интегрирования уравнений в прямоугольных координатах (\mathbf{x}), а также уравнений Шперлинга–Бодэ (\mathbf{sb}) и Кустаанхеймо–Штифеля (\mathbf{u}) классическим методом Рунге–Кутты с постоянным шагом (1024 шагов за оборот) на временном интервале 10 оборотов при моделировании эллиптических орбит ($a = 1$, $n = 1$). Серым цветом (для уравнений в прямоугольных координатах) представлены результаты, полученные с переменным шагом

местном интегрировании дифференциального уравнения для временного элемента τ , который в эллиптическом случае меняется линейно. Разумеется, использование временного элемента вместо физического времени повышает точность интегрирования не только самого времени, но и координат, хотя это повышение точности незначительное, приблизительно в 2–3 раза.

Выводы об эффективности уравнений по результатам интегрирования с постоянным шагом могут восприниматься как сомнительные, поскольку в действительности численное моделирование орбит выполнялось в разных временных шкалах. Между тем постоянный шаг в физическом времени абсолютно не равносителен постоянному шагу в фиктивном времени. Возникающие сомнения развеивают результаты, полученные для уравнений в прямо-

угольных координатах при интегрировании с переменным шагом (рис. 3.1, серый цвет). При этом величина шага выбиралась таким образом, чтобы она соответствовала величине постоянного шага в фиктивном времени. Как видно, использование переменного шага значительно повышает точность моделирования, однако в то же время она остается еще достаточно низкой в сравнении с точностью для регулярных уравнений.

3.2. Сглаженные уравнения

Сглаживающие преобразования применяются к дифференциальным уравнениям, для того чтобы понизить скорость изменения величины их правых частей, что позволяет заметно повысить эффективность численного интегрирования, в особенности для сильно вытянутых эллиптических орбит, а также для орбит, имеющих тесные сближения с массивными телами.

Сглаживающие преобразования — это фактически временные преобразования, подобные используемым выше, которые задаются посредством дифференциального соотношения (Baumgarte, Stiefel, 1974)

$$dt = f(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) ds. \quad (3.12)$$

Их применение предполагает переход к новой независимой переменной s , фиктивному времени, после чего уравнения движения принимают вид

$$\mathbf{x}'' = f^2(\mathbf{F} + \mathbf{P}) + f^{-1} f' \mathbf{x}'.$$

Для приведения в соответствие координат \mathbf{x} с физическим временем t систему сглаженных уравнений необходимо пополнить уравнением $t' = f$. Если же взятие производной f' затруднительно, прибегают к аналогичной системе уравнений первого порядка

$$\mathbf{x}' = f \dot{\mathbf{x}}, \quad \dot{\mathbf{x}}' = f(\mathbf{F} + \mathbf{P}), \quad t' = f. \quad (3.13)$$

При сближении с массивным телом величины $|\dot{\mathbf{x}}|$ и $|\mathbf{F} + \mathbf{P}|$ стремительно возрастают. Следовательно, для сглаживания величина f , согласно (3.13), должна подбираться таким образом,

чтобы компенсировать (демпфировать) их быстрые изменения. Например, в случае слабозмущенного, но сильноэллиптического движения может быть выбрано сглаживающее преобразование с $f = |\mathbf{x}|^\beta$, где $\beta > 0$.

Поскольку интегрирование уравнений (3.13) выполняется по фиктивному времени s , возникает проблема выхода на заданный момент физического времени t . На практике она разрешается тем же способом, что и для регулярных уравнений (см. с. 42).

В небесно-механической практике широко используются такие сглаживающие преобразования, где в качестве f выбираются следующие величины: $|\mathbf{x}|$ (эксцентриская аномалия); $|\mathbf{x}|^{3/2}$ (эллиптическая аномалия) (Ferrándiz et al., 1987); $|\mathbf{x}|^2$ (истинная аномалия); $|\dot{\mathbf{x}}|^{-1}$ (дуга орбиты) (Brumberg, Brumberg, 1999).

Эффективность сглаживающих преобразований зависит от того, к какому типу орбит они применяются. На рис. 3.2–3.4 показаны точности $\lg |\Delta \mathbf{x}|$ численного интегрирования эллиптических орбит классическим методом Рунге–Кутты при одинаковом объеме вычислений в зависимости от применяемых временных преобразований (3.12), где $f = |\mathbf{x}|^\beta$, $f = |\dot{\mathbf{x}}|^{-\beta}$ и $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2}$. В частности, при $\beta = 1$ независимая переменная s становится обобщенной эксцентриской аномалией, физической и фазовой скоростью соответственно. При $\beta = 0$, очевидно, s — это физическое время во всех случаях.

Строгий анализ результатов показывает, что невозможно выбрать оптимальный параметр β , который бы обеспечивал наивысшую точность для любого типа орбит. Единственное, что определенно можно сказать, так это то, что для умеренных эксцентриситетов, приблизительно до $e = 0.4$, применение преобразований $f = |\mathbf{x}|^\beta$ и $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2}$ не имеет смысла, поскольку они вообще не дают повышения эффективности.

Следует также отметить одну особенность, которая может вызывать недоумение, а именно разная точность $\lg |\Delta \mathbf{x}|$ для $e = 0$ при разных значениях параметра β . Действительно для круговых орбит сглаживающие преобразования не работают и, казалось бы, их применение не должно влиять на эффективность численного

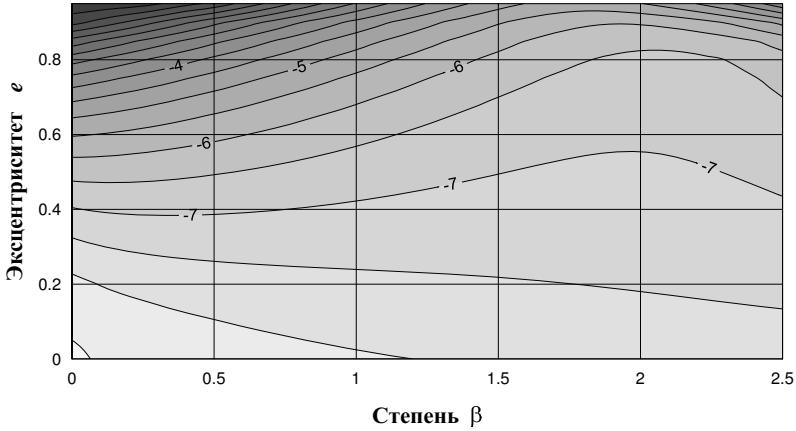


Рис. 3.2. Ошибка в векторе положения $\lg |\Delta \mathbf{x}|$ для $f = |\mathbf{x}|^\beta$ при численном интегрировании эллиптических орбит ($a = 1$, $n = 1$) классическим методом Рунге–Кутты с постоянным шагом Δs (1024 шагов за оборот) на временном интервале 10 оборотов

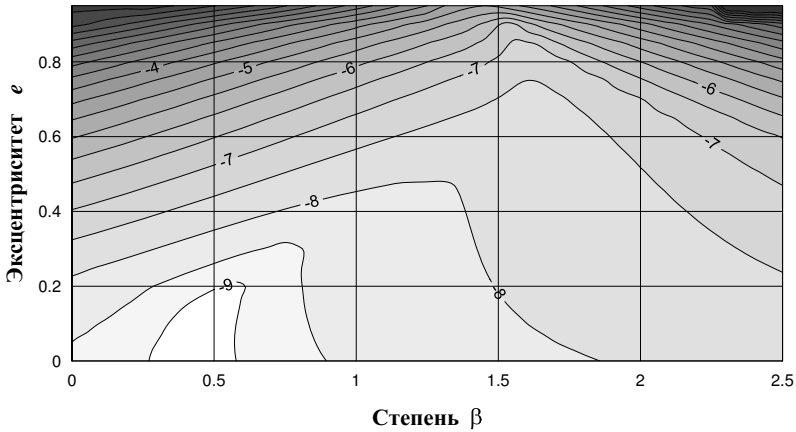


Рис. 3.3. То же, что и на рис. 3.2, но для $f = |\dot{\mathbf{x}}|^{-\beta}$

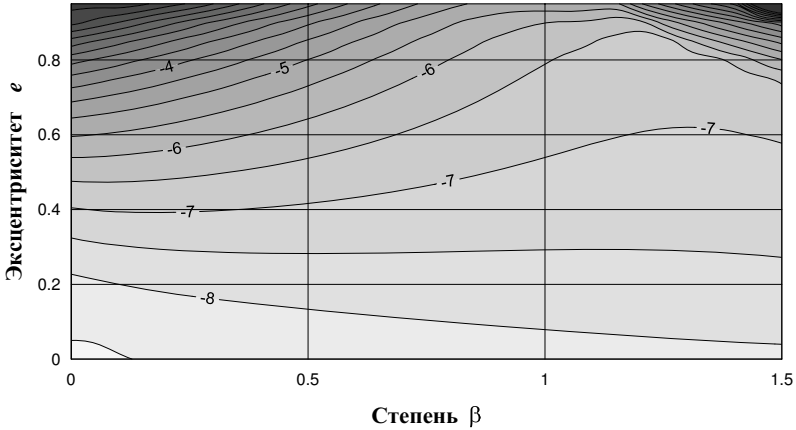


Рис. 3.4. То же, что и на рис. 3.2, но для $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2}$

интегрирования, поскольку теоретически $f = \text{const}$. Тем не менее дифференциальные уравнения, описывающие круговое движение, принципиально разные и на практике величина f как функция интегрируемых переменных вследствие ошибок не является постоянной, а потому точность интегрирования будет меняться в зависимости от выбираемых значений β (впрочем, как и от выбираемых методов интегрирования).

Между тем нестрогий анализ результатов позволяет определиться в следующем выборе. При численном моделировании сильновытянутых орбит (т.е. когда имеются тесные сближения с центральным телом) следует рекомендовать: $\beta = 3/2$ для $f = |\mathbf{x}|^\beta$; $\beta = 1 - 3/2$ для $f = |\dot{\mathbf{x}}|^{-\beta}$; $\beta = 3/4 - 1$ для $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2}$. Иначе говоря, эллиптическая аномалия в качестве независимой переменной — оптимальный выбор среди других аномалий и, кроме того, для эффективного сглаживания уравнений возможен выбор дуг физической и фазовой орбит ($\beta = 1$). Хотя, если учесть, что $(\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2} \sim |\mathbf{x}|^{2\beta}$ при $|\mathbf{x}| \rightarrow 0$ (что, кстати, объясняет похожесть характеристик на рис. 3.2 и 3.4), для последнего преобразования целесообразен выбор $\beta = 3/4$.

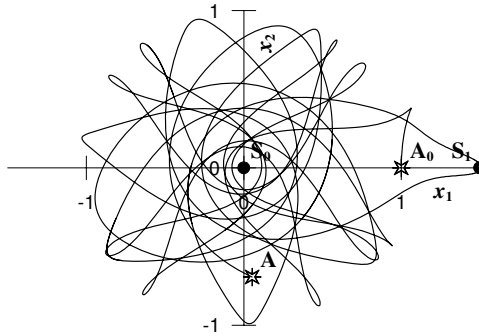


Рис. 3.5. Орбита третьего тела \mathbf{A} (\mathbf{A}_0 — начальное положение) в ограниченной задаче трех тел \mathbf{S}_0 – \mathbf{S}_1 – \mathbf{A} во вращающейся системе координат

В астероидных задачах, где исследование орбиты астероида усложняется заметным влиянием от больших планет, например, при тесных сближениях, можно использовать следующее преобразование (Авдюшев, 2006b):

$$dt = \left(\sum_{i=0}^N \frac{M_i^\alpha}{|\mathbf{x} - \mathbf{x}_i|^\beta} \right)^{-1} ds \equiv f ds. \quad (3.14)$$

где α и β — параметры преобразования; N — число планет; M_i и \mathbf{x}_i — соответственно масса и вектор положения i -й планеты, причем масса Солнца $M_0 = 1$, а $\mathbf{x}_0 = \mathbf{0}$. Интересно заметить, что при тесном сближении с k -й планетой ($|\mathbf{x} - \mathbf{x}_k| \rightarrow 0$) $f \sim |\mathbf{x} - \mathbf{x}_k|^\beta$, поэтому временное преобразование (3.14) вырождается в преобразование типа $dt = |\mathbf{x}|^\beta ds$, записанное относительно планеты, выступающей в качестве центрального тела.

В ограниченной задаче трех тел ($N = 1$, $M_0 = 1$, $M_1 = 0.1$; $\min |\mathbf{x}| = 0.075$, $\min |\mathbf{x} - \mathbf{x}_1| = 0.010$; орбита представлена на рис. 3.5) установлено, что оптимальными параметрами являются $\alpha = 1$ и $\beta = 3/2$. В задаче также тестировались преобразования с функциями $f = |\dot{\mathbf{x}}|^{-\beta}$ и $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-\beta/2}$. Первое дало очень

низкую эффективность для любых β . Второе преобразование оказалось столь же хорошим как и (3.14). Его наивысшая точность достигается при значении параметра β , близком к $3/4$, почти как в случае интегрирования эллиптических орбит. Поскольку это преобразование значительно проще (3.14) с точки зрения программной реализации, именно его при оптимальном параметре $\beta = 3/4$ можно рекомендовать в качестве сглаживающего для численного моделирования сложных орбит.

3.3. Стабилизированные уравнения

Как уже отмечалось, ляпуновская неустойчивость дифференциальных уравнений при численном интегрировании создает благоприятные условия для культивирования всевозможных ошибок, неизбежно сопровождающих любой численный процесс. Ошибки на текущем шаге интегрирования становятся ошибками начальных данных следующего, которые в дальнейшем усиливаются неустойчивостью шаг за шагом. Поэтому устойчивые дифференциальные уравнения для численного интегрирования более предпочтительны.

Задача стабилизации заключается как раз в том, чтобы ослабить (или в лучшем случае устранить) влияние ляпуновской неустойчивости на численное решение и улучшить таким образом поведение неустраимых ошибок интегрирования.

Родоначальниками стабилизации в небесной механике бесспорно можно считать Й. Баумгарте (Baumgarte, 1972a; Baumgarte, 1972b) и П. Накози (Nacozy, 1971). Их стабилизирующие методы основаны на применении известных интегралов, которые содержат дополнительную информацию о решении и рассматриваются как ограничения, налагаемые на решение.

Технически стабилизация достигается путем исправления численного решения за его отклонение от интегральной поверхности в фазовом пространстве интегрируемых переменных. В возмущенной задаче, когда интегральная поверхность динамична, а ее интегральный параметр становится переменным, для оценки отклоне-

ний решения используются те же интегральные соотношения, однако система уравнений дополняется уравнением для интегрального параметра, которое интегрируется численно совместно со всей системой (Baumgarte, 1972b; Бордовицына, Сухоплюева, 1980).

Стабилизация применяется как непосредственно к самому решению в процессе интегрирования (Nacozy, 1971), так и посредством введения дополнительных так называемых стабилизирующих членов в дифференциальные уравнения (Baumgarte, 1972b). Очевидно, в этом случае стабилизация тем эффективнее, чем слабее возмущения и чем медленнее меняются параметры опорных интегральных поверхностей. Из всех интегральных соотношений авторы стабилизирующих методов выделяют энергетические, поскольку именно стабилизация по энергии, как показывает практика, является наилучшим подспорьем в борьбе с ляпуновской неустойчивостью.

Исторически так сложилось, что методы стабилизации не были удостоены должным вниманием со стороны небесных механиков, несмотря на то, что их исключительные возможности открываются именно в задачах небесной механики и прежде всего в задачах исследования долговременной эволюции орбитального движения. В результате ниша эволюционных задач была занята симплектическими интеграторами, которые по сути обладают достоинствами стабилизирующих методов, но при этом значительно сложнее для практической реализации. Вместе с тем, как только появились первые стабилизирующие методы, они стали широко применяться в классической механике, где и продолжают развиваться до настоящего времени¹³.

Тем не менее, конечно же, нельзя не отметить некоторые не связанные друг с другом работы после Баумгарте и Накози (Му-

¹³Стоит, однако, сказать, что, по-видимому, в отличие от Баумгарте, Накози не уделил должного внимания своему методу и особо не предпринял какие-либо усилия для его популяризации. В результате метод был успешно забыт и фактически переоткрыт заново спустя некоторое время в классической механике, где он сейчас известен как метод постстабилизации (Chin, 1995) или проекционный метод (Hairer et al., 2002).

gison, 1989; Avdyushev, 2003; Fukushima, 2003; Zhu, Wu, 2007), в которых авторы раскрывают достоинства стабилизации применительно к численному моделированию орбитальной динамики.

Несмотря на то что приоритет в изобретении численной стабилизации приписывается Баумгарте и Накози, первые стабилизированные уравнения орбитального движения были предложены гораздо раньше них, еще Ж.Л. Лагранжем. Действительно, известные дифференциальные уравнения Лагранжа в орбитальных элементах обладают стабилизирующими свойствами: элементы как интегральные параметры в невозмущенном движении постоянные величины и численно они интегрируются без методических ошибок, поэтому численное решение удовлетворяет всем интегральным соотношениям, а в слабозвозмущенном движении методические ошибки достаточно малы, за счет чего, собственно, и достигается стабилизирующий эффект. Тем не менее, когда речь заходит о стабилизации, Лагранжа никто не вспоминает, поскольку его уравнения по сути не являются продуктом целенаправленного решения проблемы неустойчивости.

Уместно упомянуть и другие стабилизированные уравнения, которые появились в то же время, что и уравнения Баумгарте, но со стабилизацией обычно не ассоциируются, поскольку создавались авторами для достижения иных свойств: это, конечно же, уже рассмотренные нами уравнения в SB- (3.9) и KS-переменных (3.11), а также уравнения Роя (3.35), в SB- (3.36) и KS-элементах (3.37) из следующего разд. 3.4, где переменная энергии либо связанные с ней интегрируемые переменные рассматриваются как самостоятельные.

Хотелось бы также отметить, что стабилизацию не следует рассматривать как нечто привнесенное из математики, как сугубо математический прием. Она также существует в природе. Однако ни Баумгарте, ни Накози, если судить по их работам, не воспользовались подсказками природы, и пришли к своим оригинальным идеям в результате формального решения проблемы неустойчивости, хотя в то же время механизмы математической и физической стабилизаций очень схожи.

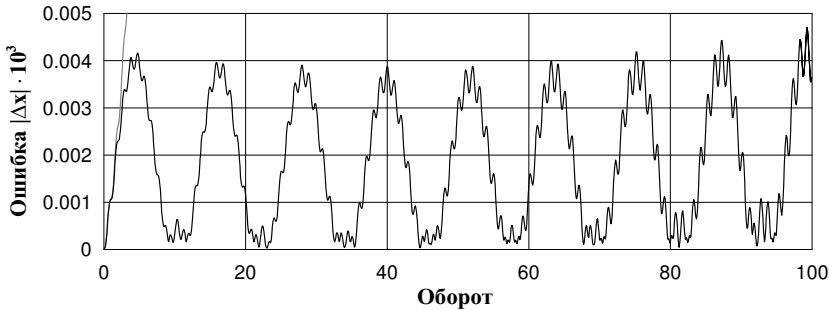


Рис. 3.6. Поведение ошибки в векторе положения при численном интегрировании возмущенного движения около треугольной (устойчивой) равновесной точки классическим методом Рунге–Кутты с шагом $\Delta t = 2\pi/128$. Возмущающее тело массы $M_1 = 0.001$ ($M_0 = 1$) движется по круговой орбите единичного радиуса с единичной частотой. Серым цветом показана точность при отсутствии возмущений

Стабилизирующие эффекты, например, имеют место в движении троянских астероидов в окрестности устойчивых равновесных точек (Дубошин, 1964), вызванных гравитацией соответствующих больших планет. На рис. 3.6 показано поведение ошибки численного моделирования возмущенного движения в прямоугольных координатах около устойчивой равновесной точки в ограниченной безразмерной задаче трех тел. Как видно из рисунка, возмущающая сила от второго тела играет роль стабилизатора, и ошибка интегрирования возрастает существенно медленнее, нежели даже в невозмущенном случае (см. серая кривая). Более подробный анализ показывает, что ее рост приблизительно такой же, как и при численном интегрировании устойчивых уравнений гармонического осциллятора, описывающих круговую орбиту единичного радиуса. Здесь уместно упомянуть известный факт, что если даже невозмущенное (кеплеровское) движение неустойчиво, то еще нельзя утверждать, что неустойчивым будет и возмущенное дви-

жение, какими бы малыми не были возмущения. Кстати, этот факт подтверждает диссипативная стабилизация Баумгарте.

3.3.1. Неустойчивость кеплеровского движения

Пользуясь формулами задачи двух тел, нетрудно показать (Audyushev, 2003), что если средние движения двух близких кеплеровских решений отличаются на величину Δn в некоторый начальный момент времени t_0 , то при $t - t_0 \rightarrow \infty$ это приводит к расхождению решений $\Delta \mathbf{x}$ и $\Delta \dot{\mathbf{x}}$ в соответствии с приближенной оценкой

$$\Delta \mathbf{x} \approx \frac{\Delta n}{n} \dot{\mathbf{x}}(t - t_0), \quad \Delta \dot{\mathbf{x}} \approx \frac{\Delta n}{n} \ddot{\mathbf{x}}(t - t_0), \quad (3.15)$$

где n — среднее движение одной из орбит. Учитывая, что среднее движение непосредственно зависит от энергии H и с точностью до малых первого порядка для отклонений энергетических переменных справедливо равенство $2H\Delta n = 3n\Delta H$, формулу (3.15) можно переписать как

$$\Delta \mathbf{x} \approx \frac{3}{2} \frac{\Delta H}{H} \dot{\mathbf{x}}(t - t_0), \quad \Delta \dot{\mathbf{x}} \approx \frac{3}{2} \frac{\Delta H}{H} \ddot{\mathbf{x}}(t - t_0). \quad (3.16)$$

В частности, из (3.15) следует, что

$$|\Delta \mathbf{x}| \approx \frac{\Delta n}{n} |\dot{\mathbf{x}}|(t - t_0). \quad (3.17)$$

Оценка (3.17) показывает, что эволюция $|\Delta \mathbf{x}|$ носит периодически вековой характер: расхождение положений на орбитах со временем растет по величине, то локально увеличиваясь к перицентру, когда скорость возрастает, то локально уменьшаясь к апоцентру, когда скорость убывает. Также видно, что средняя скорость роста $|\Delta \mathbf{x}|$ явно зависит от относительного отклонения в среднем движении: чем меньше $\Delta n/n$, тем ниже скорость $|\Delta \mathbf{x}|$. Однако какой бы малой ни была Δn , всегда наступит такой момент времени, когда $|\Delta \mathbf{x}|$ станет недопустимо большой. Фактически это

говорит о том, что кеплеровское движение неустойчиво по Ляпунову. Впрочем, если для двух орбит $\Delta n = 0$, то расхождение положений $|\Delta \mathbf{x}|$ будет ограничено. Этот факт как раз принимается во внимание для построения эффективных методов стабилизации, применяемых в численном интегрировании орбит.

Если одно из двух решений рассматривать как точное, а другое — ошибочное, обусловленное начальными ошибками $\Delta \mathbf{x}_0$ и $\Delta \dot{\mathbf{x}}_0$, которые вызывают ошибку в среднем движении Δn , то формулу (3.17) можно использовать как оценку ошибок $\Delta \mathbf{x}$ и $\Delta \dot{\mathbf{x}}$ на любой другой момент времени. При численном интегрировании кеплеровской орбиты малые ошибки типа $\Delta \mathbf{x}_0$ и $\Delta \dot{\mathbf{x}}_0$ возникают на каждом шаге интегрирования, вследствие чего ошибка в среднем движении растет почти линейным образом (рис. 3.7).

Допустим, что $\Delta n = \kappa(t - t_0)$, где $\kappa = \text{const}$, тогда, согласно (3.17), величина ошибки в векторе положения будет

$$|\Delta \mathbf{x}| \approx \frac{\kappa}{n} \int_{t_0}^t (t - t_0) |\dot{\mathbf{x}}| dt.$$

Следовательно,

$$\mathcal{Q}(t-t_0) \sqrt{\frac{1-e}{1+e}} \leq |\Delta \mathbf{x}| \leq \mathcal{Q}(t-t_0) \sqrt{\frac{1+e}{1-e}}, \quad \mathcal{Q}(t-t_0) \approx \frac{a\kappa}{2}(t-t_0)^2, \quad (3.18)$$

где a и e — соответственно большая полуось и эксцентриситет орбиты.

В соответствии с (3.18) в круговом случае ($e = 0$) ошибка в векторе положения $|\Delta \mathbf{x}|$ будет вести себя квадратичным образом: $|\Delta \mathbf{x}| \approx \mathcal{Q}(t-t_0)$. В эллиптическом случае квадратичное поведение ошибки сопровождается колебаниями с орбитальным периодом и увеличивающейся амплитудой (рис. 3.7), причем $|\Delta \mathbf{x}|$ изменяется между минорантой и мажорантой (3.18). Подобным образом ошибка $|\Delta \mathbf{x}|$ ведет себя и при численном интегрировании слабо-возмущенной орбиты.

Фактически, квадратичное поведение $|\Delta \mathbf{x}|$ вызывается линейным ростом Δn . Таким образом, чтобы ослабить влияние ляпу-

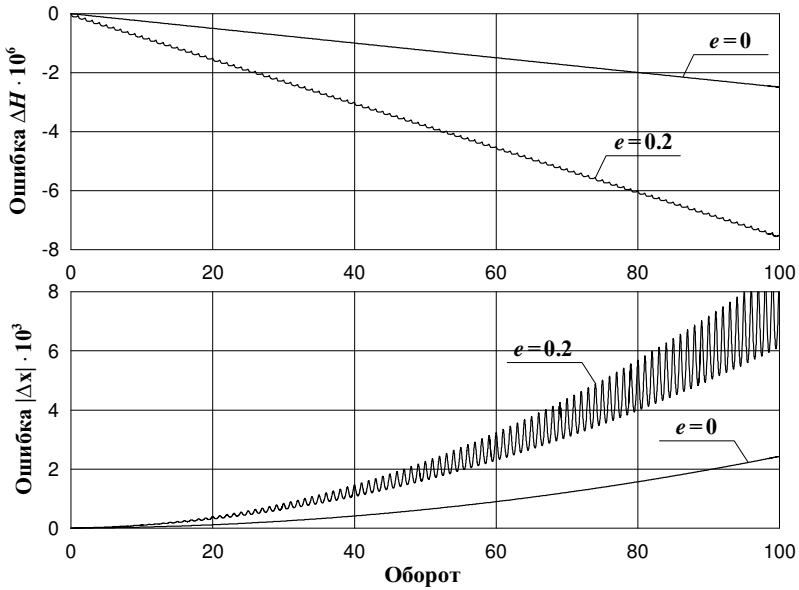


Рис. 3.7. Поведение ошибок в векторе положения и в энергии при численном интегрировании круговой и эллиптической ($e = 0.2$) орбит ($a = 1$, $n = 1$) классическим методом Рунге–Кутты ($\Delta t = 2\pi/128$)

новской неустойчивости на численное решение и уменьшить скорость роста ошибки в векторе положения, необходимо в процессе интегрирования тем или иным способом подавлять ошибку в среднем движении (или в любой другой энергетической переменной).

3.3.2. Диссипативная стабилизация Баумгарте

Простой, но мощный метод численной стабилизации был предложен Й. Баумгарте (Baumgarte, 1972a; Baumgarte, 1972b). Метод Баумгарте основан на идее искусственного введения в дифференциальные уравнения движения (3.1) так называемых стабилизирующих (возмущающих) членов, компенсирующих отклонения

численного (ошибочного) решения от опорной интегральной поверхности, задаваемой энергией. При численном интегрировании таких уравнений решение будет оставаться около интегральной поверхности, что обеспечивает ограничение ошибки в энергии.

Расширенная система уравнений возмущенного движения (3.1)

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{\mu}{|\mathbf{x}|^3}\mathbf{x} + \mathbf{P}, \quad \frac{dh}{dt} = \dot{\mathbf{x}} \cdot \mathbf{P} \quad (3.19)$$

имеет следующее интегральное соотношение:

$$\mathcal{H} = H(\mathbf{x}, \dot{\mathbf{x}}) - h = \frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|} - h \equiv 0, \quad (3.20)$$

которое при численном интегрировании не выполняется вследствие ошибок в интегрируемых переменных \mathbf{x} , $\dot{\mathbf{x}}$ и h , т.е. $\mathcal{H} \neq 0$. В слабозвозмущенном движении энергетический параметр h интегрируется с высокой точностью, поэтому значения \mathcal{H} будут определяться главным образом ошибками в энергии H , вызванными ошибками в положении \mathbf{x} и скорости $\dot{\mathbf{x}}$.

Согласно Баумгарте, система стабилизированных уравнений принимает вид

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{\mu}{|\mathbf{x}|^3}\mathbf{x} + \mathbf{P} - \gamma\mathcal{H}\frac{\dot{\mathbf{x}}}{\dot{\mathbf{x}}^2}, \quad \frac{dh}{dt} = \dot{\mathbf{x}} \cdot \mathbf{P}, \quad \mathcal{H} = \frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|} - h, \quad (3.21)$$

где $\gamma = \text{const}$ — так называемый стабилизирующий параметр. Из (3.21) нетрудно показать, что \mathcal{H} удовлетворяет уравнению

$$\frac{d\mathcal{H}}{dt} = -\gamma\mathcal{H}, \quad \text{следовательно,} \quad |\mathcal{H}| = |\mathcal{H}_0|e^{-\gamma(t-t_0)}. \quad (3.22)$$

Отсюда, если $\gamma > 0$, то $|\mathcal{H}| \rightarrow 0$ при $t - t_0 \rightarrow \infty$, что фактически указывает на асимптотическую устойчивость уравнений по H . Поэтому при численном интегрировании уравнений (3.21) любые отклонения H относительно h , вызываемые ошибками в \mathbf{x} и $\dot{\mathbf{x}}$, будут сходиться на нет. Для того чтобы система оставалась асимптотически устойчивой при ретроспективном интегрировании ($t < t_0$),

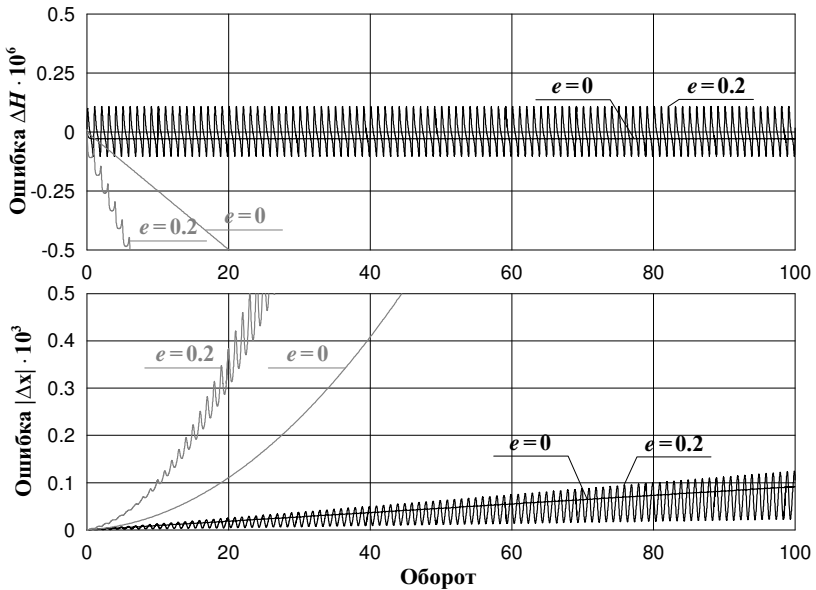


Рис. 3.8. Поведение ошибок в векторе положения и в энергии при численном интегрировании круговой и эллиптической стабилизированных орбит ($\gamma = 1$, $a = 1$, $n = 1$) классическим методом Рунге–Кутты ($\Delta t = 2\pi/128$) (серым цветом показаны результаты рис. 3.7)

параметр γ должен быть отрицательным. Таким образом, введенные в уравнения движения стабилизирующие возмущения будут удерживать вычисляемую орбиту около интегральной поверхности $H(\mathbf{x}, \dot{\mathbf{x}}) = h$, что, очевидно, позволяет ограничить возникающие при интегрировании ошибки в энергии и улучшить тем самым поведение ошибок в \mathbf{x} и $\dot{\mathbf{x}}$ (см. рис. 3.8).

Заметим, что стабилизация Баумгарте может быть выполнена по любому интегральному соотношению вида $\mathcal{H} = H(\mathbf{x}, \dot{\mathbf{x}}) - h \equiv 0$. В этом случае вид системы стабилизированных уравнений будет

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{\mu}{|\mathbf{x}|^3}\mathbf{x} + \mathbf{P} - \gamma\mathcal{H} \left(\frac{\partial\mathcal{H}}{\partial\dot{\mathbf{x}}} \right)^T \left(\frac{\partial\mathcal{H}}{\partial\dot{\mathbf{x}}} \cdot \frac{\partial\mathcal{H}}{\partial\dot{\mathbf{x}}} \right)^{-1},$$

$$\frac{dh}{dt} = \frac{\partial H}{\partial \dot{\mathbf{x}}} \cdot \mathbf{P}, \quad \mathcal{H} = H(\mathbf{x}, \dot{\mathbf{x}}) - h.$$

В круговом случае можно получить приближенные оценки для \mathcal{H} и $|\Delta \mathbf{x}|$. При интегрировании нестабилизированных уравнений кругового движения ошибка в энергии ΔH ведет себя линейно (рис. 3.7): $\mathcal{H} = \Delta H = K(t - t_0)$, поэтому \mathcal{H} будет удовлетворять уравнению

$$\frac{d\mathcal{H}}{dt} = K. \quad (3.23)$$

Тогда величина ошибки в векторе положения $|\Delta \mathbf{x}|$, согласно (3.16), будет развиваться как

$$|\Delta \mathbf{x}| \approx \frac{3}{4} \frac{|\dot{\mathbf{x}}|}{H} K(t - t_0)^2. \quad (3.24)$$

Учитывая (3.22) и (3.23), будем иметь уравнение для \mathcal{H} при интегрировании стабилизированной системы (3.21) (Авдюшев, 2003а):

$$\frac{d\mathcal{H}}{dt} = K - \gamma \mathcal{H}, \quad \text{следовательно,} \quad \mathcal{H} = K \frac{1 - e^{-\gamma(t-t_0)}}{\gamma}. \quad (3.25)$$

Отсюда при $\gamma > 0$ и $t \rightarrow \infty$ величина \mathcal{H} стремится к постоянному значению K/γ . В соответствии с (3.25) оценку величины ошибки в векторе положения можно представить в следующем виде:

$$\begin{aligned} |\Delta \mathbf{x}| &\approx \frac{3}{2} \frac{K}{\gamma} \frac{|\dot{\mathbf{x}}|}{H} \int_{t_0}^t [1 - e^{-\gamma(t-t_0)}] dt = \\ &= \frac{3}{2} \frac{K}{\gamma} \frac{|\dot{\mathbf{x}}|}{H} \left[(t - t_0) - \frac{1 - e^{-\gamma(t-t_0)}}{\gamma} \right]. \end{aligned}$$

Поэтому

$$|\Delta \mathbf{x}| \approx \frac{3}{2} \frac{K}{\gamma} \frac{|\dot{\mathbf{x}}|}{H} (t - t_0) \quad \text{при} \quad t \rightarrow \infty. \quad (3.26)$$

С другой стороны, нетрудно видеть, что оценка (3.26) вырождается в (3.24), если параметр γ устремить к нулю. Сравнивая (3.24) и

(3.26), можно получить характеристику $\delta = \gamma(t - t_0)/2$, указывающую, во сколько раз может повыситься точность интегрирования при введении в уравнения стабилизирующих возмущений. Очевидно, полученные оценки будут справедливы и в случае почти круговых слабозвозмущенных орбит.

Как показывает практика, оценка (3.25) достаточно хорошо представляет поведение \mathcal{H} только при использовании численных методов высоких порядков (Авдюшев, 2003а), тогда как оценка (3.26) учитывает лишь те ошибки в положении, которые непосредственно вызваны ошибкой ΔH и развиваемые ляпуновской неустойчивостью. Так или иначе, оценки вполне можно использовать, по крайней мере, для качественного описания поведения ошибок ΔH и $|\Delta \mathbf{x}|$ в зависимости от степени влияния стабилизирующих членов.

Особо заметим, что оценки создают превратное впечатление, что ошибки ΔH и $|\Delta \mathbf{x}|$ можно свести на нет, увеличивая параметр γ . В действительности при достаточно больших γ уравнения движения становятся жесткими (Chin, 1995; Ascher et al., 1995), что вызывает определенные сложности для численного интегрирования, в особенности при использовании интеграторов высоких порядков (Авдюшев, 2003а). Кроме того, увеличение стабилизирующего параметра приводит к большим стабилизирующим возмущениям, способным существенно исказить динамическую картину, описываемую исходными уравнениями (3.1).

Вообще говоря, до численного интегрирования стабилизирующий параметр γ не известен и его следует выбирать экспериментально по достижении наилучших результатов, однако в (Авдюшев, 2003а; Avdyushev, 2003) показано, что при исследовании слабозвозмущенных почти круговых орбит в качестве оптимального стабилизирующего параметра следует брать среднее движение n . Попутно заметим, что при использовании интеграторов высоких порядков оценки (3.25) и (3.26) будут хорошо представлять поведение ошибок, если $\gamma \in [0, n]$.

3.3.3. Консервативная стабилизация Баумгарте

Й. Баумгарте (Baumgarte, 1976a; Baumgarte, 1976b) предложил другой метод стабилизации, который также преобразует уравнения движения, но в отличие от (3.21) получаемые стабилизированные уравнения не обременены какими-либо неопределенностями.

Согласно (3.15), если Δn — величина переменная, оценки для $\Delta \mathbf{x}$ и $\Delta \dot{\mathbf{x}}$ будут

$$\begin{pmatrix} \Delta \mathbf{x} \\ \Delta \dot{\mathbf{x}} \end{pmatrix} = \int_{t_0}^t \frac{\Delta n}{n} \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix} dt, \quad \text{поэтому} \quad \frac{d}{dt} \begin{pmatrix} \Delta \mathbf{x} \\ \Delta \dot{\mathbf{x}} \end{pmatrix} = \frac{\Delta n}{n} \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix}.$$

Следовательно, уравнения для стабилизированных решений можно представить в виде (Авдюшев, 2006b)

$$\frac{d}{dt} \begin{pmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix} - \frac{d}{dt} \begin{pmatrix} \Delta \mathbf{x} \\ \Delta \dot{\mathbf{x}} \end{pmatrix} = \bar{n} \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix} = \left(\frac{h}{H} \right)^{3/2} \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix}, \quad (3.27)$$

где $\bar{n} = n - \Delta n$ — опорное среднее движение (точной орбиты). Тогда общая стабилизированная система уравнений движения будет

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{pmatrix} &= \begin{pmatrix} h \\ H \end{pmatrix}^{3/2} \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix}, \quad \frac{dh}{dt} = (\dot{\mathbf{x}} \cdot \mathbf{P}), \\ \ddot{\mathbf{x}} &= \mathbf{F} + \mathbf{P}, \quad H = \frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|}. \end{aligned} \quad (3.28)$$

Уравнения (3.28) впервые вывел Баумгарте (Baumgarte, 1976b), хотя и несколько иным способом. Данный метод стабилизации, в отличие от предыдущего, не приводит к изменению энергии, поэтому автор назвал его консервативным. К сожалению, в задачах небесной механики консервативный метод оказался неэффективным. Это главным образом связано с наличием в уравнениях (3.28) множителя $(h/H)^{3/2}$, который усложняет поведение правых частей.

3.3.4. Стабилизация по времени

От нежелательного множителя в (3.28) можно избавиться, если выполнять интегрирование по фиктивному времени \bar{t} , которое связано с истинным временем t как (Авдюшев, 2006b)

$$dt = d\bar{t} \left(\frac{H}{h} \right)^{3/2}. \quad (3.29)$$

В итоге новые уравнения примут вид

$$\frac{d}{d\bar{t}} \begin{pmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{pmatrix}, \quad \frac{dh}{d\bar{t}} = (\dot{\mathbf{x}} \cdot \mathbf{P}) \left(\frac{H}{h} \right)^{3/2}, \quad \frac{dt}{d\bar{t}} = \left(\frac{H}{h} \right)^{3/2}. \quad (3.30)$$

Недостатком такой стабилизации является то, что в систему уравнений необходимо включать дополнительное уравнение для t . Кроме того, поскольку t рассматривается как интегрируемая переменная, возникает проблема выхода на заданный момент истинного времени. Эту проблему можно решить следующим образом. Сначала интегрируем систему уравнений до значения независимой переменной \bar{t} , равного заданному моменту времени t . Если в результате интегрирования величина разности $t - \bar{t}$ оказывается достаточно малой, находим решение по приближенной формуле: $\mathbf{x}(t) = \mathbf{x}(\bar{t}) + \dot{\mathbf{x}}(\bar{t})(t - \bar{t}) + \ddot{\mathbf{x}}(\bar{t})(t - \bar{t})^2$. Если $t - \bar{t}$ — большая величина, выполняем интегрирование от \bar{t} на интервале $t - \bar{t}$ и так далее, пока разность $t - \bar{t}$ не станет пренебрежимо малой.

Впрочем, уравнение времени можно исключить из системы (3.30), если предположить, что ошибка ΔH ведет себя линейно. Тогда, интегрируя уравнение аналитически, получим связь между временами t и \bar{t} :

$$t = \bar{t} + \frac{3}{4} \frac{H - h}{h} (\bar{t} - t_0). \quad (3.31)$$

Даже если поведение ошибки ΔH не строго линейно, использование (3.31) вместо интегрирования соответствующего дифференциального уравнения будет также полезным, поскольку в разности $t - \bar{t}$ будет учитываться главная квадратическая составляющая, вызываемая ляпуновской неустойчивостью.

3.3.5. Стабилизация Накози

Альтернативный способ сохранения интегральных соотношений был предложен П. Накози (Nacozy, 1971). Следуя Накози, уравнения (3.19) интегрируются обычным образом, однако после нескольких шагов интегрирования выполняется коррекция переменных \mathbf{x} и $\dot{\mathbf{x}}$ за отклонение H от h по приближенной формуле

$$\bar{\mathbf{x}} = \mathbf{x} \left(1 - \frac{\mathcal{H}}{D} \frac{\mu}{|\mathbf{x}|^3} \right), \quad \dot{\bar{\mathbf{x}}} = \dot{\mathbf{x}} \left(1 - \frac{\mathcal{H}}{D} \right), \quad D = \dot{\mathbf{x}}^2 + \frac{\mu^2}{|\mathbf{x}|^4}, \quad (3.32)$$

где $\bar{\mathbf{x}}$ и $\dot{\bar{\mathbf{x}}}$ — исправленные (стабилизированные) значения интегрируемых переменных. Если $\mathcal{H} = H(\bar{\mathbf{x}}, \dot{\bar{\mathbf{x}}}) - h \neq 0$, процедуру коррекции (3.32) выполняют снова при $\mathbf{x} = \bar{\mathbf{x}}$ и $\dot{\mathbf{x}} = \dot{\bar{\mathbf{x}}}$, и так далее до выполнения равенства с заданной точностью. Затем продолжают интегрирование, отправляясь от значений $\bar{\mathbf{x}}$ и $\dot{\bar{\mathbf{x}}}$.

Займствуя идею Накози, Т. Фукушима (Fukushima, 2003) предложил подобный метод коррекции интегрируемых переменных:

$$\bar{\mathbf{x}} = \mathbf{x}\mathcal{S}, \quad \dot{\bar{\mathbf{x}}} = \dot{\mathbf{x}}\mathcal{S},$$

где \mathcal{S} определяется кубическим уравнением

$$\frac{\dot{\mathbf{x}}^2}{2} \mathcal{S}^3 - h\mathcal{S} - \frac{\mu}{|\mathbf{x}|} = 0.$$

Несмотря на то что коррекция Фукушимы отличается от (3.32), она также обеспечивает сохранение интегрального соотношения $\mathcal{H} = 0$.

3.4. Уравнения в орбитальных элементах

Метод вариации произвольных постоянных, впервые предложенный Л. Эйлером (Lovett, 1899) и детально разработанный Ж.Л. Лагранжем (Lagrange, 1783), изначально применялся в теории обших возмущений, хотя он успешно используется и при численном интегрировании.

В методе вариации постоянных орбита представляется в невозмущенном виде, где орбитальные параметры (элементы) рассматриваются как переменные. В соответствии с формой представления орбиты составляются дифференциальные уравнения для орбитальных параметров, которые затем интегрируются приближенными методами. Следует заметить, что методику Лагранжа можно применять не только к постоянным параметрам, но также и к различным функциям этих параметров и времени.

В случае малых и гладких возмущающих сил параметры — медленноменяющиеся переменные, поэтому их уравнения при использовании численных методов будут интегрироваться гораздо эффективнее, нежели классические уравнения (3.1).

Рассмотрим основные положения метода Лагранжа. Предположим, что невозмущенная \mathbf{p}_K и возмущенная \mathbf{p} орбиты небесного тела описываются уравнениями первого порядка

$$\frac{d\mathbf{p}_K}{dt} = \mathcal{P}_K(t, \mathbf{p}_K), \quad \frac{d\mathbf{p}}{dt} = \mathcal{P}(t, \mathbf{p}) = \mathcal{P}_K(t, \mathbf{p}) + \mathcal{P}(t, \mathbf{p}) - \mathcal{P}_K(t, \mathbf{p}), \quad (3.33)$$

причем известен закон невозмущенного движения $\mathbf{p}_K = \mathbf{p}_K(t, \mathbf{q})$, где \mathbf{q} — орбитальные параметры суть постоянные. Представим возмущенное движение как $\mathbf{p} = \mathbf{p}_K(t, \mathbf{q})$, где, однако, орбитальные параметры уже переменные величины. Тогда

$$\frac{d\mathbf{p}}{dt} = \frac{\partial \mathbf{p}_K}{\partial t} + \frac{\partial \mathbf{p}_K}{\partial \mathbf{q}} \frac{d\mathbf{q}}{dt} = \mathcal{P}_K(t, \mathbf{p}) + \frac{\partial \mathbf{p}_K}{\partial \mathbf{q}} \frac{d\mathbf{q}}{dt}$$

и в соответствии с (3.33)

$$\frac{\partial \mathbf{p}_K}{\partial \mathbf{q}} \frac{d\mathbf{q}}{dt} = \mathcal{P}(t, \mathbf{p}) - \mathcal{P}_K(t, \mathbf{p}). \quad (3.34)$$

Выражая из (3.34) временные производные от \mathbf{q} , получим в итоге уравнения в орбитальных параметрах.

Далее представим наиболее подходящие с точки зрения численного интегрирования дифференциальные уравнения в орбитальных элементах, а именно уравнения Роя и уравнения в регулярных SB- и KS-элементах.

3.4.1. Уравнения Роя

С вычислительной точки зрения уравнения в кеплеровских элементах неудобны тем, что они довольно сложны, содержат много тригонометрических функций и, кроме того, имеют особенность для нулевых наклонений и эксцентриситетов. Чтобы обойти эти недостатки, А. Рой предложил в качестве орбитальных параметров рассматривать момент количества движения \mathbf{c} , вектор Лапласа \mathbf{g} и истинную долготу $\lambda = \Omega + \omega + v$ (см. Приложение). Их уравнения можно записать в виде (Рой, 1981)

$$\frac{d\mathbf{c}}{dt} = \mathbf{x} \times \mathbf{P}, \quad \frac{d\mathbf{g}}{dt} = \mathbf{P} \times \mathbf{c} + \dot{\mathbf{x}} \times \dot{\mathbf{c}}, \quad \frac{d\lambda}{dt} = \frac{|\mathbf{c}|}{|\mathbf{x}|^2} + \frac{c_1 \dot{c}_2 - c_2 \dot{c}_1}{|\mathbf{c}|(|\mathbf{c}| + c_3)}. \quad (3.35)$$

При интегрировании уравнений (3.35) для вычисления возмущающей силы \mathbf{P} требуется переход от переменных \mathbf{c} , \mathbf{g} и λ к \mathbf{x} и $\dot{\mathbf{x}}$. Он выполняется по формулам

$$\begin{aligned} \gamma &= g_1 - \frac{g_3 c_1}{|\mathbf{c}| + c_3}, & \delta &= g_2 - \frac{g_3 c_2}{|\mathbf{c}| + c_3}; \\ |\mathbf{g}| \cos v &= \gamma \cos \lambda + \delta \sin \lambda, & |\mathbf{g}| \sin v &= \gamma \sin \lambda - \delta \cos \lambda; \\ |\mathbf{x}| &= \frac{|\mathbf{c}|^2}{\mu + |\mathbf{g}| \cos v}, & \mathbf{x} \cdot \dot{\mathbf{x}} &= \frac{|\mathbf{x}| |\mathbf{g}| \sin v}{|\mathbf{c}|}; \\ x_1 &= |\mathbf{x}| \cos \lambda + \frac{x_3 c_1}{|\mathbf{c}| + c_3}, & x_2 &= |\mathbf{x}| \sin \lambda + \frac{x_3 c_2}{|\mathbf{c}| + c_3}, \\ x_3 &= -\frac{|\mathbf{x}|}{|\mathbf{c}|} (c_1 \cos \lambda + c_2 \sin \lambda); \\ \dot{\mathbf{x}} &= \frac{\mathbf{c} \times \mathbf{x} + (\mathbf{x} \cdot \dot{\mathbf{x}}) \mathbf{x}}{|\mathbf{x}|^2}. \end{aligned}$$

Система (3.35) имеет серьезный недостаток, который проявляется главным образом при интегрировании сильновытянутых орбит. Последнее уравнение в (3.35) сингулярно в начале координат, что приводит к неравномерному поведению его правой части. Чтобы разрешить эту трудность, достаточно заменить λ на среднюю долготу $l = \Omega + \omega + M$, которая в слабозвмущенном движении меняется почти линейно.

3.4.2. Уравнения в SB-элементах

Весьма полезными для практики могут быть также уравнения в так называемых SB- (Авдюшев, Бордовицына, 2006) и KS-элементах (Stiefel, Scheifele, 1971). Если положение и скорость представить как решения невозмущенных уравнений (3.9):

$$\mathbf{x} = \boldsymbol{\alpha} \cos E + \boldsymbol{\beta} \sin E - \frac{\mathbf{g}}{4\omega^2}, \quad \mathbf{x}' = -\boldsymbol{\alpha} \sin E + \boldsymbol{\beta} \cos E,$$

тогда уравнения в SB-элементах будут

$$\boldsymbol{\alpha}' = \mathcal{G} \cos E - \mathcal{P} \sin E, \quad \boldsymbol{\beta}' = \mathcal{G} \sin E + \mathcal{P} \cos E,$$

$$\mathcal{P} = \frac{1}{4\omega^2} [|\mathbf{x}|^2 \mathbf{P} + (\mathbf{x}' \cdot \mathbf{P}) \mathbf{x}'], \quad \mathcal{G} = \frac{1}{4\omega^2} \left[\mathbf{g}' - 2\mathbf{g} \frac{\omega'}{\omega} \right],$$

$$\mathbf{g}' = 2\mathbf{x}(\mathbf{x}' \cdot \mathbf{P}) - \mathbf{x}'(\mathbf{x} \cdot \mathbf{P}) - \mathbf{P}(\mathbf{x} \cdot \mathbf{x}'), \quad \omega' = -\frac{1}{4\omega} (\mathbf{x}' \cdot \mathbf{P}), \quad (3.36)$$

$$\tau' = \frac{1}{8\omega^3} \left[\mu + |\mathbf{x}|(\mathbf{x} \cdot \mathbf{P}) + 2 \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|} (\mathbf{x}' \cdot \mathbf{P}) \right], \quad t = \tau - \frac{1}{2\omega} \frac{(\mathbf{x} \cdot \mathbf{x}')}{|\mathbf{x}|}.$$

В соответствии с формульным представлением положения и скорости начальные условия (при $E = 0$) для элементов $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ определяются как

$$\boldsymbol{\alpha}_0 = \mathbf{x}_0 + \frac{\mathbf{g}_0}{4\omega_0^2}, \quad \boldsymbol{\beta}_0 = \mathbf{x}'_0 = \frac{|\mathbf{x}|_0 \dot{\mathbf{x}}_0}{2\omega_0}.$$

3.4.3. Уравнения в KS-элементах

Для получения уравнений в KS-элементах представим положение и скорость в 4-мерном KS-пространстве (см. подразд. 3.1.2) в виде

$$\mathbf{u} = \boldsymbol{\alpha} \cos \frac{E}{2} + \boldsymbol{\beta} \sin \frac{E}{2}, \quad \mathbf{u}' = -\frac{1}{2} \boldsymbol{\alpha} \sin \frac{E}{2} + \frac{1}{2} \boldsymbol{\beta} \cos \frac{E}{2}.$$

Тогда уравнения орбитального движения в KS-элементах запишутся как

$$\boldsymbol{\alpha}' = -\mathcal{P} \sin \frac{E}{2}, \quad \boldsymbol{\beta}' = \mathcal{P} \cos \frac{E}{2}, \quad \omega' = -\frac{1}{2\omega} (\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}),$$

$$\tau' = \frac{1}{8\omega^3} [\mu + |\mathbf{u}|^2(\mathbf{u} \cdot \mathbf{L}^T \mathbf{P}) + 8(\mathbf{u} \cdot \mathbf{u}')(\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P})], \quad (3.37)$$

$$t = \tau - \frac{1}{\omega}(\mathbf{u} \cdot \mathbf{u}'), \quad \mathcal{P} = \frac{1}{2\omega^2} \left[\frac{|\mathbf{u}|^2}{2} \mathbf{L}^T \mathbf{P} + 2(\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}) \mathbf{u}' \right].$$

При этом $\boldsymbol{\alpha}_0 = \mathbf{u}_0$ и $\boldsymbol{\beta}_0 = 2\mathbf{u}'_0$.

Следует заметить, что элементы $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ в системах (3.36) и (3.37) принципиально разные (хотя бы потому что имеют разные размерности), несмотря на то что для них используются одинаковые обозначения.

3.4.4. Уравнения в элементах Пайнса

Координаты и скорости на начальную эпоху также можно рассматривать как орбитальные элементы некоторой невозмущенной (кеплеровской) орбиты, оскулирующей возмущенную не только на начальную эпоху, а на какой-либо произвольный момент времени вообще. Такую систему элементов в методе вариации параметров предложил использовать С. Пайнс (Pines, 1961). Выведем дифференциальные уравнения для этих столь необычных элементов.

Примем $\mathbf{p} = (\mathbf{x}, \dot{\mathbf{x}})^T$, а $\mathbf{q} = (\mathbf{x}_0, \dot{\mathbf{x}}_0)^T$. Между векторами динамического состояния \mathbf{p} (или \mathbf{p}_K) и \mathbf{q} имеет место взаимоднозначное соответствие, поэтому (3.34) можно переписать как

$$\frac{d\mathbf{q}}{dt} = \frac{\partial \mathbf{q}}{\partial \mathbf{p}_K} (\mathcal{P} - \mathcal{P}_K).$$

Поскольку $\mathcal{P} - \mathcal{P}_K = (\mathbf{0}, \mathbf{P})^T$, получаем окончательную систему уравнений в элементах Пайнса

$$\frac{d\mathbf{x}_0}{dt} = \frac{\partial \mathbf{x}_0}{\partial \dot{\mathbf{x}}_K} \mathbf{P}, \quad \frac{d\dot{\mathbf{x}}_0}{dt} = \frac{\partial \dot{\mathbf{x}}_0}{\partial \dot{\mathbf{x}}_K} \mathbf{P}. \quad (3.38)$$

Начальные условия для этой системы очевидны.

Текущий вектор возмущенного динамического состояния \mathbf{p} задается положением и скоростью на кеплеровской орбите с изменяемым начальным динамическим состоянием \mathbf{q} : $\mathbf{p} = \mathbf{p}_K(t, \mathbf{q})$,

определяемым системой уравнений (3.38). Что касается частных производных в системе, то они задаются аналитически и получаются путем прямого дифференцирования формул задачи двух тел (см. Приложение).

Главный недостаток системы (3.38), однако, является то, что с отдалением времени t от начальной эпохи t_0 частные производные возрастают и это существенно может усложнять поведение правых частей уравнений (3.38) даже при малых возмущающих силах \mathbf{P} . Впрочем, разрешается этот недостаток довольно просто: как только величины правых частей уравнений при численном интегрировании становятся больше величины возмущающей силы, скажем, в несколько раз, необходимо переопределять систему (3.38), приняв за начальные параметры t_0 , \mathbf{x}_0 и $\dot{\mathbf{x}}_0$ их текущие аналоги t , \mathbf{x} и $\dot{\mathbf{x}}$.

3.5. Уравнения Энке

Идея метода Энке (Encke, 1852; Херрик, 1977; Рой, 1981) состоит в том, чтобы вместо координат интегрировать их возмущения (отклонения) относительно заранее выбранной (как правило, кеплеровской) орбиты. Применение метода Энке будет эффективным лишь в том случае, если интегрируемые возмущения будут не только меньше, но и более гладкими, чем сама возмущенная орбита, поскольку это условие обеспечивает уменьшение методической ошибки, а потому увеличение шага интегрирования. Свойство гладкости в методе Энке первично, и возмущения обязательно должны быть малыми, тем не менее на практике их малость является обязательным условием, поскольку это позволяет уменьшить влияние ошибок округления на представление вычисляемой орбиты.

Формально метод Энке состоит в следующем. Предположим, что движение небесного тела описывается уравнениями первого порядка

$$\frac{d\mathbf{p}}{dt} = \mathcal{P}(t, \mathbf{p}), \quad \mathbf{p}_0 = \mathbf{p}(t_0). \quad (3.39)$$

В пространстве \mathbf{p} выбирается опорная орбита $\mathbf{p}_K = \mathbf{p}_K(t, \mathbf{q})$, удовлетворяющая уравнениям

$$\frac{d\mathbf{p}_K}{dt} = \mathcal{P}_K(t, \mathbf{p}_K), \quad \mathbf{p}_{K0} = \mathbf{p}_K(t_0). \quad (3.40)$$

Прямым вычитанием уравнений (3.40) из (3.39) получаем уравнения для отклонений $\delta\mathbf{p} = \mathbf{p} - \mathbf{p}_K$:

$$\frac{d\delta\mathbf{p}}{dt} = \mathcal{P}(t, \mathbf{p}_K + \delta\mathbf{p}) - \mathcal{P}_K(t, \mathbf{p}_K), \quad \delta\mathbf{p}_0 = \mathbf{p}_0 - \mathbf{p}_{K0}. \quad (3.41)$$

Если удастся в \mathcal{P} явно выделить невозмущенную \mathcal{P}_K и возмущенную $\mathcal{P} - \mathcal{P}_K$ части, путем соответствующих преобразований разрешается проблема вычитания близких величин:

$$\mathcal{P}_K(t, \mathbf{p}_K + \delta\mathbf{p}) - \mathcal{P}_K(t, \mathbf{p}_K). \quad (3.42)$$

Таким образом, интегрирование системы (3.41) дает возмущения $\delta\mathbf{p}$, из которых возмущенные координаты представляются как $\mathbf{p} = \mathbf{p}_K(t, \mathbf{q}) + \delta\mathbf{p}$.

Обычно в качестве начальных данных в системе (3.40) выбирают $\mathbf{p}_{K0} = \mathbf{p}_0$, т.е. $\delta\mathbf{p}_0 = \mathbf{0}$, поэтому вблизи эпохи оскуляции возмущения $\delta\mathbf{p}$ малы. Тем не менее со временем они возрастают и на достаточно длинных временных интервалах увеличиваются настолько, что разность (3.42), так называемый член Энке, становится сравнимой с возмущающими членами. В этом случае прибегают к процедуре исправления орбиты путем перевычисления параметров опорного движения на новую эпоху.

Причиной возрастания $\delta\mathbf{p}$, очевидно, является неточное представление движения промежуточной орбитой, которое, кроме того, усиливается неустойчивостью движения. Представление можно улучшить путем подбора параметров промежуточной орбиты \mathbf{q} (каковыми также являются начальные данные \mathbf{p}_{K0}) таким образом, чтобы на длительном интервале времени она учитывала основные тенденции динамической эволюции. Однако трудность указанного подхода заключается в том, что уже сама необходимость применения численных методов как альтернативы аналитическим методам для решения динамической задачи предполагает

сложность исследуемой орбитальной динамики и при этом мало что известно заранее о характере движения. Второй подход более фундаментальный и состоит в поиске новых математических моделей для конструирования промежуточных орбит.

3.5.1. Классические уравнения Энке

В классическом методе Энке (Encke, 1852) в качестве опорной выбирается кеплеровская оскулирующая (на начальный момент времени) орбита $\mathbf{x}_K = \mathbf{x}_K(t)$ (см. Приложение), описываемая дифференциальными уравнениями

$$\frac{d^2 \mathbf{x}_K}{dt^2} = \mathbf{F}(\mathbf{x}_K). \quad (3.43)$$

Тогда дифференциальные уравнения для возмущений $\delta \mathbf{x} = \mathbf{x} - \mathbf{x}_K$ принимают вид

$$\frac{d^2 \delta \mathbf{x}}{dt^2} = \mathbf{F}(\mathbf{x}_K + \delta \mathbf{x}) - \mathbf{F}(\mathbf{x}_K) + \mathbf{P}(t, \mathbf{x}_K + \delta \mathbf{x}, \dot{\mathbf{x}}_K + \delta \dot{\mathbf{x}}). \quad (3.44)$$

При малых $\delta \mathbf{x}$ кеплеровские члены в (3.44) — близкие величины. В то же время их малая разность будет значительно меньше каждого из них и, следовательно, из-за ошибок округления она будет вычисляться с неудовлетворительной точностью. Для разрешения этой проблемы проводят дополнительные преобразования, которые приводят разность кеплеровских членов к некой функции Φ , величина которой явно выражается через малые возмущения $\delta \mathbf{x}$, и уравнения приобретают окончательный вид

$$\frac{d^2 \delta \mathbf{x}}{dt^2} = \Phi(\delta \mathbf{x}) + \mathbf{P}(t, \mathbf{x}_K + \delta \mathbf{x}, \dot{\mathbf{x}}_K + \delta \dot{\mathbf{x}}). \quad (3.45)$$

В (3.45) функцию Φ можно привести, например, к следующему виду (Conte, 1962):

$$\Phi(\delta \mathbf{x}) = -\frac{\mu}{|\mathbf{x}|^3} \mathbf{x} + \frac{\mu}{|\mathbf{x}_K|^3} \mathbf{x}_K = \frac{\mu}{|\mathbf{x}_K|^3} (\mathcal{D}\mathbf{x} - \delta \mathbf{x}),$$

$$D = \left(1 + \frac{\zeta^2}{1 + \zeta}\right) \frac{(\mathbf{x}_K + \mathbf{x}) \cdot \delta \mathbf{x}}{|\mathbf{x}|^2}, \quad \text{где } \zeta = \frac{|\mathbf{x}_K|}{|\mathbf{x}|}.$$

Уравнения (3.45) называются уравнениями Энке, несмотря на то что впервые их получил Дж.П. Бонд (Bond, 1849) еще двумя годами раньше Энке. Впрочем, в литературе иногда метод получения уравнений в возмущениях называют методом Бонда–Энке.

Главный недостаток системы уравнений Энке (3.45), как, собственно, и системы уравнений Пайнса (3.38), состоит в ее сложности в сравнении с (3.1) из-за наличия в ней опорного решения $\mathbf{x}_K = \mathbf{x}_K(t)$. Кроме того, в случаях сильновытянутых орбит вычисление опорных положений усложняется необходимостью численного решения уравнения Кеплера¹⁴.

Любой механик в (3.44) без труда узнает уравнения движения, записанные относительно неинерциальной системы координат с близким к исследуемому небесному телу началом, движущимся по закону $\mathbf{x}_K = \mathbf{x}_K(t)$. Тогда величина $-\mathbf{F}(\mathbf{x}_K)$ становится не чем иным, как инерциальной силой, возникающей вследствие неинерциальности координатной системы. Выбранная система отсчета в принципе может быть инерциальной, в то время как условие оскуляции не является обязательным. Продолжая рассуждать в том же направлении, можно прийти к интересному выводу, что в динамическом смысле классические уравнения в прямоугольных координатах (3.1) формально оказываются уравнениями Энке, где $\mathbf{x}_K = \mathbf{0}$, $\mathbf{F}(\mathbf{x}_K) = \mathbf{0}$, а $\delta \mathbf{x} = \mathbf{x}$.

На рис. 3.9 показано, насколько можно повысить точность моделирования возмущенной орбиты, если вместо координат интегрировать малые возмущения. Моделировалась почти круговая орбита единичного радиуса в поле тяготения центрального тела массы $\mu = 1$ под действием притяжения третьего тела массы $\mu_P = 10^{-3}$, двигающегося по круговой орбите радиуса $a_P = 2.08$. Параметры задачи были подобраны так, чтобы они приближенно соответствовали орбитальной динамике астероида главного поя-

¹⁴Хотя этот недостаток не столь существен, поскольку в настоящее время имеются быстрые методы решения уравнения Кеплера (Fukushima, 1996).

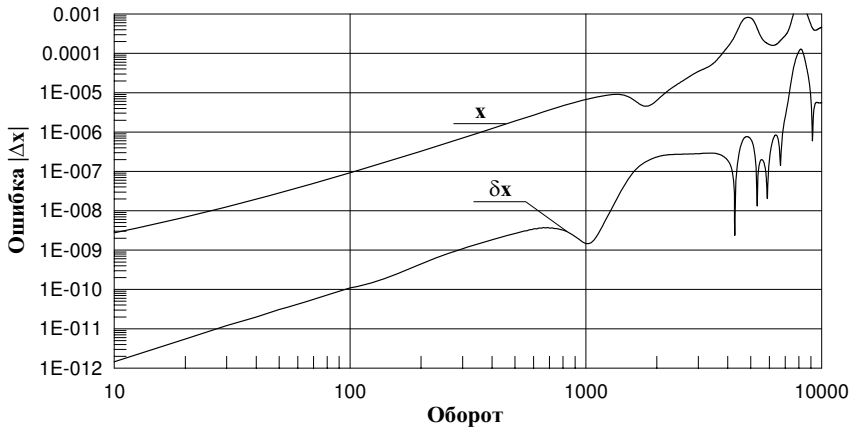


Рис. 3.9. Поведение ошибок в векторе положения при численном интегрировании слабозмущенной почти круговой орбиты классическим методом Рунге–Кутты с шагом $\Delta t = 2\pi/1024$ (\mathbf{x} — уравнения в прямоугольных координатах; $\delta\mathbf{x}$ — уравнения Энке)

са в рамках тройной системы астероид–Солнце–Юпитер. Опорные положения \mathbf{x}_K вычислялись по формулам (П.13), тогда как опорная орбита в уравнениях исправлялась на каждом шаге интегрирования. Как видно из рисунка, метод Энке позволяет существенно повысить точность численного интегрирования. Однако его эффективность высока только до 1000 оборотов объекта. Далее она падает, по-видимому, из-за значительного влияния постепенно накапливающихся вычислительных ошибок в параметрах опорных орбит \mathbf{q} , каковыми в данном случае являются \mathbf{x}_{K0} и $\dot{\mathbf{x}}_{K0}$, перевычисляемые на каждом шаге интегрирования.

3.5.2. Уравнения Энке в KS-переменных

Метод Энке в KS-интерпретации имеет несколько преимуществ. Во-первых, благодаря линейности исходных уравнений движения в KS-переменных преобразование Энке не требует сложных манипуляций для устранения вычитаний близких величин. Во-вторых,

в результате преобразования формульный вид дифференциальных уравнений принципиально не усложняется. В частности, уравнения в отклонениях, описывающие орбитальное движение в KS-пространстве, сохраняют вид возмущенного гармонического осциллятора. В-третьих, опорное решение выражается явно через независимую переменную, фиктивное время s , либо обобщенную эксцентрическую аномалию E (Stiefel, Scheifele, 1971). Заметим, что те же преимущества будут иметь и уравнения Энке в SB-переменных (Авдюшев, Бордовицына, 2006).

Если в качестве опорного взять невозмущенное решение (3.10)

$$\mathbf{u}_K = \mathbf{u}_0 \cos \omega s + \frac{\mathbf{u}'_0}{\omega} \sin \omega s,$$

$$\omega = \sqrt{-\frac{h_K}{2}}, \quad h_K = h_0, \quad \tau_K = -\frac{\mu}{2h_K}s - \frac{1}{h_K}(\mathbf{u}_0 \cdot \mathbf{u}'_0) + t_0,$$

уравнение Энке в KS-переменных будут иметь вид (Бордовицына и др., 1998b; Avdyushev, Bordovitsyna, 2000; Bordovitsyna et al., 2001)

$$\delta \mathbf{u}'' - \frac{h_K}{2} \delta \mathbf{u} = \frac{\delta h}{2} \mathbf{u} + \frac{|\mathbf{u}|^2}{2} \mathbf{L}^T \mathbf{P}, \quad \delta h' = 2(\mathbf{u}' \cdot \mathbf{L}^T \mathbf{P}), \quad (3.46)$$

$$\delta \tau' = -\frac{1}{2h} \left[-\frac{\delta h}{h_K} \mu + |\mathbf{u}|^2 (\mathbf{u} \cdot \mathbf{L}^T \mathbf{P}) - 2(\mathbf{u} \cdot \mathbf{u}') \frac{h'}{h} \right],$$

где возмущенные величины определяются как сумма их опорных аналогов и соответствующих возмущений.

3.5.3. Улучшение опорной орбиты

В классическом методе Энке в качестве опорной используется орбита невозмущенной задачи двух тел. В последнее время предпринимались попытки усовершенствовать метод Энке путем использования лучшей опорной орбиты.

Перспективным для построения таких опорных орбит оказался подход, основанный на идее введения фиктивного притягивающего центра (Shaikh, 1966). Этот подход получил развитие в

работах Ю.В. Батракова (Батраков, Макарова, 1979; Батраков, 1981). Построенные автором промежуточные траектории в методе Энке задают движение по кеплеровской орбите относительно фиктивного притягивающего центра.

Развивая идеи Батракова, В.А. Шефер (Шефер, 1998) предложил в качестве опорного решения в методе Энке рассматривать движение по промежуточной траектории относительно фиктивного центра, которое при этом не является кеплеровским, а описывается уравнениями задачи Гюльдена–Меццерского.

У.Т. Кайнер и М.М. Беннет (Kuyner, Bennett, 1966) показали, что при интегрировании уравнений движения низкого спутника Земли метод Энке можно улучшить, если при построении опорной траектории учесть эффект первого порядка от сжатия Земли.

В основе построения любых промежуточных решений для метода Энке лежат геометрические и динамические свойства орбит, которые могут быть известны заранее. Именно это обеспечивает им практическую ценность в задачах численного исследования орбитального движения. Однако, к сожалению, чаще всего до интегрирования нам мало что известно об исследуемом движении не только в количественном, но и в качественном отношении.

Весьма полезными для построения промежуточных решений могут являться некоторые задачи, имеющие аналитическое решение в частных случаях. Рассмотрим данный подход на примере близких спутников планет. Одним из примечательных свойств в динамике близких спутников планет является то, что они движутся по почти круговым экваториальным орбитам. Как известно, в своем движении спутники испытывают действия двух основных факторов: силы притяжения массы центральной планеты и гравитационное влияние ее экваториального утолщения за счет сжатия. Причем для близких спутников фактор сжатия вносит настолько существенный вклад в орбитальное движение, что использование решения задачи двух тел для предсказания положений спутников уже становится совершенно неприемлемым даже на малых интервалах времени.

Движение близкого спутника в поле тяготения сжатой планеты можно с высокой точностью интерпретировать общей задачей двух неподвижных центров (Аксенов, 1977), которая имеет аналитическое решение в замкнутых формулах. В общем случае для обеспечения высокой точности и малости возмущений именно это решение и целесообразно принять в качестве промежуточного (Сорокин, 1991). Однако, поскольку орбиты рассматриваемых спутников довольно просты и регулярны, использование такой «тяжеловесной» теории для построения опорных решений было бы весьма нерациональным.

С. Херрик показал (Херрик, 1978), что увеличение центральной силы или гравитационного параметра в опорной орбите позволяет уменьшить возмущения. В нашем случае фактор сжатия в динамике близких спутников можно также трактовать как эффект дополнительной массы центральной планеты. Поэтому для построения опорных решений здесь может быть достаточно использовать формулы задачи двух тел с модифицированным гравитационным параметром. Изложенную идею легко продемонстрировать на примере уравнений в прямоугольных координатах. Для этого предварительно рассмотрим одну простую задачу.

Пусть спутник движется в экваториальной плоскости планеты по круговой орбите. Тогда $x_3 = 0$ и потенциал несферичности планеты до второй зональной гармоники (см. разд. 2.1) будет

$$V = V(\mathbf{x}) = -\frac{\mathcal{J}}{2|\mathbf{x}|^3},$$

где $\mathcal{J} = \mu C_2^0 b^2$ (b — экваториальный радиус планеты), а дифференциальные уравнения движения (3.1) можно переписать в виде

$$\frac{d^2 \mathbf{x}}{dt^2} + \frac{\mu}{|\mathbf{x}|^3} \mathbf{x} = \mathbf{P} = \left(\frac{\partial V}{\partial \mathbf{x}} \right)^T = \frac{3}{2} \frac{\mathcal{J}}{|\mathbf{x}|^5} \mathbf{x}. \quad (3.47)$$

Группируя коэффициенты при кеплеровском множителе, получим (Херрик, 1978)

$$\frac{d^2 \mathbf{x}}{dt^2} + \frac{\mu}{|\mathbf{x}|^3} \left(1 - \frac{3}{2} \frac{C_2^0 b^2}{|\mathbf{x}|^2} \right) \mathbf{x} = \mathbf{0}. \quad (3.48)$$

Таким образом, круговое движение экваториального спутника в поле тяготения сжатой планеты описывается уравнениями задачи двух тел с измененным (увеличенным) гравитационным параметром. Очевидно, что решение полученных уравнений в сравнении с кеплеровским может достаточно хорошо представлять движение экваториальных спутников с почти круговыми орбитами (в особенности класса близких спутников, когда сжатие центральной планеты существенно изменяет частоту спутникового движения по орбите). Следовательно, в алгоритмах типа Энке при исследовании движения данных объектов именно это решение целесообразно использовать в качестве опорного вместо кеплеровского.

3.6. Гравитационные системы координат и уравнения синхронного слежения

Метод Энке является эффективным средством для уменьшения ошибок округления. Динамический смысл уравнений Энке состоит в том, что они описывают динамику материальной точки относительно некоторой подвижной системы координат с близким к точке началом. Это обеспечивает малость интегрируемых переменных, а потому является залогом малых содержащихся в них ошибок округления.

Основной проблемой в численном исследовании орбитального движения вблизи массивных тел является существенная потеря вычислительной точности. Особенно неприятно при численном интегрировании такое явление, как сближение астероида с большой планетой. Удивительно, что даже при оптимальном шаге интегрирования во время сближения с планетой (гелиоцентрическая) орбита астероида обрастает ошибками округления более интенсивно, нежели даже после многократных сближений с центральным массивным телом, каким является Солнце¹⁵ (Авдюшев, 2003b).

¹⁵Если уравнивать величины притягивающих сил от Солнца и Юпитера, то получим, что сближению с Юпитером 0.01 а.е. равнозначным будет сближение с Солнцем приблизительно 0.3 а.е., однако с точки зрения численного

Проблема потери точности при сближениях, как уже было сказано, оказывается, состоит просто в неудачной формализации исследуемой орбиты, а именно, когда она рассматривается в рамках только одной координатной системы, начало которой достаточно удалено от массивного тела, с которым происходит сближение (Авдюшев, 2000; Авдюшев, 2003b). В то же время успех в сохранении точности интегрирования обеспечивается только при совместном использовании различных координатных систем, связанных с массивными телами, определяющими главным образом орбитальную динамику.

В данном разделе обосновывается выбор гравитационной системы координат в задаче двух тел, а также комбинированное использование гравитационных систем в задачах численного исследования орбитального движения при наличии тесных сближений с массивными телами. Здесь под гравитационной понимается координатная система, начало которой помещено в гравитационный центр, каким может являться звезда или планета, рассматриваемые как материальные точки.

3.6.1. Ошибки округления в кеплеровских членах.

Задача двух тел

Если спутниковая задача решается численным методом, то выбор планетоцентрической системы координат становится не столь очевидным, поскольку все динамически неподвижные системы координат, в том числе и планетоцентрическая, могут рассматриваться как равнозначные по той причине, что характер поведения правых частей динамических уравнений при переходе от одной системы к другой сохраняется. Тем не менее уравнения, записанные относительно планетоцентрической системы, качественно лучше, поскольку, как оказывается, менее подвержены влиянию ошибок округления. Покажем это на примере задачи двух тел.

интегрирования эти сближения не равнозначны, поскольку вычислительные ошибки для первого будут гораздо больше, чем для второго.

Рассмотрим движение тестовой частицы в поле тяготения точечной массы с вектором положения \mathbf{c} в некоторой неподвижной системе координат и гравитационным параметром μ . Тогда уравнения движения примут вид

$$\frac{d^2\mathbf{x}}{dt^2} = -\mu \frac{\mathbf{x} - \mathbf{c}}{|\mathbf{x} - \mathbf{c}|^3} = \mathbf{F}. \quad (3.49)$$

Предположим, что в векторе \mathbf{x} содержится ошибка $\Delta\mathbf{x}$. Оценим величину ошибки $\Delta\mathbf{F}$ в кеплеровском члене \mathbf{F} , производимую ошибкой $\Delta\mathbf{x}$. С точностью до малых первого порядка будем иметь

$$\Delta\mathbf{F} = -\mu \frac{\Delta\mathbf{x}}{|\mathbf{x} - \mathbf{c}|^3} + 3\mu \frac{\mathbf{x} - \mathbf{c}}{|\mathbf{x} - \mathbf{c}|^5} (\mathbf{x} - \mathbf{c}) \cdot \Delta\mathbf{x}. \quad (3.50)$$

Теперь представим векторы $\Delta\mathbf{x}$ и $\mathbf{x} - \mathbf{c}$ как

$$\Delta\mathbf{x} = |\Delta\mathbf{x}|\mathbf{y} \quad \text{и} \quad \mathbf{x} - \mathbf{c} = |\mathbf{x} - \mathbf{c}|\mathbf{z}, \quad (3.51)$$

где \mathbf{y} и \mathbf{z} — единичные векторы. Пусть φ — угол между векторами \mathbf{y} и \mathbf{z} . Тогда, используя (3.50), получим оценку для величины $\Delta\mathbf{F}$:

$$|\Delta\mathbf{F}| = \mu \frac{|\Delta\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3} |\mathbf{y} - 3\mathbf{z} \cos \varphi|. \quad (3.52)$$

Величину $|\mathbf{y} - 3\mathbf{z} \cos \varphi|$ можно легко определить, если ввести 2-мерную систему координат в плоскости (\mathbf{y}, \mathbf{z}) таким образом, чтобы $\mathbf{z} = (1, 0)$. В этой системе $\mathbf{y} = (\cos \varphi, \sin \varphi)$ и, следовательно,

$$|\mathbf{y} - 3\mathbf{z} \cos \varphi| = \sqrt{1 + 3 \cos^2 \varphi}.$$

Тогда, согласно (3.52), справедливо неравенство

$$\mu \frac{|\Delta\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3} \leq |\Delta\mathbf{F}| \leq 2\mu \frac{|\Delta\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3}. \quad (3.53)$$

Если $\Delta\mathbf{x}$ — вектор ошибок округления, направление которого изотропно, то $|\Delta\mathbf{x}| \approx \varepsilon|\mathbf{x}|$, где ε — машинный эпсилон, характеризующий точность компьютерной арифметики; и среднее значение (математическое ожидание) $|\Delta\mathbf{F}|$ будет

$$\overline{|\Delta\mathbf{F}|} \approx 1.38\mu\varepsilon \frac{|\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3}.$$

В дальнейшем для простоты изложения будем использовать нижний предел неравенства (3.53), т.е. будем считать, что

$$|\Delta \mathbf{F}| = \mu \varepsilon \frac{|\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3}. \quad (3.54)$$

Заметим, что ошибка $|\Delta \mathbf{F}|$ также может быть вызвана округлением самой величины \mathbf{F} . В этом случае

$$|\Delta \mathbf{F}| = \varepsilon |\mathbf{F}| = \frac{\mu \varepsilon}{|\mathbf{x} - \mathbf{c}|^2}. \quad (3.55)$$

Поэтому для оценки $|\Delta \mathbf{F}|$ следует выбирать наибольшее из величин (3.54) и (3.55):

$$|\Delta \mathbf{F}| = \max \left(\frac{\mu \varepsilon}{|\mathbf{x} - \mathbf{c}|^2}, \frac{\mu \varepsilon |\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|^3} \right) = \frac{\mu \varepsilon}{|\mathbf{x} - \mathbf{c}|^2} \max \left(1, \frac{|\mathbf{x}|}{|\mathbf{x} - \mathbf{c}|} \right). \quad (3.56)$$

Оценка (3.56), впрочем, как и (3.54), говорит о том, что с удалением системы координат от тестовой частицы при фиксированном гравитационном векторе положения $\mathbf{x} - \mathbf{c}$ ошибка $|\Delta \mathbf{F}|$ будет неограниченно возрастать. Таким образом, только неудачный выбор начала отсчета может привести к нежелательным возмущающим силам $\Delta \mathbf{F}$, не имеющим никакого отношения к физике задачи и способным существенно исказить истинную динамику тестовой частицы.

С другой стороны, если зафиксировать начало координатной системы около гравитационного центра ($\mathbf{c} \neq \mathbf{0}$), то при $|\mathbf{x} - \mathbf{c}| \rightarrow 0$ в соответствии с (3.54) $|\Delta \mathbf{F}| \sim |\mathbf{x} - \mathbf{c}|^{-3}$, но в то же самое время, если гравитационный центр взять за начало координат ($\mathbf{c} = \mathbf{0}$), фактически получим, что $|\Delta \mathbf{F}| \sim |\mathbf{x} - \mathbf{0}|^{-2}$. Отсюда следует, что в гравитационной системе при достаточно малых $|\mathbf{x} - \mathbf{c}|$ ошибка $|\Delta \mathbf{F}|$ будет существенно меньше. Это и объясняет, чем хорош гравитационный центр в качестве начала отсчета для численного интегрирования близких к нему орбит.

На рис. 3.10 показано, как ухудшается точность интегрирования спутниковой орбиты при удалении системы координат от

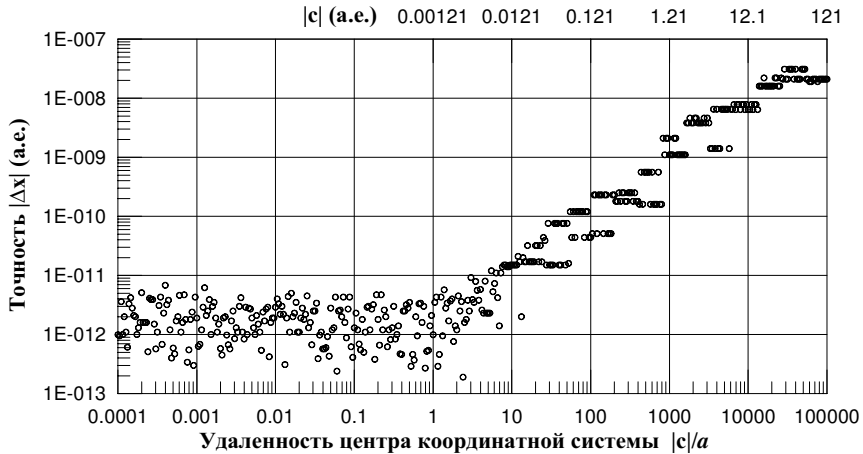


Рис. 3.10. Точность интегрирования орбиты Амальтеи, пятого спутника Юпитера ($a = 0.00121$ а.е., $T = 0.5$ сут), в зависимости от удаленности $|c|$ начала координатной системы от центральной планеты

центральной планеты. Орбита интегрировалась методом Эверхарта 15-го порядка в арифметике с двойной точностью ($\varepsilon \approx 10^{-16}$). Ошибка вычислений $|\Delta \mathbf{x}|$ оценивалась методом прямого и обратного интегрирования на интервале времени 10 лет. Интегрирование выполнялось с постоянным шагом $5 \cdot 10^{-3}$ сут, который в результате уравнивания ошибок усечения и округления обеспечивал наивысшую точность в йовицентрической системе координат.

До тех пор, пока $|c| < a$ (начало координатной системы находится внутри орбиты), точность интегрирования $|\Delta \mathbf{x}|$ остается в пределах одного порядка. В этом случае все координатные системы равнозначны. Однако в дальнейшем ($|c| > a$) с увеличением $|c|$ ошибка $|\Delta \mathbf{x}|$ возрастает пропорционально, что хорошо согласуется с оценкой (3.54). Здесь необходимо учесть, что планетоцентрическое расстояние $|\mathbf{x} - \mathbf{c}|$ почти постоянно.

Таким образом, если бы мы интегрировали спутниковую орбиту в гелиоцентрической системе координат, то заплатились бы за это ухудшением точности интегрирования почти на 4 порядка.

В астероидной задаче, где исследуемый объект сближается с планетой и временно становится спутником, выбор планетоцентрической координатной системы как будто очевиден. Но встает вопрос: в какой момент астероид меняет свое амплуа и когда необходимо прибегать к планетоцентрической формализации его орбиты, чтобы тем самым уменьшить влияние злополучных ошибок округления? Об этом речь пойдет ниже.

3.6.2. Ограниченная задача трех тел. Астероидная задача

Будем рассматривать движение астероида в гравитационном поле Солнца и (для определенности) Юпитера. Запишем уравнения движения относительно гелиоцентра

$$\frac{d^2 \mathbf{x}_S}{dt^2} = - \underbrace{\mu_S \frac{\mathbf{x}_S - \mathbf{0}}{|\mathbf{x}_S - \mathbf{0}|^3}}_{\mathbf{F}_{SS}} - \underbrace{\mu_J \frac{\mathbf{x}_S - \mathbf{c}_{JS}}{|\mathbf{x}_S - \mathbf{c}_{JS}|^3}}_{\mathbf{F}_{JS}} - \underbrace{\mu_J \frac{\mathbf{c}_{JS} - \mathbf{0}}{|\mathbf{c}_{JS} - \mathbf{0}|^3}}_{\mathbf{F}_{IS}} \quad (3.57)$$

и йовицентра

$$\frac{d^2 \mathbf{x}_J}{dt^2} = - \underbrace{\mu_J \frac{\mathbf{x}_J - \mathbf{0}}{|\mathbf{x}_J - \mathbf{0}|^3}}_{\mathbf{F}_{JJ}} - \underbrace{\mu_S \frac{\mathbf{x}_J - \mathbf{c}_{SJ}}{|\mathbf{x}_J - \mathbf{c}_{SJ}|^3}}_{\mathbf{F}_{SJ}} - \underbrace{\mu_S \frac{\mathbf{c}_{SJ} - \mathbf{0}}{|\mathbf{c}_{SJ} - \mathbf{0}|^3}}_{\mathbf{F}_{IJ}}. \quad (3.58)$$

Здесь μ_S и μ_J — гравитационные параметры Солнца и Юпитера соответственно; \mathbf{x}_S и \mathbf{c}_{JS} — гелиоцентрические векторы положения астероида и Юпитера, а \mathbf{x}_J и \mathbf{c}_{SJ} — йовицентрические векторы положения астероида и Солнца. Причем заметим, что

$$\mathbf{x}_S - \mathbf{c}_{JS} = \mathbf{x}_J, \quad \mathbf{x}_J - \mathbf{c}_{SJ} = \mathbf{x}_S \quad \text{и} \quad \mathbf{c}_{JS} = -\mathbf{c}_{SJ}. \quad (3.59)$$

При сближении с Юпитером $|\mathbf{x}_J| < |\mathbf{x}_S|$. Тогда в соответствии с (3.56) и (3.59) для ошибок вычисления сил в правых частях уравнений (3.57) и (3.58) будут справедливы следующие оценки:

$$|\Delta \mathbf{F}_{SS}| = \frac{\mu_S \varepsilon}{|\mathbf{x}_S|^2}, \quad |\Delta \mathbf{F}_{JS}| = \frac{\mu_J \varepsilon}{|\mathbf{x}_J|^2} \frac{|\mathbf{x}_S|}{|\mathbf{x}_J|}, \quad |\Delta \mathbf{F}_{IS}| = \frac{\mu_J \varepsilon}{|\mathbf{c}_{JS}|^2}; \quad (3.60)$$

$$|\Delta \mathbf{F}_{JJ}| = \frac{\mu_J \varepsilon}{|\mathbf{x}_J|^2}, \quad |\Delta \mathbf{F}_{SJ}| = \frac{\mu_S \varepsilon}{|\mathbf{x}_S|^2}, \quad |\Delta \mathbf{F}_{IJ}| = \frac{\mu_S \varepsilon}{|\mathbf{c}_{SJ}|^2}. \quad (3.61)$$

Поскольку $\mu_J < \mu_S$, а $|\mathbf{c}_{JS}| \approx |\mathbf{x}_S|$ ввиду малости $|\mathbf{x}_J|$, то в (3.60) ошибка $|\Delta \mathbf{F}_{IS}|$ будет меньше двух других, а $|\Delta \mathbf{F}_{IJ}| \approx \approx |\Delta \mathbf{F}_{SJ}|$. Поэтому $|\Delta \mathbf{F}_{IS}|$ и $|\Delta \mathbf{F}_{IJ}|$ можно исключить из рассмотрения и сосредоточить внимание на оставшихся оценках. Тогда, согласно (3.60), условием использования уравнений (3.57), очевидно, будет неравенство $|\Delta \mathbf{F}_{SS}| > |\Delta \mathbf{F}_{JS}|$, а условием перехода к уравнениям (3.58) — равенство $|\Delta \mathbf{F}_{SS}| = |\Delta \mathbf{F}_{JS}|$, которое дает

$$|\mathbf{x}_J| = |\mathbf{x}_S| \sqrt[3]{\frac{\mu_J}{\mu_S}}. \quad (3.62)$$

Как видно, поверхность, задаваемая уравнением (3.62), близка к сфере влияния Юпитера ($|\mathbf{x}_J| = \text{const} \approx 0.5$ а.е.) (Кислик, 1964).

Сравнивая между собой $|\Delta \mathbf{F}_{JJ}|$, $|\Delta \mathbf{F}_{SJ}|$ и $|\Delta \mathbf{F}_{JS}|$, получаем, что при

$$|\mathbf{x}_J| < |\mathbf{x}_S| \sqrt{\frac{\mu_J}{\mu_S}}$$

переход к йовицентрической системе координат способен фактически повысить точность вычисления правой части уравнений в $|\mathbf{x}_S|/|\mathbf{x}_J|$ раз.

На рис. 3.11 (серый цвет) показано, насколько эффективным может быть использование планетоцентрической координатной системы в задачах исследования орбитальной динамики астероидов, сближающихся с планетами. Здесь и ниже результаты получены в рамках ограниченной плоской круговой задачи трех тел: Солнце–Юпитер–астероид (рис. 3.12). На интервале времени 1000 сут рассматривалось семейство параболических¹⁶ относительно планеты астероидных орбит с перицентрическими расстояниями $\min |\mathbf{x}_J| = 1 - 0.001$ а.е. и моментом прохождения перигелия τ . Орбиты интегрировались численно методом Эверхарта в арифметике с двойной точностью. Интегрирование выполнялось

¹⁶В эпоху оскуляции τ .

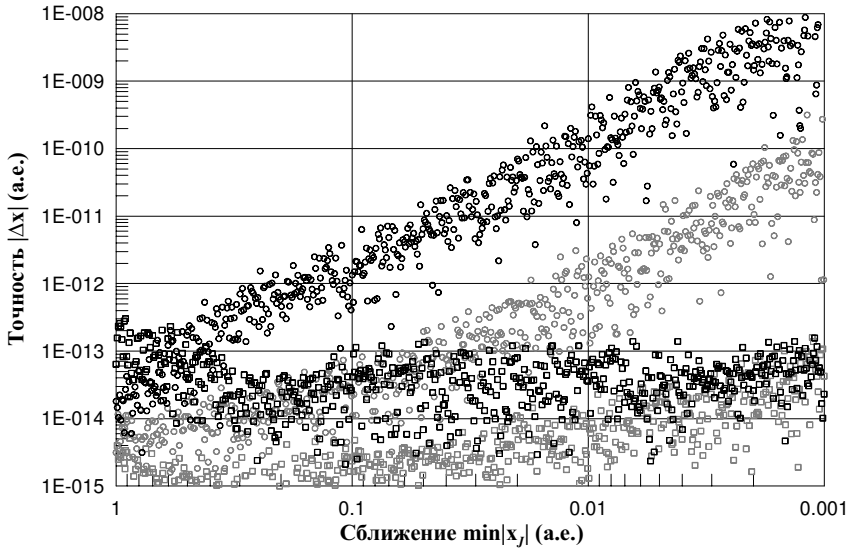


Рис. 3.11. Ошибки численного интегрирования астероидных орбит в разных системах координат в зависимости от минимального расстояния до планеты $\min |\mathbf{x}_J|$: \circ — в гелиоцентрической системе ($|\Delta \mathbf{x}_S|$); \square — в планетоцентрической ($|\Delta \mathbf{x}_J|$). Серым цветом показаны результаты для $\tau = 0$, черным — для $\tau = 2451544.5$ JD

с переменным шагом, обеспечивающим наивысшую точность, т.е. когда влияния ошибок усечения и округления на результаты интегрирования уравниваются. Вычислительная точность оценивалась методом прямого и обратного интегрирования. Таким образом, в процессе моделирования каждой орбиты имели место два однотипных сближения (одно — при прямом интегрировании, другое — при обратном).

Итак, из рис. 3.11 видно, что при тесном сближении планетоцентрическая система позволяет сохранить точность численного интегрирования на 2 порядка и выше ($\min |\mathbf{x}_J| < 0.5$). При сближениях вне сферы влияния ($\min |\mathbf{x}_J| > 0.5$) ошибки $|\Delta \mathbf{x}_J|$ стано-

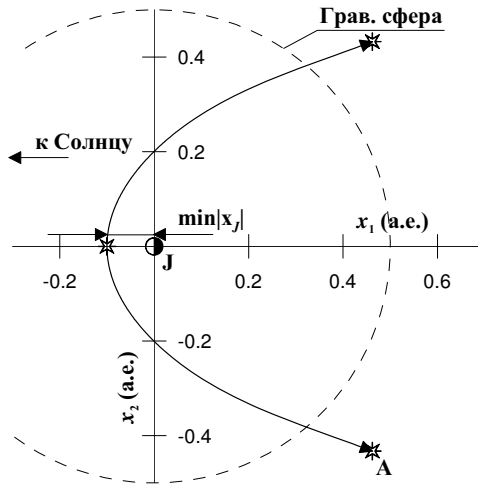


Рис. 3.12. Сближение астероида с Юпитером

вдаться сравнимы с ошибками $|\Delta \mathbf{x}_S|$, причем имеет место тенденция к ухудшению точности интегрирования в планетоцентрической системе координат с увеличением $\min |\mathbf{x}_J|$. Так, переход к планетоцентрической координатной системе становится совершенно бессмысленным, если астероид не приближается к Юпитеру ближе чем на 1 а.е.

3.6.3. Временные ошибки

Численный эксперимент, изложенный выше, был выполнен при условии, что момент прохождения через перигеум равен нулю. Поскольку векторы положения возмущающих тел \mathbf{c}_{JS} и \mathbf{c}_{SJ} в эксперименте вычислялись по аналитическим формулам (кругового движения) как функции времени, то условие $\tau = 0$ обеспечивало им минимальные ошибки, вызываемые ошибкой округления Δt во времени. Их можно оценить как

$$\Delta \mathbf{c}_{JS} = \dot{\mathbf{c}}_{JS} \Delta t \quad \text{и} \quad \Delta \mathbf{c}_{SJ} = \dot{\mathbf{c}}_{SJ} \Delta t, \quad \text{где} \quad \Delta t \approx \varepsilon t. \quad (3.63)$$

Однако в практических задачах при использовании более точных аналитических представлений движения возмущающих тел как в DE405 (Standish, 1998) векторы типа \mathbf{c}_{JS} и \mathbf{c}_{SJ} вычисляются на полную юлианскую дату, и поэтому ошибки в них, вызванные временной ошибкой, здесь гораздо больше. Более того, они будут превышать ошибки округления $\Delta\mathbf{x}_S$ и $\Delta\mathbf{x}_J$. Для Юпитера, например, при $\varepsilon \sim 10^{-16}$ $\Delta\mathbf{c}_{JS} \sim 10^{-11}$ а.е., тогда как при сближении астероида с планетой $\Delta\mathbf{x}_S \sim 10^{-15}$ а.е.

Тем не менее условие перехода (3.62) остается в силе. Действительно, в задачах сближения ошибки $\Delta\mathbf{c}_{JS}$ и $\Delta\mathbf{c}_{SJ}$ наиболее заметно должны проявлять себя в членах \mathbf{F}_{JS} и \mathbf{F}_{SJ} . Поэтому, уравнивая $|\Delta\mathbf{F}_{JS}|$ и $|\Delta\mathbf{F}_{SJ}|$ с учетом только ошибок $\Delta\mathbf{c}_{JS}$ и $\Delta\mathbf{c}_{SJ}$, получим

$$\mu_J \frac{|\Delta\mathbf{c}_{JS}|}{|\mathbf{x}_J|^3} = \mu_S \frac{|\Delta\mathbf{c}_{SJ}|}{|\mathbf{x}_S|^3},$$

что фактически дает (3.62).

Мы повторили предыдущий эксперимент, где в качестве момента прохождения через перигеум взяли полную юлианскую дату, для определенности 2000 г.: $\tau = 2451544.5$ JD. Результаты приведены на рис. 3.11 (черный цвет).

Как видно, точность интегрирования в обеих системах координат становится хуже: в гелиоцентрической системе до двух порядков при $\min |\mathbf{x}_J| < 0.1$ а.е., тогда как в йовицентрической системе до одного порядка при $\min |\mathbf{x}_J| > 0.01$ а.е. Таким образом, результаты эксперимента показывают, насколько нежелательны в уравнениях, описывающих орбиты сближения, члены, явно зависящие от времени.

Кроме того, следует заметить, что большие временные ошибки приводят к скачкообразному движению возмущающих тел и к разрывам в правых частях уравнений движения, которые интегратором воспринимаются как сверхбыстрые изменения. Предположим, что мы интегрируем орбиту астероида, сближающегося с планетой, в гелиоцентрической системе координат. До сближения изменения в правых частях, вызванные временной ошибкой, не представляют опасности для численного интегрирования. Однако

при вхождении астероида в сферу влияния временные ошибки будут усиливаться сингулярностью $|\mathbf{x}_S - \mathbf{c}_{JS}|^{-3}$, причем настолько, что окажутся способными влиять не только на точность интегрирования, но и на выбор переменного шага.

На рис. 3.13 показано, как соотносятся скорости интегрирования в гелиоцентрической (NS_S) и планетоцентрической (NS_J) системах координат при различных значениях параметра интегратора L (см. подразд. 4.3.7: $\|\mathbf{e}\|_{tol} = 10^{-L}$), определяющего величину переменного шага в интеграторе. Здесь скорость интегрирования оценивалась по количеству выполненных шагов NS за весь интервал интегрирования. Как видно из рисунка, при повышении задаваемой локальной точности интегратор становится более чувствительным к влиянию временных ошибок. Так, при $L = 14$ быстродействие интегрирования орбиты глубоко внутри гравитационной сферы относительно гелиоцентрической системы координат ниже в 4–5 раз, нежели относительно ювицентрической системы; и, наоборот, вне гравитационной сферы эффективность по быстродействию в гелиоцентрической системе выше. В то же время следует иметь в виду, что в данном случае $L = 13$ и 14 — избыточно большие величины, поскольку уже $L = 12$ является оптимальным параметром, при котором уравниваются влияния ошибок усечения и округления, и, кроме того, скорости вычислений в обеих координатных системах становятся почти одинаковыми. Поэтому, как мы видим, только неудачный выбор параметра L может существенно понизить быстродействие интегрирования орбиты астероида в тот период, когда он сближается с планетой. Интересно также заметить, что все характеристики на рис. 3.13 пересекаются приблизительно в значении $\min |\mathbf{x}_J| = 0.6$. Это говорит о том, что на практике радиус гравитационной сферы, определяющей критерий смены координатных систем, можно выбирать немного больше, чем задает формула (3.62).

Необходимо также обратить внимание на одну деталь, связанную с временной ошибкой. Переход к новой системе координат в задачах сближения, как мы выяснили, необходим для повышения точности вычисления правых частей уравнений движения. В ре-

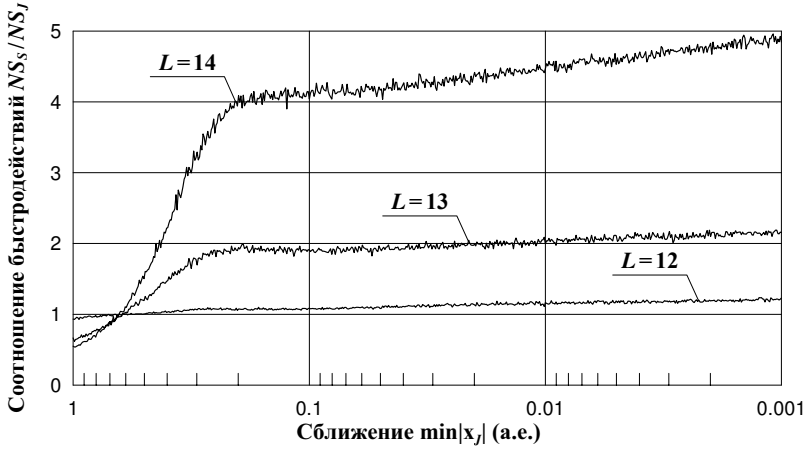


Рис. 3.13. Соотношение быстродействий численного интегрирования в гелио- и йовицентрической координатных системах при различных параметрах L интегратора Эверхарта 15-го порядка ($\tau = 2451544.5$ JD)

в результате смены системы координат только при одном сближении интегрируемая задача Коши разбивается на три подзадачи: до сближения, во время сближения и после сближения с планетой. Каждая из подзадач требует перевычисления начальных данных. Например, для перехода от гелиоцентрической координатной системы к йовицентрической используются формулы

$$\mathbf{x}_{J0} = \mathbf{x}_{S0} - \mathbf{c}_{JS0}, \quad \dot{\mathbf{x}}_{J0} = \dot{\mathbf{x}}_{S0} - \dot{\mathbf{c}}_{JS0}. \quad (3.64)$$

Следовательно, в новые начальные данные через векторы \mathbf{c}_{JS0} и $\dot{\mathbf{c}}_{JS0}$ будут вноситься ошибки типа (3.63). Таким образом, улучшая за счет перехода уравнения подзадачи Коши, мы будем ухудшать ее начальные условия.

Эксперименты, изложенные выше, были выполнены в рамках одной из двух координатных систем. Поэтому характеристики для йовицентрической системы на рис. 3.11 в действительности не да-

ют ответа на вопрос, какова истинная оценка точности интегрирования при использовании переходов от одной системы к другой. К этому вопросу вернемся несколько позже.

3.6.4. Метод синхронного слежения

Временную зависимость членов \mathbf{F}_{JS} и \mathbf{F}_{SJ} в уравнениях (3.57) и (3.58) удастся избежать в результате объединения этих уравнений. Действительно, произведя замену (3.59), получим систему (Авдюшев, 2000; Авдюшев, 2003b)

$$\begin{aligned}\frac{d^2\mathbf{x}_S}{dt^2} &= -\mu_S \frac{\mathbf{x}_S}{|\mathbf{x}_S|^3} - \mu_J \frac{\mathbf{x}_J}{|\mathbf{x}_J|^3} - \mu_J \frac{\mathbf{c}_{JS}}{|\mathbf{c}_{JS}|^3}, \\ \frac{d^2\mathbf{x}_J}{dt^2} &= -\mu_J \frac{\mathbf{x}_J}{|\mathbf{x}_J|^3} - \mu_S \frac{\mathbf{x}_S}{|\mathbf{x}_S|^3} - \mu_S \frac{\mathbf{c}_{SJ}}{|\mathbf{c}_{SJ}|^3}.\end{aligned}\quad (3.65)$$

Она представляет собой формализацию движения, рассматриваемого одновременно в разных координатных системах. Поэтому назовем ее системой синхронного слежения.

Помимо того, что уравнения (3.65) освобождают важные для численного интегрирования члены от временной зависимости, они также хороши еще и тем, что избавляют нас от переходов, какие необходимы при использовании систем (3.57) и (3.58) по отдельности, а это, в свою очередь, позволяет обезопасить численное решение от появления больших ошибок, не связанных с интегрированием.

Единственным недостатком системы (3.65) является ее большой порядок, который в два раза превышает порядок каждой из систем (3.57) либо (3.58). Поэтому уравнения синхронного слежения интегрируются медленнее. Однако этот недостаток может быть компенсирован введением временного (сглаживающего) преобразования типа (3.14):

$$dt = \frac{|\mathbf{x}_S||\mathbf{x}_J|}{\mu_S|\mathbf{x}_J| + \mu_J|\mathbf{x}_S|} ds,$$

где s — новая независимая переменная, так называемое фиктивное время. Как показывают численные результаты, временное преобразование позволяет повысить быстродействие интегрирования в несколько раз (Бордовицына и др., 1998а; Бордовицына и др., 2003).

Следует особо отметить, что для корректного использования системы (3.65) орбиты $\mathbf{x}_S = \mathbf{x}_S(t)$ и $\mathbf{x}_J = \mathbf{x}_J(t)$ должны быть адекватны друг другу, что выражается соотношением

$$\mathbf{c}_{JS} = \mathbf{x}_S - \mathbf{x}_J. \quad (3.66)$$

В соответствии с (3.66) вектор $\mathbf{c}_{JS} = \mathbf{c}_{JS}(t)$ должен являться решением задачи Коши с уравнением

$$\frac{d^2 \mathbf{c}_{JS}}{dt^2} = \frac{d^2 \mathbf{x}_S}{dt^2} - \frac{d^2 \mathbf{x}_J}{dt^2} \quad (3.67)$$

и начальными условиями

$$\mathbf{c}_{JS0} = \mathbf{x}_{S0} - \mathbf{x}_{J0}, \quad \dot{\mathbf{c}}_{JS0} = \dot{\mathbf{x}}_{S0} - \dot{\mathbf{x}}_{J0}. \quad (3.68)$$

Условия (3.68) выполняются автоматически при использовании системы синхронного слежения. Следовательно, адекватность орбит будет иметь место при выполнении равенства (3.67). В нашем случае (3.67) с учетом (3.65) будет иметь вид уравнения кеплеровского движения планеты относительно Солнца

$$\frac{d^2 \mathbf{c}_{JS}}{dt^2} = -(\mu_S + \mu_J) \frac{\mathbf{c}_{JS}}{|\mathbf{c}_{JS}|^3}. \quad (3.69)$$

Поэтому в рамках классической постановки ограниченной задачи трех тел принцип адекватности выполняется. Однако, если для вычисления положений планеты мы бы пользовались эфемеридой, скажем, DE405, то тогда вектор \mathbf{c}_{JS} не удовлетворял бы уравнению (3.69), поскольку на самом деле он является решением более сложной задачи. В этом случае для выполнения принципа адекватности движение астероида следует рассматривать в тех условиях, в которых была получена эфемерида планеты.

Наконец, в системе синхронного слежения (3.65) можно рассмотреть ту неприятную особенность, что правая часть системы ввиду наличия \mathbf{c}_{JS} и \mathbf{c}_{SJ} еще зависит от времени. Впрочем, избежать эту особенность удастся путем простого объединения уравнений (3.65) и (3.69). Интегрируемые численно переменные \mathbf{c}_{JS} и \mathbf{c}_{SJ} будут существенно точнее их аналитических аналогов, вычисляемых на большие значения времени, поэтому и ошибки интегрирования расширенной системы (3.65) и (3.69) будут меньше. Разумеется, тот же способ можно применить к уравнениям (3.57) и (3.58), что тем самым позволяет качественно улучшить их интегрирование. Использование интегрируемых переменных \mathbf{c}_{JS} и \mathbf{c}_{SJ} в (3.64) также повышает точность вычисления начальных данных, необходимых при переходе от одной координатной системы к другой. Очевидно, главным недостатком таких расширенных систем является понижение быстродействия интегрирования.

3.6.5. Эффективность интегрирования при использовании гравитационных координатных систем

Чтобы выявить эффективность численного интегрирования астероидных орбит с использованием различных гравитационных координатных систем, проанализируем результаты следующего численного эксперимента. В рассматриваемой астероидной задаче выполнялось численное моделирование тесных сближений астероида с Юпитером в рамках системы уравнений (3.57), систем (3.57) и (3.58) с переходом от одной к другой во время прохождения астероида через сферу влияния планеты, а также в рамках системы (3.65), используемой на протяжении всего интегрирования. Путем сравнения результатов прямого и обратного интегрирования в эксперименте оценивалась точность моделирования гелиоцентрического вектора положения \mathbf{x}_S . Причем для корректной оценки точности в случае последовательного использования координатных систем переходы при обратном интегрировании от одной системы к другой производились на те же моменты времени, что

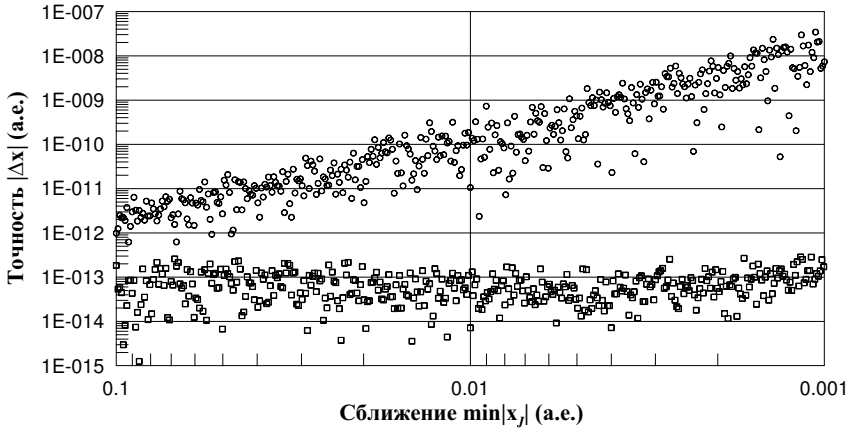


Рис. 3.14. Точность интегрирования астероидных орбит в зависимости от сближения: ○ — относительно гелиоцентра; □ — с переходом от гелиоцентрической системы координат к йовицентрической

и при прямом интегрировании. Заметим, что таким образом исключалось влияние ошибок типа (3.63), вносимых в начальные данные каждой из подзадач Коши с уравнениями (3.57) и (3.58).

Результаты эксперимента приведены на рис. 3.14. Как видно из рисунка, временное использование йовицентрической координатной системы в период пребывания астероида внутри гравитационной сферы Юпитера позволяет повысить точность интегрирования на несколько порядков. Так, при наиболее тесных сближениях $\min |x_J| \approx 0.001$ а.е. переход к планетоцентрической системе дает улучшение точности на 5 порядков и выше.

Между тем заметим, что точность численного интегрирования с последовательным использованием координатных систем при различных сближениях остается приблизительно на одном уровне. Это, по-видимому, обусловлено тем, что в интегрируемых уравнениях (3.58) имеются члены, явно зависящие от времени и вместе с тем слабо зависящие от малых величин x_J .

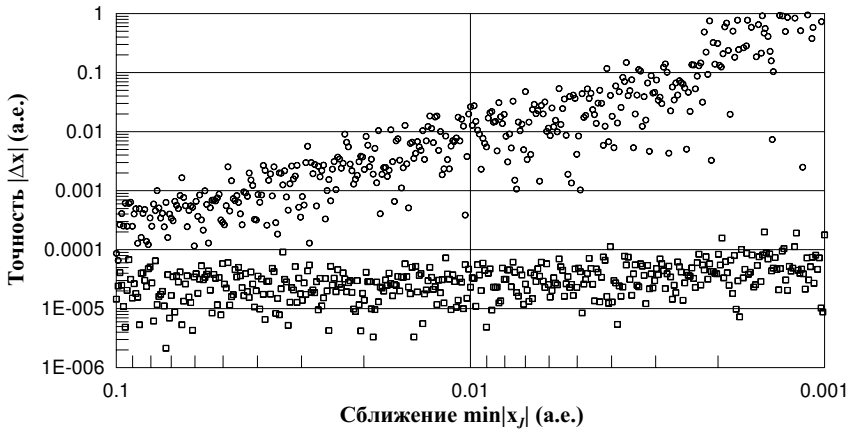


Рис. 3.15. То же, что и на рис. 3.14, но при $\varepsilon \approx 10^{-8}$, и орбиты интегрировались методом Эверхарта 7-го порядка

Тот же эксперимент был проведен в арифметике с одинарной точностью $\varepsilon \approx 10^{-8}$. Результаты приведены на рис. 3.15. Здесь эффективность йовицентрической системы немного ниже и в отличие от того, что было показано на рис. 3.14, при тесных сближениях удается повысить точность максимум на 4 порядка.

Следует отметить, что здесь не приводятся результаты для метода синхронного слежения, поскольку ошибки интегрирования уравнений (3.65) и (3.57)–(3.58) оказываются в пределах одного порядка.

Как уже говорилось выше, особенность оценки точности при последовательном использовании координатных систем состояла в том, что моменты входа в гравитационную сферу и выхода из нее при прямом и обратном интегрировании совпадали. Это позволило исключить влияние ошибок (3.63), вносимых в начальные данные систем (3.57) и (3.58), что в действительности привело к завышенным оценкам точности. На рис. 3.16 показано, насколько серьезно ухудшается оценка в случае, когда указанные моменты

становятся разными и выбираются автоматически внутри интегратора при регистрации пересечения вычисляемой орбитой границы гравитационной сферы. Как видно, за счет внесения ошибок в начальные данные при переходе от одной системы координат к другой общая точность падает почти на 2 порядка.

Выше отмечалось, что вычислительную точность можно повысить, если совместно с уравнениями (3.57) и (3.58) интегрировать уравнение (3.69). Действительно, как показывает тот же рис. 3.16, точность последовательного интегрирования расширенных систем (3.57)–(3.69) и (3.58)–(3.69) становится значительно лучше, причем независимо от того, как выбираются моменты, определяющие длительность пребывания астероида внутри гравитационной сферы.

Таким образом, было показано, что комплексный подход в использовании гравитационных координатных систем позволяет существенно уменьшить влияние ошибок округления на результаты численного интегрирования. Тем не менее прибегать к нему имеет смысл только в тех случаях, когда в уравнениях отсутствуют члены, явно зависящие от времени, либо при наличии таковых, если их влияние пренебрежимо мало. Например, такие уравнения имеют место в задачах двух неподвижных центров либо в ограниченной круговой задаче трех тел, где движение рассматривается во вращающейся системе координат. Впрочем, гравитационные системы могут быть полезны и в прикладных задачах астероидной динамики, где, однако, ее необходимо рассматривать уже как задачу многих тел (Солнце–планеты–астероид), с тем чтобы описывающие ее уравнения не зависели явно от времени. В данном случае использование фондов типа DE405 для приближенного представления движения планет в динамических уравнениях астероида совершенно недопустимо, поскольку это нарушает условие адекватности, о чем было сказано выше.

Далее покажем, насколько существенно искажается вычисляемая орбита сближающегося астероида при отклонении планеты от своего истинного положения, т.е. когда положение планеты задается с некоторой ошибкой, как, например, в фонде DE405.

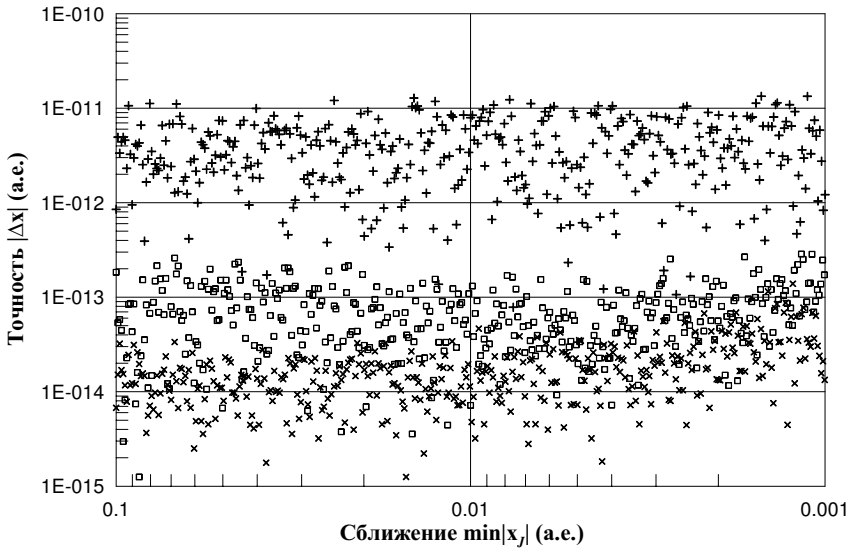


Рис. 3.16. Точность интегрирования астероидных орбит с переходом от гелиоцентрической системы координат к ювицентрической и обратно: \square — моменты времени входа в гравитационную сферу и выхода из нее при прямом и обратном интегрировании одинаковые, $+$ и \times — разные, причем \times — при совместном интегрировании уравнения (3.69)

3.6.6. Влияние ошибок в положении планеты на точность моделирования астероидной орбиты

Использование эфемерид типа DE405 (Standish, 1998) позволяет существенно упростить динамическую модель и избежать интегрирования более сложной задачи нескольких тел. Однако очевидно, что неточное положение гравитирующих планет (главным образом за счет ошибок интерполирования) неизбежно искажает представление динамики малого тела Солнечной системы, в особенности когда оно имеет тесные сближения с планетами. Насколько существенны могут быть такие искажения в динамической картине?

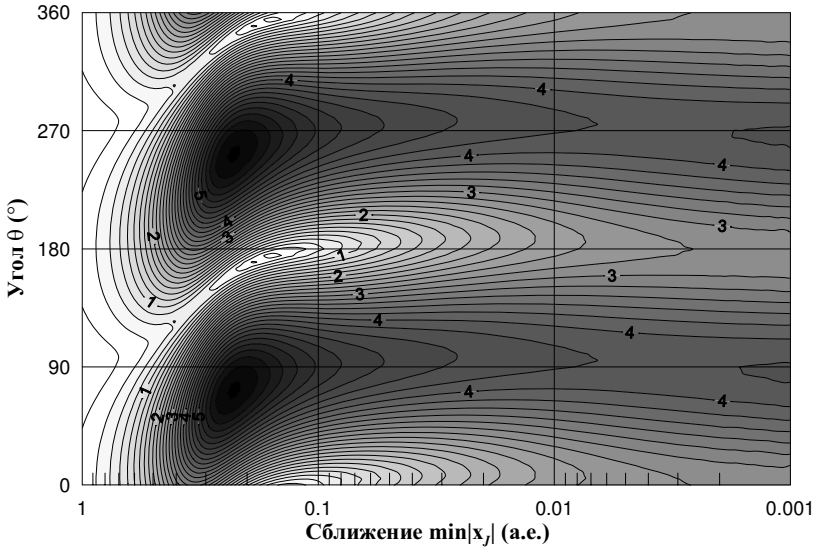


Рис. 3.17. Относительные возмущения $|\Delta \mathbf{x}_S|/|\Delta \mathbf{c}_{JS}|$ в астероидном движении, вызванные отклонениями Юпитера $\Delta \mathbf{c}_{JS}$ от его действительного положения, за 1000 сут

На рис. 3.17 показаны относительные гелиоцентрические возмущения $|\Delta \mathbf{x}_S|/|\Delta \mathbf{c}_{JS}|$ в астероидном движении (см. рис. 3.12), вызванные отклонениями Юпитера $\Delta \mathbf{c}_{JS}$ от его действительного положения, в зависимости от минимального расстояния $\min |\mathbf{x}_J|$ и угла направления ошибки θ . Таким образом, величины $\Delta \mathbf{c}_{JS}$ задавались как

$$\Delta c_{JS1} = |\Delta \mathbf{c}_{JS}| \cos \theta, \quad \Delta c_{JS2} = |\Delta \mathbf{c}_{JS}| \sin \theta.$$

Явление тесного сближения, как правило, не долговременно, и ошибки используемых планетных эфемерид в течение сближения меняются незначительно. В связи с этим отклонение планеты $\Delta \mathbf{c}_{JS}$ за каждый вариант моделирования астероидной орбиты принималось как постоянный вектор, определяемый величиной

$|\Delta \mathbf{c}_{JS}|$ и углом θ . Причем экспериментально обнаруживается, что характеристика $|\Delta \mathbf{x}_S|/|\Delta \mathbf{c}_{JS}|$ на рис. 3.17 хорошо сохраняется относительно изменения величины $|\Delta \mathbf{c}_{JS}|$. Поэтому для получения характеристики было достаточно произвести расчет с одной определенной величиной отклонений $|\Delta \mathbf{c}_{JS}|$.

Из рис. 3.17 видно, что при наиболее тесных сближениях величина возмущений $|\Delta \mathbf{x}_S|$ (только за 1000 сут) становится в 3–4 раза больше величины самих ошибок в положении планеты $|\Delta \mathbf{c}_{JS}|$. При других сближениях относительные возмущения самые разнообразные: от 0.04 до 6. Причем явно прослеживается тенденция к уменьшению $|\Delta \mathbf{x}_S|$ с увеличением $\min |\mathbf{x}_J|$. Стоит также обратить внимание на то, что характеристика возмущений π -периодична по θ , т.е. диаметрально противоположные по направлению отклонения планеты производят одинаковые возмущения в астероидной орбите.

Кроме того, наличие ошибок в положении планеты нарушает условие адекватности орбит в гелио- и планетоцентрической координатных системах. Из рис. 3.18 видно, что расхождения орбит \mathbf{x}_S и \mathbf{x}_J , вызванные ошибками в положении Юпитера $\Delta \mathbf{c}_{JS}$, довольно большие: от 0.5 до 4 величин $|\Delta \mathbf{c}_{JS}|$. Это фактически означает, что если ошибки $\Delta \mathbf{c}_{JS}$ достаточно большие (превышающие ошибки округления), то после временного использования йовицентрической координатной системы при тесном сближении астероида с Юпитером мы получим орбиту, отличающуюся от определяемой в гелиоцентрической формализации на величину порядка $|\Delta \mathbf{c}_{JS}|$.

В фонде координат больших планет DE405 ошибки интерполяции эфемерид составляют порядка 10^{-10} а.е. Поэтому, учитывая результаты рис. 3.14, можно полагать, что прибегать к подходу поочередного использования гравитационных систем с целью уменьшения влияния ошибок округления имеет смысл только при сближениях астероида с минимальными расстояниями, меньшими 0.01 а.е. В противном случае, переход к планетоцентрической системе будет ухудшать точность результатов численного моделирования.

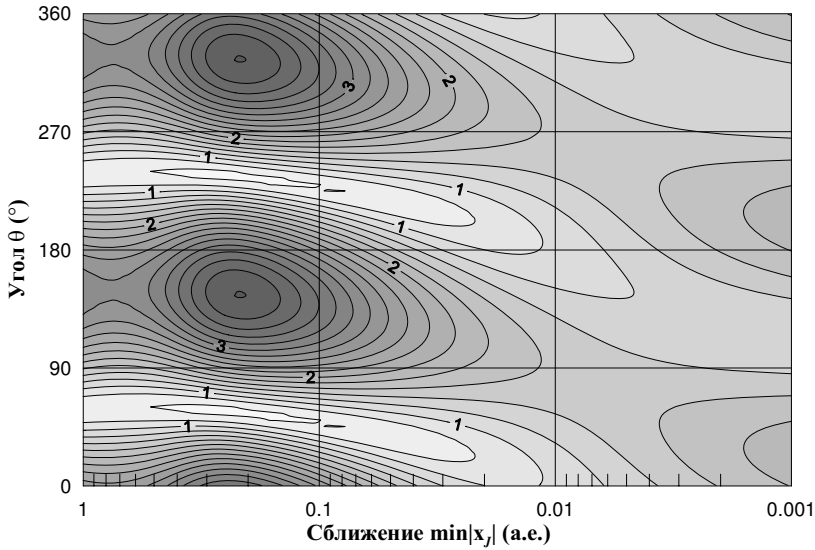


Рис. 3.18. Расхождения астероидных орбит $|\mathbf{x}_J - \mathbf{x}_S|/|\Delta\mathbf{c}_{JS}|$, вызванные отклонениями Юпитера $\Delta\mathbf{c}_{JS}$ от его действительного положения, за 1000 сут

3.7. Проблема короткопериодических возмущений и усредненные уравнения

Как уже отмечалось, прибегать к методу Энке имеет смысл, если возмущения более гладкие, нежели сама орбита. «Шероховатость» в орбите, как правило, вызывают короткопериодические возмущающие силы (Авдюшев, 2006а; Авдюшев, 2006b). На примере ограниченной круговой задачи трех тел покажем, как короткопериодические возмущения влияют на быстродействие численного интегрирования (Авдюшев, 2006b).

В рассматриваемой задаче орбита третьего тела возмущается внутренним массивным телом, быстро обращающимся около центрального. Дифференциальные уравнения движения третьего

тела можно представить в виде

$$\frac{d^2\mathbf{x}}{dt^2} = \underbrace{-\mu \frac{\mathbf{x}}{|\mathbf{x}|^3}}_{\mathbf{F}} - \underbrace{\mu_P \frac{\mathbf{x} - \mathbf{x}_P}{|\mathbf{x} - \mathbf{x}_P|^3}}_{\mathbf{P}} - \overbrace{\mu_P \frac{\mathbf{x}_P}{|\mathbf{x}_P|^3}}^{\mathbf{P}_I}, \quad (3.70)$$

где μ_P и \mathbf{x}_P — соответственно гравитационный параметр и вектор положения второго тела, причем $\mu_P \ll \mu$. Заметим, что эти уравнения записаны в координатной системе, связанной с центральным телом.

Метод Энке фактически гасит влияние кеплеровского члена \mathbf{F} и предоставляет численному методу интегрировать возмущающие силы \mathbf{P} . Следовательно, метод Энке будет эффективен только тогда, когда функция \mathbf{P} будет более гладкой, нежели \mathbf{F} , говоря иначе, когда возмущения, вызываемые силой \mathbf{P} , будут интегрироваться быстрее, чем кеплеровская орбита.

При использовании численного метода порядка p для ошибки интегрирования кеплеровской почти круговой орбиты $|\Delta\mathbf{x}_K|$ на шаге Δt справедлива оценка¹⁷

$$|\Delta\mathbf{x}_K| \approx a(n\Delta t)^{p+1}/(p+1)! \quad (3.71)$$

Внутреннее массивное тело, вызывающее в орбите третьего короткопериодические возмущения, должно быть достаточно близко к центральному телу. Тогда в \mathbf{P} будет доминировать инерциальная сила \mathbf{P}_I . Ошибку интегрирования на шаге Δt_P этой составляющей можно оценить как

$$|\Delta\mathbf{x}_P| \approx \beta a_P (n_P \Delta t_P)^{p+1}/(p+1)!, \quad (3.72)$$

где $\beta = \mu_P/(\mu + \mu_P)$; a и a_P — большие полуоси соответственно третьего и второго тел, а n и n_P — их средние движения. Введем малый параметр $\alpha = a_P/a$, причем $\alpha^{-3/2} = n_P/n$. Тогда, согласно

¹⁷В соответствии с главным членом погрешности (4.10).

(3.71) и (3.72), одинаковая точность интегрирования \mathbf{F} и \mathbf{P} достижима только в том случае, когда величины шагов Δt и Δt_P будут удовлетворять соотношению

$$\nu \equiv |\Delta t / \Delta t_P| = (\alpha\beta)^{1/(p+1)} \alpha^{-3/2}. \quad (3.73)$$

Таким образом, если $\Delta t_P > \Delta t$ или $\nu < 1$, использование метода Энке целесообразно, иначе нет. Более того, эффективность метода Энке будет тем выше, чем меньше величина ν . Эта оценка может также служить для обоснования использования других рассмотренных нами методов.

Следует заметить, что интегрирование круговых орбит выполняется с постоянным шагом, поэтому коэффициент (3.73) можно определить как $\nu = M_P/M$, где M и M_P — число шагов Δt и Δt_P , выполненных на всем интервале интегрирования.

В эллиптическом случае замечено (Авдюшев, 2006b), что для обеспечения определенной локальной ошибки шаг интегрирования следует выбирать в соответствии с постоянным шагом по эллиптической аномалии. При этом число шагов M_e за оборот увеличивается с повышением эксцентricности орбиты как

$$M_e = M\sigma, \quad \sigma = \frac{1}{\sqrt{1-e}} \frac{2}{\pi} \int_0^{\pi/2} \frac{d\psi}{\sqrt{1-\varepsilon \sin^2 \psi}}, \quad \varepsilon = -\frac{2e}{1-e}. \quad (3.74)$$

Чтобы точность интегрирования соответствовала точности в круговом случае, величину шага по эллиптической аномалии нужно уменьшить, умножая ее на $\xi^{1/(p+1)}$, где $\xi = \sqrt{1-e}$. В итоге получаем обобщенный коэффициент

$$\nu = \frac{1}{\sigma} (\alpha\beta\xi)^{1/(p+1)} \alpha^{-3/2}.$$

Впрочем, от инерциального члена можно избавиться, если записать уравнения движения относительно барицентра массивных

тел. Эти уравнения можно представить в виде (Авдюшев, 2006а)

$$\frac{d^2 \mathbf{x}_B}{dt^2} = \underbrace{-(\mu + \mu_P)}_{\mathbf{F}} \frac{\mathbf{x}_B}{|\mathbf{x}_B|^3} - \underbrace{\left(-\mu \frac{\mathbf{x}_B - \mathbf{x}_{1B}}{|\mathbf{x}_B - \mathbf{x}_{1B}|^3} + \mu \frac{\mathbf{x}_B}{|\mathbf{x}_B|^3} \right)}_{\mathbf{P}_1} \underbrace{\left(-\mu_P \frac{\mathbf{x}_B - \mathbf{x}_{2B}}{|\mathbf{x}_B - \mathbf{x}_{2B}|^3} + \mu_P \frac{\mathbf{x}_B}{|\mathbf{x}_B|^3} \right)}_{\mathbf{P}_2}, \quad (3.75)$$

где \mathbf{x}_{1B} , \mathbf{x}_{2B} и \mathbf{x} — барицентрические векторы положения первого, второго и третьего тел соответственно. Здесь \mathbf{P}_1 и \mathbf{P}_2 можно рассматривать как возмущающую силу за отклонение массивных тел от барицентра. Из условия $\mu_P \ll \mu$ следует, что $\mathbf{P}_1 \ll \mathbf{P}_2$, и поэтому возмущающей силой \mathbf{P}_1 можно пренебречь¹⁸.

Получение оценки $|\Delta \mathbf{x}_P|$ для \mathbf{P}_2 в общем случае весьма проблематично. Однако в круговой плоской задаче при использовании численного метода нечетного порядка удается получить ее предельные значения (Авдюшев, 2006b):

$$|\Delta \mathbf{x}_P|_{\min}^{\max} \approx \pm \frac{a(n\Delta t_P)^{p+1}}{(p+1)!} \frac{\beta}{(1 \mp \alpha)^{p+1}} \times \sum_{i=0}^{p-1} (1 \mp \alpha)^i (\alpha^{-3/2} - 1)^{p-1-i} \sum_{j=1}^{p-1-i} a_{ij} (\pm \alpha)^j, \quad (3.76)$$

где a_{ij} — целые числа, зависящие от p .

Пусть при некотором $\alpha = \alpha_B$ методическая ошибка $|\Delta \mathbf{x}_P|$ (3.72) равна $|\Delta \mathbf{x}_P|^{\max}$ (3.76). Тогда для всех $\alpha > \alpha_B$ переход к барицентрической системе становится необоснованным. Эмпирически (Авдюшев, 2006b) получена следующая приближенная формула для граничного значения α_B : $\alpha_B = e^{-0.2p}$.

Для достаточно малых α

$$|\Delta \mathbf{x}_P| \approx |\Delta \mathbf{x}_P|^{\max} \approx |\Delta \mathbf{x}_P|_{\min} \approx 2a(n\Delta t_P)^{p+1} \beta \alpha (\alpha^{-3/2})^{p-1} / (p+1)!.$$

¹⁸См., например, (Авдюшев, 2006а).

Отсюда коэффициент ν_B для уравнения (3.75) будет

$$\nu_B = (2\alpha\beta)^{1/(p+1)}(\alpha^{-3/2})^{(p-1)/(p+1)}. \quad (3.77)$$

Поэтому, сравнивая (3.73) и (3.77), можно ожидать, что при переходе к барицентрической координатной системе быстродействие численного интегрирования повысится в $(2\alpha^3)^{-1/(p+1)}$ раз.

Впрочем, несмотря на то, что уравнения (3.75) будут интегрироваться быстрее, нежели (3.70), этот подход, как правило, на практике не позволяет полностью разрешить проблему короткопериодических возмущений, в особенности при использовании численных методов высокого порядка; и применение к (3.75) каких-либо методов теории специальных возмущений также будет безуспешным.

С другой стороны, интересно отметить, что сам переход к барицентрической системе координат можно рассматривать как самостоятельный метод повышения эффективности численного интегрирования, однако прибегать к нему имеет смысл только при достаточно малых α .

Подводя итог вышесказанному, можно полагать, что методы теории специальных возмущений будут эффективны только в тех задачах, где отсутствуют короткопериодические возмущающие силы, либо при наличии таковых, но лишь в случае, если $\nu \ll 1$. Например, эти методы целесообразно использовать в задачах динамики близких спутников.

В случае наличия короткопериодических возмущений, когда $\nu, \nu_B \gg 1$, некоторые методы могут даже ухудшать результаты численного интегрирования. На рис. 3.19 показано, насколько хуже становится поведение ошибки интегрирования $|\Delta \mathbf{x}|$ после применения к уравнениям (3.70) стабилизации Баумгарте.

Представленные ошибки оценивались относительно более точного численного решения системы (3.70), полученного путем уменьшения шага интегрирования. Начальные условия задачи соответствовали круговой орбите с единичным радиусом, а параметры α и β подобраны таким образом, чтобы они отвечали системе Юпитер–Ио–Гималия. Для качественного сравнения здесь также

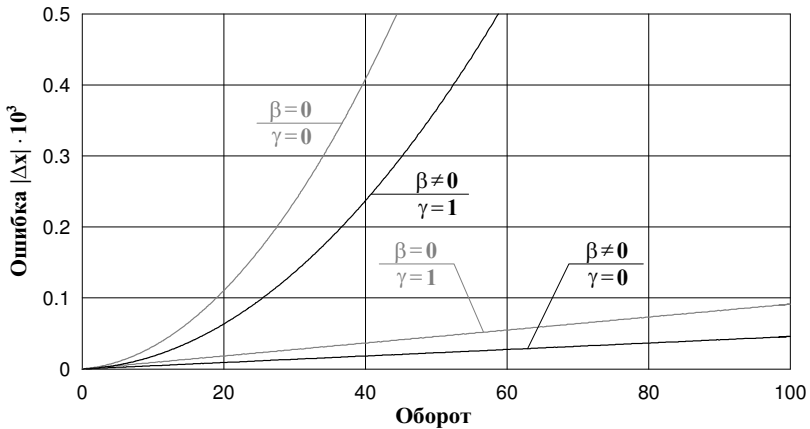


Рис. 3.19. Поведение ошибок интегрирования в ограниченной круговой задаче трех тел (3.70) с параметрами $\alpha = 0.04$ и $\beta = 10^{-4}$ ($\nu = 10.4$, $n_P/n = 125$) при использовании классического метода Рунге–Кутты в стабилизированном ($\gamma = 1$) и нестабилизированном ($\gamma = 0$) случаях (серым цветом показаны результаты рис. 3.7 и 3.8)

приведены результаты рис. 3.7 и 3.8 (для невозмущенных уравнений), хотя следует заметить, что ошибки интегрирования уравнений (3.70) были получены с шагом в $10 (\approx \nu)$ раз меньшим. Без уменьшения шага ошибки оказываются очень большими, сравнимыми с размерами самой орбиты.

Прежде всего следует обратить внимание на то, что ошибка интегрирования исходных уравнений (3.70) растет линейно, а не квадратично. Это главным образом связано с тем, что инерциальная сила \mathbf{P}_I , вызывающая короткопериодические возмущения и определяющая точность интегрирования, не зависит от интегрируемых переменных. Вместе с тем после применения стабилизации Баумгарте поведение ошибки, наоборот, становится квадратичным. При этом интегральное соотношение довольно хорошо сохраняется: $\mathcal{H} \approx 2 \cdot 10^{-9}$. Плохое поведение ошибки (как в нестабилизированном случае), по-видимому, объясняется тем, что

в возмущенной задаче сохранение интегрального соотношения \mathcal{H} уже не является достаточным для уменьшения влияния ляпуновской неустойчивости и, более того, влияние уже бесполезных стабилизирующих членов вносит дополнительные ошибки, которые ее только усиливают.

На практике для разрешения проблемы короткопериодических возмущений, как правило, прибегают к одному из двух способов. Первый способ состоит в том, что массы быстрообращающихся тел включаются в массу центрального¹⁹, во втором способе влияние от массивных объектов моделируется с помощью так называемых гауссовых колец²⁰. Оба способа приводят к упрощенным уравнениям движения: в первом — это возмущенные уравнения задачи двух тел с модифицированным гравитационным параметром:

$$\frac{d^2\mathbf{x}}{dt^2} = -(\mu + \mu_P) \frac{\mathbf{x}}{|\mathbf{x}|^3}; \quad (3.78)$$

во втором — уравнения, усредненные по быстрым аномалиям массивных объектов (Дубошин, 1961):

$$\frac{d^2\mathbf{x}}{dt^2} = -\mu \frac{\mathbf{x}}{|\mathbf{x}|^3} + \left(\frac{\partial V}{\partial \mathbf{x}} \right)^T. \quad (3.79)$$

Второе слагаемое в правой части уравнения (3.79) представляет притяжение гауссова кольца радиуса a_P с гравитационным параметром μ_P . Потенциал кольца V выражается как

$$V(\rho, x_3) = \frac{\mu_P K(\kappa)}{\sqrt{(\rho + a_P)^2 + x_3^2}},$$

$$\rho = \sqrt{x_1^2 + x_2^2}, \quad \kappa^2 = \frac{4\rho a_P}{(\rho + a_P)^2 + x_3^2},$$

где

$$K(\kappa) = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\psi}{\sqrt{1 - \kappa^2 \sin^2 \psi}}$$

¹⁹См., например, (Бордовицына, Быкова, 1978; Rocher, Chapront, 1996).

²⁰См., например, (Jacobson, 2000; Emel'yanov, Arlot, 2008).

— полный эллиптический интеграл первого рода. Уравнения (3.79) представляют орбитальное движение внешних объектов точнее, нежели (3.78), хотя в то же время усредненные уравнения гораздо сложнее.

Дифференцируя V по координатам, получим составляющие силы притяжения гауссова кольца:

$$\frac{\partial V}{\partial x_1} = \frac{x_1}{\rho} \frac{\partial V}{\partial \rho}, \quad \frac{\partial V}{\partial x_2} = \frac{x_2}{\rho} \frac{\partial V}{\partial \rho},$$

$$\frac{\partial V}{\partial \rho} = \frac{\mu_P}{\sqrt{(\rho + a_P)^2 + x_3^2}} \left[\frac{dK}{d\kappa^2} \frac{\partial \kappa^2}{\partial \rho} - \frac{(\rho + a_P)K(\kappa)}{(\rho + a_P)^2 + x_3^2} \right],$$

$$\frac{\partial V}{\partial x_3} = \frac{\mu_P}{\sqrt{(\rho + a_P)^2 + x_3^2}} \left[\frac{dK}{d\kappa^2} \frac{\partial \kappa^2}{\partial x_3} - \frac{x_3 K(\kappa)}{(\rho + a_P)^2 + x_3^2} \right],$$

где

$$\frac{\partial \kappa^2}{\partial \rho} = \frac{\kappa^2}{\rho} \frac{-\rho^2 + a_P^2 + x_3^2}{(\rho + a_P)^2 + x_3^2}, \quad \frac{\partial \kappa^2}{\partial x_3} = -\frac{2\kappa^2 x_3}{(\rho + a_P)^2 + x_3^2}.$$

Производную от функции $K(\kappa)$ можно представить как

$$\frac{dK}{d\kappa^2} = \frac{1}{2\kappa^2} \left[\frac{E(\kappa)}{1 - \kappa^2} - K(\kappa) \right],$$

где

$$E(\kappa) = \frac{2}{\pi} \int_0^{\pi/2} \sqrt{1 - \kappa^2 \sin^2 \psi} \, d\psi$$

— полный эллиптический интеграл второго рода.

Эллиптические интегралы K и E обычно вычисляются по итерационной схеме Кинга (Ланс, 1962):

$$a_0 = 1, \quad b_0 = \sqrt{1 - \kappa^2}$$

$$a_{i+1} = \frac{1}{2}(a_i + b_i), \quad b_{i+1} = \sqrt{a_i b_i}, \quad c_{i+1} = a_{i+1}^2 - b_{i+1}^2.$$

Известно, что

$$\lim_{i \rightarrow \infty} a_i = \lim_{i \rightarrow \infty} b_i = \frac{1}{K}.$$

Таким образом, после сходимости итерационного процесса с некоторой заданной точностью получаем значение эллиптического интеграла K . Затем находим функцию E из соотношения

$$\frac{E}{K} = 1 - \frac{1}{2} \sum_{i=0}^{\infty} 2^i c_i^2.$$

Итерационная схема Кинга сходится довольно быстро: обычно даже при достаточно больших κ , близких к единице, для вычисления эллиптических интегралов с точностью до последней значащей цифры (в арифметике с двойной точностью) требуется всего несколько итераций.

Во избежание вычислений эллиптических интегралов можно прибегнуть к многоточечному представлению гауссова кольца, где его влияние рассматривается как притяжение от нескольких стационарных (неподвижных в координатной системе) материальных точек, лежащих на кольце (орбите близкого массивного объекта). Многоточечную модель можно представить в виде

$$\mathbf{P} = -\frac{\mu_P}{N} \sum_{j=1}^N \frac{\mathbf{x} - \mathbf{x}_j}{|\mathbf{x} - \mathbf{x}_j|^3}, \quad (3.80)$$

где N — число материальных точек, а \mathbf{x}_j — их положения, координаты которых определяются по следующим формулам:

$$x_{j1} = a_P \cos \frac{2\pi j}{N}, \quad x_{j2} = a_P \sin \frac{2\pi j}{N}, \quad x_{j3} = 0.$$

Очевидно, что точность представления гауссова кольца (3.80) будет тем выше, чем больше будет число N . Однако, с другой стороны, увеличение N ведет к понижению скорости вычисления \mathbf{P} . Поэтому необходимо выбирать оптимальное число N , которое бы обеспечивало достаточно высокую точность моделирования силы

при высокой скорости ее вычисления. Как показывает практика, в задачах высокоточного численного моделирования движения далеких спутников, где короткопериодические возмущения вызываются притяжениями близких спутников типа галилеевых, часто бывает достаточно использовать формулу (3.80) с $N = 6$, на вычисление которой требуется приблизительно то же время, что и для гауссова кольца.

Как известно, высокая точность представления орбитального движения с помощью усредненных уравнений (3.79) достигается главным образом за счет того, что они учитывают вековые возмущения (первого порядка) в орбитальных элементах, вызванные короткопериодическими возмущающими силами. Однако после усреднения остаются еще неучтенными довольно большие вековые эффекты в быстрой переменной (Авдюшев, 2006а), которая определяет положение далекого объекта на орбите. Этот недостаток, например, становится значительным в тех задачах, где орбиты близкого (внутреннего) и далекого (внешнего) спутников достаточно близки, как, например, для Каллисто и Гималии, спутников Юпитера. Причем усредненные уравнения будут настолько плохо представлять реальную орбиту, что их использование может оказаться совершенно неоправданным на фоне простых уравнений с модифицированным гравитационным параметром.

Возникновение неучитываемых вековых эффектов при использовании усредненных уравнений объясняется в работе (Авдюшев, 2006а), тем не менее, насколько нам известно, эта проблема на данный момент фактически остается еще не разрешенной. В той же работе, в частности, показано, что при умеренном понижении модельной точности интерпретация гравитационного влияния быстрых внутренних объектов гауссовыми кольцами позволяет в несколько раз повысить быстродействие численного интегрирования при сохранении методической точности.

3.8. Сравнительный анализ эффективности уравнений

Выше были изложены основополагающие идеи и принципы построения методов теории специальных возмущений, а также обосновано их применение к решению задач небесной механики. Далее проведем сравнительный анализ эффективности некоторых из них применительно к численному моделированию спутниковых, астероидных и планетных орбит, для того чтобы выработать четкие рекомендации по использованию методов (Авдюшев, 2007).

3.8.1. Численный эксперимент

Численная эффективность уравнений, получаемых методами теории специальных возмущений (табл. 3.1), исследовалась применительно к решению задач динамики спутников, астероидов и планет (табл. 3.2 и 3.3). В табл. 3.1 ссылки даны на формулы, представленные выше; Sn — соответствующие интегрируемым переменным обозначения для характеристик численной эффективности методов на приводимых ниже рисунках; Ne — число интегрируемых уравнений. Следует заметить, что в методах Энке опорные решения перевычислялись через каждые 2 оборота. В табл. 3.2 \mathcal{T} , e и i — период, эксцентриситет и наклонение исследуемых орбит соответственно (наклонения для спутников даны относительно экватора центральной планеты, а для астероида — относительно эклиптики); в табл. 3.3 J — отношение величины возмущающей силы, вызываемой несферичностью планеты, к величине центральной силы; ν — показатель влияния короткопериодических возмущений.

В задачах учитывались следующие возмущающие факторы: \mathcal{J} — влияние несферичности центрального тела (для спутников); \bullet — притяжение от массивных тел (галилеевых спутников и планет); \ast — притяжение от Солнца (для спутников). Другие возмущающие факторы не рассматривались ввиду их незначительного влияния на эффективность численного моделирования.

Таблица 3.1. Уравнения орбит

Переменные	S_n	Ссылка	N_e
$\mathbf{x}, \dot{\mathbf{x}}$	(t)	\mathbf{x}	(3.1) 6
$\delta\mathbf{x}, \delta\dot{\mathbf{x}}$	(t)	$\delta\mathbf{x}$	(3.45) 6
$\mathbf{x}, \dot{\mathbf{x}}, h$	(t)	\mathbf{st}	(3.21) 7
$\mathbf{x}, \dot{\mathbf{x}}, h$	(t)	\mathbf{nz}	(3.32) 7
$\mathbf{x}, \dot{\mathbf{x}}, h$	(t)	\mathbf{cn}	(3.28) 7
$\mathbf{x}, \dot{\mathbf{x}}, h, t$	(t)	$\delta\mathbf{t}$	(3.30) 8
$\mathbf{c}, \mathbf{g}, l$	(t)	\mathbf{ry}	(3.35) 7
$\mathbf{x}, \mathbf{x}', \mathbf{g}, h, \tau$	(E)	\mathbf{sb}	(3.8) 11
$\mathbf{u}, \mathbf{u}', h, \tau$	(E)	\mathbf{u}	(3.10) 10
$\delta\mathbf{u}, \delta\mathbf{u}', \delta h, \delta\tau$	(E)	$\delta\mathbf{u}$	(3.46) 10
$\mathbf{x}, \dot{\mathbf{x}}, t$	(s)	\mathbf{sm}	(3.12) 7
$\mathbf{x}, \dot{\mathbf{x}}, t$	(E_G)		(3.14) 7
$\mathbf{x}, \dot{\mathbf{x}}, h, \tau$	(E)	τ	(3.8) 8
$\mathbf{x}_B, \dot{\mathbf{x}}_B$	(t)	\mathbf{br}	6

Таблица 3.2. Небесные объекты и их орбиты

Объект	\mathcal{T} (сут)	e	i ($^\circ$)	Интервал (г., об.)	
Спутники					
Амальтея	0.499	0.003	0.3	100	73000
Гималия	247.767	0.166	30.2	100	152
Фобос	0.319	0.015	1.1	100	114500
300 км	0.063	0.000	50.0	0.038	222
ГСС	0.997	0.010	10.0	40	14600
Астероиды и планеты					
Фазетон	523.609	0.890	22.2	1433	1000
Церера	1680.907	0.079	10.6	4602	1000
Меркурий	87.969	0.205	7.0	240	1000
Юпитер	4339.289	0.048	1.3	11880	1000

Для численного решения уравнений движения применялся интегратор Эверхарта 15-го порядка (Everhart, 1973; Everhart, 1974; Everhart, 1985)²¹, который основан на видоизмененных формулах неявных методов Рунге–Кутты типа Бутчера и строится на разбиениях Гаусса–Радо (Butcher, 1964).

Для оценки точности интегрирования варьировался параметр интегратора $\|e\|_{tol}$ (см. подразд. 4.3.7), который фактически задает локальную ошибку, и по решению с меньшим значением параметра оценивалась глобальная ошибка решения с большим значением параметра. В качестве показателя быстродействия рассматривалось число выполненных шагов. В результате для каждой системы интегрируемых уравнений были получены зависимости точности интегрирования от быстродействия, которые являются удобными характеристиками для исследования эффективности численного интегрирования.

Следует заметить, что принятая оценка быстродействия по числу шагов интегрирования здесь вполне обоснована. Конечно, предлагаемые методы усложняют уравнения орбитального движения, что естественно сказывается на времени вычисления их правых частей. Поэтому, казалось бы, каждый шаг интегрирования для разных систем должен быть по-разному трудоемок. Тем не менее нельзя забывать, что подавляющая часть времени в современных моделях тратится на вычисление возмущающей функции \mathbf{P} , которая присутствует во всех рассматриваемых системах; и все отличительные особенности тех или иных уравнений просто меркнут на фоне конгломерата \mathbf{P} .

Таким образом, для оценки быстродействия в анализе эффективности методов достаточно знать лишь число обращений в интеграторе к процедуре вычисления \mathbf{P} , либо пропорциональное ему число шагов, выполненных за весь процесс интегрирования. Кроме того, такая характеристика быстродействия замечательна тем, что она не зависит ни от оптимизации численной модели, ни от возможностей компьютерного процессора.

²¹См. также подразд. 4.3.7.

Таблица 3.3. Учитываемые возмущения

Объект	Возмущения			
Спутники				
Амальтея	\mathcal{J}	●	※	$J = 3.4 \cdot 10^{-3}$
Гималия		●	※	$\nu = 41$ (Ио)
Фобос	\mathcal{J}		※	$J = 3.8 \cdot 10^{-4}$
300 км	\mathcal{J}	●	※	$J = 1.5 \cdot 10^{-3}$
ГСС	\mathcal{J}	●	※	$J = 3.7 \cdot 10^{-5}$
Астероиды и планеты				
Фаэтон		●		$\nu = 0.9$ (Меркурий)
Церера		●		$\nu = 3.9$ (Меркурий)
Меркурий		●		
Юпитер		●		$\nu = 9.4$ (Меркурий)

3.8.2. Выбранные объекты

Спутниковое движение представляют 3 естественных спутника Юпитера и Марса (Амальтея, Гималия и Фобос) и 2 искусственных спутников Земли (низколетящий с высотой 300 км и геосинхронный). Динамика естественных спутников рассматривалась на 100-летнем интервале времени: приблизительно столько времени прошло с момента их открытия.

Амальтея и Фобос — близкие спутники, движущиеся вблизи плоскости экватора планеты по почти круговым орбитам. Их орбиты сильно возмущаются несферичностью центральной планеты. Кроме того, спутник Юпитера также испытывает сильное гравитационное влияние от массивных галилеевых спутников. Орбитальные периоды объектов меньше суток, и за 100-летний интервал времени они совершают довольно большое число оборотов, более 50000. Поэтому с точки зрения численного интегрирования их динамику следует рассматривать как долговременную.

Далекый спутник, Гималия, имеет большой орбитальный период и за 100 лет делает всего 152 оборота. Орбита спутника нерегулярна. Главный возмущающий фактор — притяжение от Солнца.

Несмотря на это, интегрирование орбиты Гималии существенно усложняется короткопериодическими возмущениями от галилеевых спутников, что является причиной низкого быстродействия численного процесса.

Низколетящий искусственный спутник Земли (с высотой 300 км) — это быстрый околоземный объект. За 2 недели он совершает 222 оборота. Именно на этом интервале рассматривалось его орбитальное движение: во-первых, потому что, вообще говоря, интересно, насколько целесообразно использование исследуемых методов на таком коротком интервале времени применительно к интегрированию достаточно гладкой орбиты; во-вторых, рассматривать движение низких спутников на более длительном временном интервале просто не имеет смысла, поскольку он довольно часто подвергается орбитальной коррекции за уменьшение его высоты вследствие сопротивления атмосферы. При этом ввиду незначительного влияния на эффективность численного интегрирования сопротивление атмосферы все же не учитывалось.

Геосинхронный объект (ГСС) интересен как представитель наиболее загрязненной области околоземного пространства, как один из многочисленных фрагментов космического мусора. Его динамика моделировалась на 40-летнем временном интервале, фактически соизмеримом с эрой освоения (и загрязнения) космоса.

В астероидных и планетных задачах движение объектов моделировалось на интервале 1000 оборотов. Рассматривались четыре объекта: два астероида, Фаэтон и Церера, и две планеты, Меркурий и Юпитер. Фаэтон имеет сильно вытянутую орбиту со сложной структурой возмущений. Этот объект выбран специально для того, чтобы на его примере продемонстрировать возможности регуляризирующих преобразований. Орбиты Цереры и Юпитера почти круговые. Однако их интегрирование усложнено коротко периодическими возмущениями от планет земной группы, в особенности от Меркурия. Между тем при моделировании движения самого Меркурия проблемы короткопериодических возмущений, очевидно, не возникает, что весьма благоприятно сказывается на результатах численного интегрирования.

3.8.3. Численные результаты

Характеристики точность–быстродействие представлены на рис. 3.20–3.27. Для оценки значимости ошибок интегрирования на рисунках пунктирными линиями приведены два уровня: один соответствует величине большой полуоси орбиты (a), второй — приблизительно односекундной угловой ошибке относительно земного наблюдателя ($1''$).

Полученные результаты для спутниковых задач говорят о том, что методы теории специальных возмущений наиболее эффективны и могут быть рекомендованы к применению только для долгосрочного численного моделирования спутниковых орбит с гладкой структурой возмущений (рис. 3.20, 3.23 и 3.25). Впечатляющие результаты получаются при использовании КS-уравнений (\mathbf{u} , $\delta\mathbf{u}$) и уравнений Роя (\mathbf{ry}). Так, при сохранении точности интегрирования с их помощью удается повысить быстродействие в 3–7 раз. Кроме того, в случае Фобоса метод Энке в КS-переменных ($\delta\mathbf{u}$) за счет ослабления влияния ошибок округления позволяет повысить уровень наивысшей точности почти на 2 порядка.

Высокая эффективность этих уравнений связана прежде всего с тем, что они обладают стабилизирующим эффектом. Несмотря на это, использование стабилизированных уравнений (\mathbf{st}) существенно менее эффективно. По-видимому, это объясняется наличием в этих уравнениях искусственно введенных членов (стабилизирующих возмущений), поведение которых связано с ошибками интегрирования, не имеющими никакого отношения к физике задачи. Поэтому данный артефакт, вполне обоснованный с точки зрения стабилизации как средство для борьбы с ляпуновской неустойчивостью, все же имеет побочный эффект, который со временем заметно искажает динамическую картину задачи.

На примере Амальтеи (рис. 3.21) оценена также эффективность различных стабилизирующих преобразований (\mathbf{st} , \mathbf{nz} , \mathbf{cn} , $\delta\mathbf{t}$). В результате оказалось, что все стабилизирующие подходы, кроме канонического (\mathbf{cn}), одинаковы хороши. Низкая эффективность канонической стабилизации объяснялась выше.

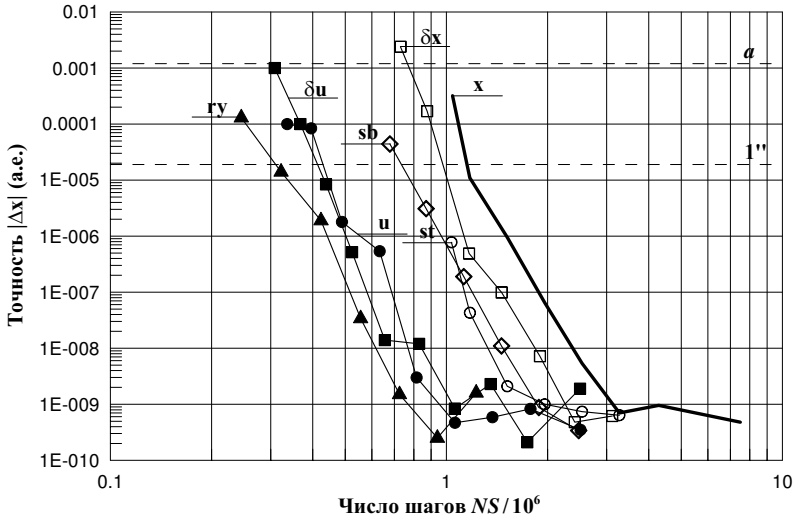


Рис. 3.20. Амальтея

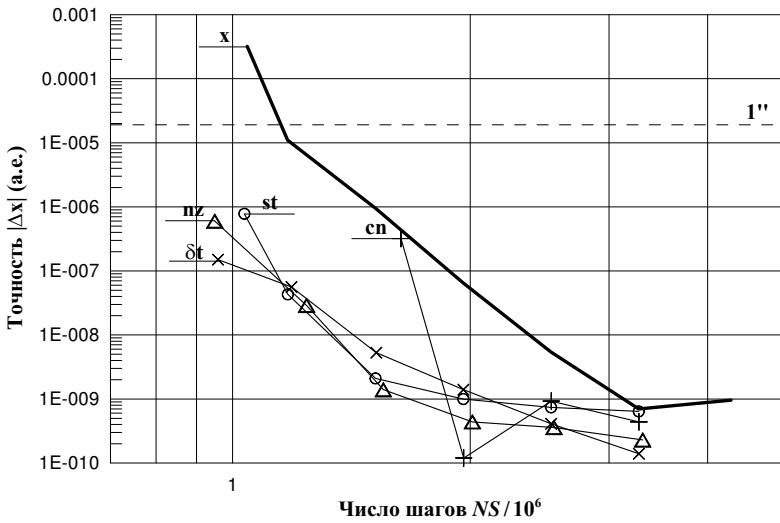


Рис. 3.21. Амальтея (стабилизирующие преобразования)

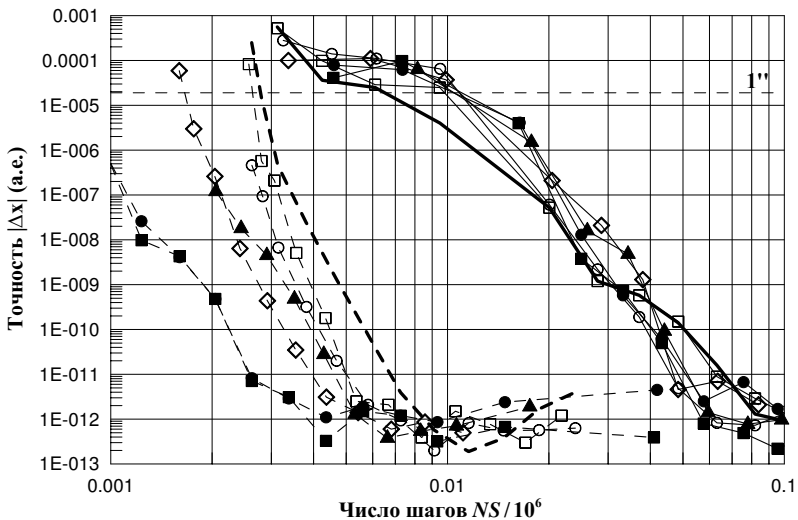


Рис. 3.22. Гималия

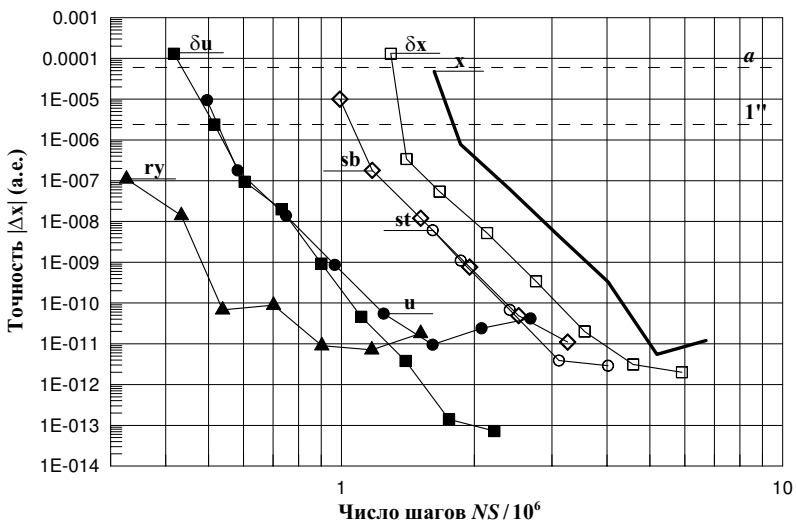


Рис. 3.23. Фобос

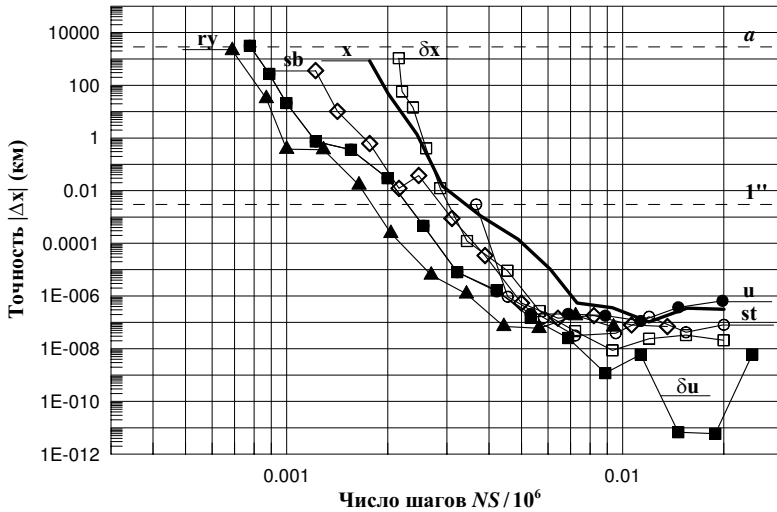


Рис. 3.24. Низколетящий искусственный спутник Земли

Далее, как показывают результаты на рис. 3.22, при моделировании движения далекого спутника (Гималии) все попытки повысить эффективность численного интегрирования с использованием рассматриваемых методов терпят неудачу вследствие короткопериодических возмущений от галилеевых спутников. Из рисунка также видно, что при отсутствии короткопериодических возмущений (пунктирные линии) эффективность интегрирования существенно повышается, причем характеристики распределяются здесь почти в том же порядке значимости, как и для близких спутников.

Наконец, следует заметить, что с уменьшением временного интервала эффективность методов теории специальных возмущений заметно понижается (рис. 3.24). Впрочем, несмотря на это, их применение все еще остается целесообразным.

Для астероида Фэтон показатель значимости короткопериодических возмущений ν меньше 1. Поэтому применение методов

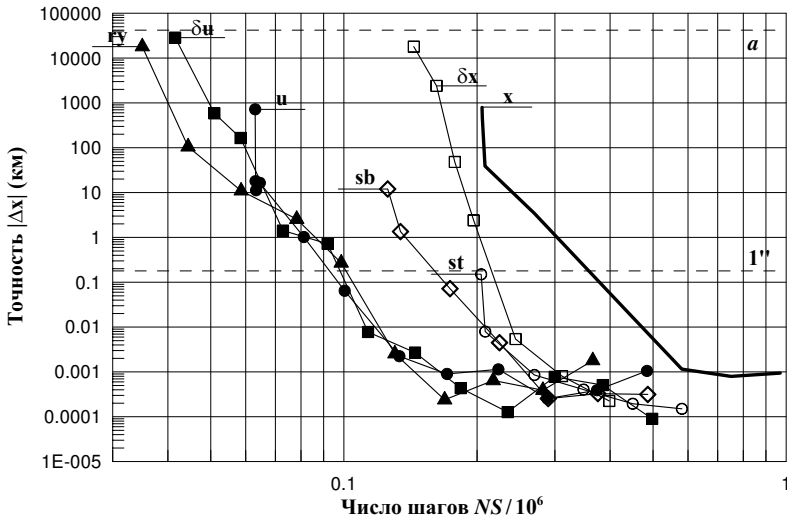


Рис. 3.25. Геосинхронный спутник Земли

теории специальных возмущений вполне обосновано, что подтверждают характеристики на рис. 3.26. В данном случае наивысшая эффективность численного интегрирования достигается путем использования регуляризирующих преобразований (u , δu , sb), что является ожидаемым результатом, поскольку Фаэтон имеет очень вытянутую орбиту с эксцентриситетом $e = 0.89$.

На примере Фаэтона также исследованы возможности сглаживающих преобразований (sm). Результаты приведены на рис. 3.27. Здесь характеристики соответствуют следующим независимым переменным: t — время; E , ε , v и E_G — эксцентрическая, эллиптическая, истинная и обобщенная эксцентрическая (3.14) аномалии соответственно; l — дуга орбиты; $E(\tau)$ — эксцентрическая аномалия (с интегрируемой переменной τ). Как видно из рисунка, для интегрирования сильновытянутых орбит со сложной структурой возмущений, как у Фаэтона, имеет смысл прибегать только к таким сглаживающим преобразованиям, где в качестве независимых переменных выступают эксцентрическая аномалия и ее

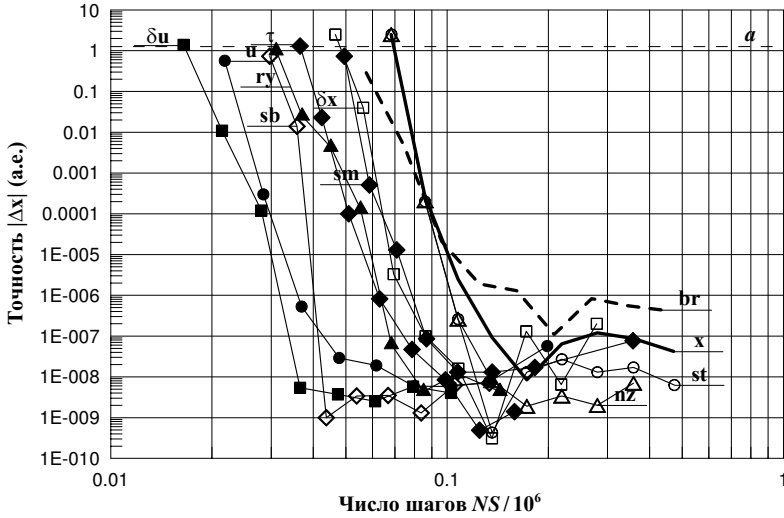


Рис. 3.26. Фэтон

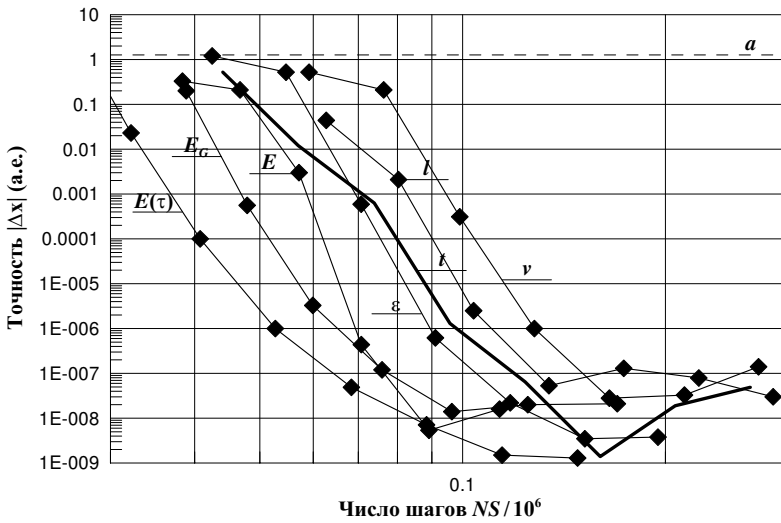


Рис. 3.27. Фэтон (сглаживающие преобразования)

аналоги. Повышение эффективности при использовании обобщенной эксцентрической аномалии объясняется тем, что орбитальная динамика вдоль нее сглаживается не только за эллиптичность орбиты, но и за нерегулярные и довольно большие возмущения со стороны больших планет.

Как показывает практика, потеря точности численного решения связана в основном с большими ошибками в переменной времени за счет ее неравномерного поведения вдоль аномалии. Из рис. 3.27 видно, что введение временного элемента τ , который ведет себя линейно в невозмущенном случае, позволяет существенно повысить эффективность численного интегрирования.

Для Цереры и Юпитера показатель ν довольно высок, отсюда использование рассматриваемых методов фактически ничего не дает (рис. 3.28 и 3.30). Более того, для Юпитера, например, замечательные в случае Фазтона методы (KS- и SB-регуляризация) (\mathbf{u} , $\delta\mathbf{u}$, \mathbf{sb}) даже ухудшают эффективность численного интегрирования. В то же время следует заметить, что переход к барицентрической координатной системе (\mathbf{br}) в планетной задаче может существенно ослабить влияние короткопериодических возмущений от Меркурия и тем самым повысить точность интегрирования на несколько порядков (рис. 3.30), тогда как в астероидной задаче результаты качественно не улучшаются.

В случае Меркурия результаты показывают (рис. 3.29), что для численного исследования его динамики исключительно хороши уравнения в KS-переменных (\mathbf{u} , $\delta\mathbf{u}$). Так, при сохранении довольно высокой точности можно повысить быстродействие численного интегрирования почти в 7 раз. Так же, как и для Фобоса, метод Энке ($\delta\mathbf{u}$) позволяет повысить уровень наивысшей точности почти на 2 порядка. Низкая эффективность метода вариации произвольных постоянных, который представляют уравнения Роя (\mathbf{ry}), по-видимому, обусловлена тем, что орбита Меркурия достаточно быстро прецессирует. Это, в свою очередь, приводит к быстрому изменению вектора количества движения \mathbf{c} , который, возможно, интегрируется с недостаточно высокой точностью.

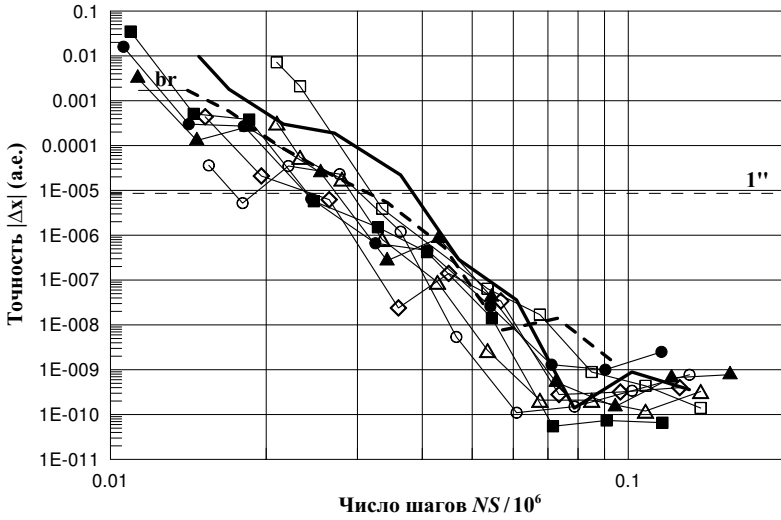


Рис. 3.28. Церера

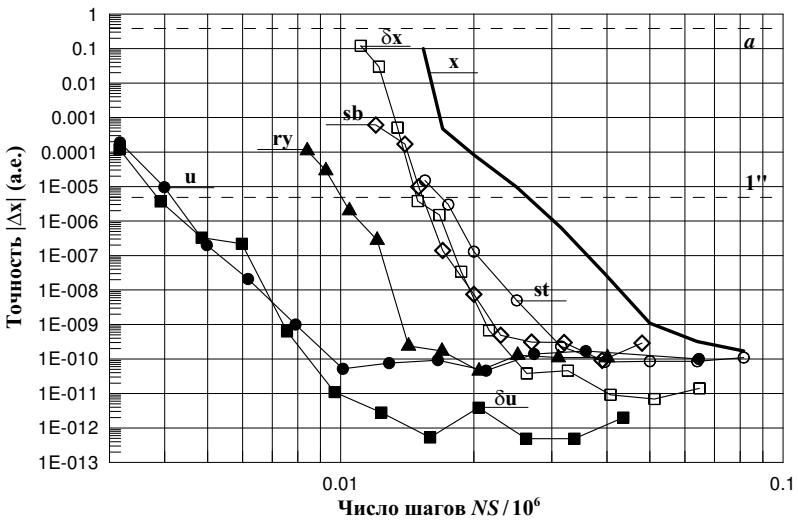


Рис. 3.29. Меркурий

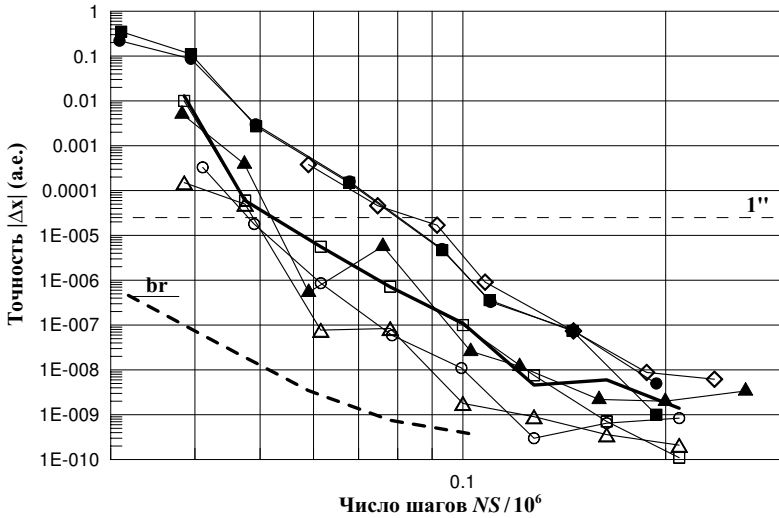


Рис. 3.30. Юпитер

Наконец, что касается метода Энке (δx), который до сих пор не был удостоен наших комментариев, то, как показывают результаты, из всех рассматриваемых в эксперименте методов он оказался наименее эффективен.

Таким образом, методы теории специальных возмущений целесообразно применять только в тех задачах, где отсутствуют короткопериодические возмущающие силы, либо при наличии таковых, но лишь в случае, если $\nu < 1$. При отсутствии короткопериодических возмущений исключительно эффективным оказывается применение KS-регуляризации (\mathbf{u}) и метода Энке в KS-переменных ($\delta \mathbf{u}$). Кроме того, для численного исследования динамики близких спутников можно также рекомендовать метод вариации параметров в интерпретации Роя (\mathbf{ry}).

Метод важнее открытия.

Л.Д. Ландау

Глава 4. МЕТОДЫ ИНТЕГРИРОВАНИЯ

Почти все важные для современной практики дифференциальные уравнения, описывающие орбитальное движение, не интегрируются аналитически. Поэтому для их решения прибегают к приближенным методам интегрирования, которые условно делят на аналитические и численные. Рассмотрим их основные принципы на примере метода малого параметра и метода разложения в ряд Тейлора.

Пусть необходимо решить уравнение орбиты вида $\mathbf{x}' = \mathbf{f}(t, \mathbf{x}, \varepsilon)$ при условии $\mathbf{x}(0) = \mathbf{x}_0$. Здесь \mathbf{x} — искомое решение; t — независимая переменная; ε — некоторый малый параметр, причем такой, что при $\varepsilon = 0$ уравнение имеет аналитическое решение, выраженное через элементарные функции.

Именно это решение $\bar{\mathbf{x}} : \bar{\mathbf{x}}' = \mathbf{f}(t, \bar{\mathbf{x}}, 0)$, принимается в качестве опорного в аналитических методах, а приближенное решение представляется в виде усеченного ряда по степеням ε :

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{i=1}^p \mathbf{x}_i \varepsilon^i,$$

где p — порядок точности решения, а \mathbf{x}_i — неизвестные коэффициенты, которые определяются путем подстановки в дифференциальное уравнение степенного ряда и уравнивания коэффициентов при одинаковых степенях ε . При этом конструкция решения не зависит от начального условия, т.е. решение можно рассматривать как общее. Известно, что оно будет пригодным на временном интервале порядка ε^{-p} . Впрочем, эти замечательные особенности аналитического решения дискредитируются существенными недостатками, которые уже в настоящее время не могут быть не приняты во внимание.

Обязательным условием для получения аналитического приближенного решения является наличие хорошего опорного, которое бы обеспечивало малость возмущений $\delta\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$. Однако это не всегда возможно, как, например, в задачах астероидной динамики. Кроме того, для достижения достаточно высокой точности приближенного решения требуется большой порядок аппроксимации, что существенно усложняет решение. Именно эти недостатки становятся главной причиной, почему в современной практике все чаще прибегают к численной альтернативе в решении дифференциальных уравнений.

В методе разложения в ряд Тейлора в качестве опорного $\bar{\mathbf{x}}$ выступает начальное \mathbf{x}_0 , а решение представляется в виде усеченного степенного ряда по t :

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{i=1}^p \mathbf{x}_i t^i,$$

где p — порядок метода, а \mathbf{x}_i — коэффициенты ряда Тейлора. Несмотря на то, что численное решение пригодно лишь для малых t , эта трудность разрешается путем пошагового интегрирования: большой интервал времени делится на малые подынтервалы (шаги), на которых последовательно шаг за шагом определяются приближенные решения, при этом на каждом следующем шаге в качестве опорного выбирается решение предыдущего шага.

Отсюда главное достоинство численных методов состоит в том, что с их помощью почти всегда можно получить решение дифференциальных уравнений. Кроме того, методическая точность может быть повышена не только увеличением порядка аппроксимирующей формулы, но и уменьшением шага интегрирования. В связи с этим формульное представление численного решения на шаге может быть довольно простым. Хотя необходимо заметить, что вследствие пошагового интегрирования компьютерная реализация каждого численного метода дает только частное решение. «Свободен только первый шаг» с произвольными начальными данными, «но мы рабы другого . . . » (И. Гете).

В данной главе излагаются основные принципы построения методов численного интегрирования, которые широко применяются в небесно-механической практике. В частности, рассматриваются методы Рунге–Кутты, экстраполяционные и многошаговые методы, а также так называемые геометрические методы.

4.1. Терминология

Дифференциальное уравнение первого порядка — это уравнение вида

$$x' = f(t, x) \quad (4.1)$$

с заданной функцией f . Здесь штрих означает полную производную по независимой переменной.

Функция $x = x(t)$ называется решением (4.1), если для всех t выполняется равенство

$$x'(t) = f(t, x(t)).$$

На самом деле, уравнение (4.1) имеет не одно, а целое семейство решений с одним свободным параметром. Этот параметр определяется единственным образом, если задано начальное условие

$$x_0 = x(t_0).$$

Дифференциальное уравнение второго порядка имеет вид

$$x'' = f(t, x, x'). \quad (4.2)$$

Решение этого уравнения содержит уже два параметра, которые определяются из начального условия

$$x_0 = x(t_0), \quad x'_0 = x'(t_0).$$

Обычно при численном решении (4.2) вводят новые переменные $x_1 = x$ и $x_2 = x'$, и это уравнение приводят к системе уравнений

$$\begin{aligned} x'_1 &= x_2, & x_1(t_0) &= x_0, \\ x'_2 &= f(t, x_1, x_2), & x_2(t_0) &= x'_0. \end{aligned} \quad (4.3)$$

Следует заметить, что подобным способом можно привести любую систему, состоящую из уравнений второго порядка, к системе уравнений первого порядка. Очевидно, (4.3) является частным случаем системы уравнений первого порядка общего вида

$$\begin{aligned}x'_1 &= f_1(t, x_1, \dots, x_m), & x_1(t_0) &= x_{10}, \\ \dots & \\ x'_m &= f_m(t, x_1, \dots, x_m), & x_m(t_0) &= x_{m0};\end{aligned}$$

или

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad (4.4)$$

где $\mathbf{x} = (x_1, \dots, x_m)^T$, а $\mathbf{f} = (f_1, \dots, f_m)^T$. Дифференциальные уравнения (4.4) вместе со своими начальными условиями составляют задачу Коши.

В данной главе мы сосредоточим основное внимание именно на тех методах, которые были специально разработаны для численного интегрирования систем (4.4).

При этом будем полагать, что вектор-функция \mathbf{f} удовлетворяет условию Липшица, т.е. для любых решений \mathbf{x}_1 и \mathbf{x}_2 справедливо неравенство

$$\|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)\| \leq L \|\mathbf{x}_1 - \mathbf{x}_2\|, \quad (4.5)$$

где L — так называемая постоянная Липшица.

4.2. Метод разложения в ряд Тейлора

Одним из старейших методов решения дифференциальных уравнений является метод разложения в ряд Тейлора.

Обратимся к скалярному случаю. Пусть требуется найти решение уравнения (4.1)

$$x' = f(t, x)$$

для заданного значения независимой переменной $t_0 + \Delta t$ при начальном условии $x_0 = x(t_0)$. Потребуем, чтобы f была аналитична в точке (t_0, x_0) . Дифференцируя уравнение по t , будем иметь

$$x'' = f'_t + f'_x x', \quad x''' = f''_{tt} + 2f''_{tx} x' + f''_{xx} x'^2 + f'_x x'', \dots$$

Подставляя t_0 и x_0 , последовательно получаем

$$x'(t_0), \quad x''(t_0), \quad x'''(t_0), \dots \quad (4.6)$$

Таким образом, при достаточно малом Δt решение дифференциального уравнения можно представить в виде ряда Тейлора

$$x(t_0 + \Delta t) = x_0 + \sum_{i=1}^{\infty} \frac{x^{(i)}(t_0)}{i!} \Delta t^i. \quad (4.7)$$

Оборвем (усечем) ряд (4.7) на p -ом члене. Тогда получим приближенную формулу p -го порядка

$$x_1 = x_0 + \sum_{i=1}^p \frac{x^{(i)}(t_0)}{i!} \Delta t^i \approx x(t_0 + \Delta t) \quad (4.8)$$

с погрешностью

$$\Delta x = x(t_0 + \Delta t) - x_1 = \sum_{i=p+1}^{\infty} \frac{x^{(i)}(t_0)}{i!} \Delta t^i = O(\Delta t^{p+1}). \quad (4.9)$$

Нетрудно видеть, что при малых Δt погрешность Δx (ошибка усечения) будет определяться главным образом первым членом ряда (4.9):

$$\frac{x^{(p+1)}(t_0)}{(p+1)!} \Delta t^{p+1}. \quad (4.10)$$

Его называют главным членом погрешности.

Если ряд Тейлора (4.7) расходится или ошибка приближенной формулы (4.8) довольно большая, то прибегают к пошаговому интегрированию. Разбивают отрезок $[t_0, t_0 + \Delta t]$ на малые подотрезки $[t_n, t_{n+1}]$ ($n = 0, \dots, N-1$), где $t_N = t_0 + \Delta t$, и к каждому подотрезку последовательно применяют формулу (4.8):

$$x_{n+1} = x_n + \sum_{i=1}^p \frac{x^{(i)}(t_n)}{i!} h_{n+1}^i \quad (n = 0, \dots, N-1). \quad (4.11)$$

Здесь $h_{n+1} = t_{n+1} - t_n$ — шаг интегрирования. В результате искомого решения будет $x_N \approx x(t_N)$.

Рассмотрим частные случаи. Для $p = 1$ и $p = 2$ с постоянным шагом h имеем формулы

$$\begin{aligned} x_{n+1} &= x_n + hf(t_n, x_n) \quad (\text{метод Эйлера}), & (4.12) \\ x_{n+1} &= x_n + hf(t_n, x_n) + \frac{h^2}{2}[f'_t(t_n, x_n) + f'_x(t_n, x_n)f(t_n, x_n)]. \end{aligned}$$

Главным недостатком метода разложения в ряд Тейлора является то, что для получения производных (4.6) необходимо знать производные от f как функции t и x . Очевидно, если f — сложная функция, получение этих производных может представлять собой весьма утомительное занятие. В связи с этим методы разложения в ряд Тейлора редко используются на практике.

Впрочем, в небесной механике имеет место ряд важных с прикладной точки зрения задач, где, используя специфику дифференциальных уравнений, удается получить достаточно простые рекуррентные соотношения для временных производных в разложениях Тейлора. В 1956 г. К. Стеффенсен (Steffensen, 1956) путем введения новых вспомогательных переменных вывел подобные соотношения для интегрирования уравнений планетной задачи²².

Рассмотрим идею метода Стеффенсена на примере нормализованных уравнений задачи двух тел:

$$\mathbf{x}'' = -\frac{\mathbf{x}}{|\mathbf{x}|^3}. \quad (4.13)$$

Если ввести вспомогательные переменные

$$u = |\mathbf{x}|^{-3}, \quad w = |\mathbf{x}|^{-2}, \quad s = \mathbf{x} \cdot \dot{\mathbf{x}} \quad \text{и} \quad \sigma = ws,$$

²²Впрочем, по утверждению авторов (Hairer et al., 1993) этот прием был известен еще Ньютону. Но самое интересное, что и метод Стеффенсона был переоткрыт спустя несколько десятков лет Э. Паркером и Дж. Сочаки (Parker, Sochacki, 1996), и сейчас соответственно в тех или иных научных работах его порой можно встретить под названием переоткрывателей, т.е. как метод Паркера–Сочаки.

то уравнение (4.13) можно привести к дифференциально-алгебраической системе уравнений с квадратичными правыми частями

$$\mathbf{x}'' = -u\mathbf{x}, \quad u' = -3u\sigma, \quad w' = -2u\sigma, \quad s = \mathbf{x} \cdot \dot{\mathbf{x}}, \quad \sigma = ws. \quad (4.14)$$

Тогда, подставляя в (4.14) степенные ряды по малому параметру h для \mathbf{x} , u , w , s и σ и приравнявая коэффициенты при одинаковых степенях h , получим для них рекуррентные формулы

$$\mathbf{x}_{j+2} = -\frac{1}{(j+1)(j+2)} \sum_{i=0}^j \mathbf{x}_i u_{j-i}, \quad u_j = -\frac{3}{j} \sum_{i=0}^{j-1} u_i \sigma_{j-i-1},$$

$$w_j = -\frac{2}{j} \sum_{i=0}^{j-1} w_i \sigma_{j-i-1}, \quad s_j = \sum_{i=0}^j (i+1) \mathbf{x}_{i+1} \cdot \mathbf{x}_{j-1}, \quad \sigma_j = \sum_{i=0}^j w_i s_{j-1}.$$

При этом $\mathbf{x}_i = \mathbf{x}^{(i)}(t_0)/i!$ Коэффициенты нулевого порядка u_0 , w_0 , s_0 и σ_0 вычисляются по начальным данным \mathbf{x}_0 и $\dot{\mathbf{x}}_0$.

Подобный подход также применяется и для численного интегрирования более сложных дифференциальных уравнений небесной механики, в частности, задачи нескольких тел в классической постановке.

4.3. Методы Рунге–Кутты

На рубеже двух последних прошедших столетий сначала К. Рунге (Runge, 1895), а затем К. Хойн (Heun, 1900) и М.В. Кутта (Kutta, 1901) предложили подход, основанный на построении приближенной формулы, близкой к (4.8), не содержащей производных от правых частей дифференциальных уравнений.

4.3.1. Явные методы Рунге–Кутты

Рунге (Runge, 1895) продемонстрировал идею получения новых методов только для низких порядков, но именно Кутта (Kutta, 1901) сформулировал общую схему того, что теперь называется

методами Рунге–Кутты. Дадим определение явных методов применительно к системе (4.4).

Метод

$$\begin{aligned}
 \mathbf{k}_1 &= \mathbf{f}(t_0, \mathbf{x}_0), \\
 \mathbf{k}_2 &= \mathbf{f}(t_0 + hc_2, \mathbf{x}_0 + ha_{21}\mathbf{k}_1), \\
 \mathbf{k}_3 &= \mathbf{f}(t_0 + hc_3, \mathbf{x}_0 + h(a_{31}\mathbf{k}_1 + a_{32}\mathbf{k}_2)), \\
 &\dots \\
 \mathbf{k}_s &= \mathbf{f}(t_0 + hc_s, \mathbf{x}_0 + h(a_{s1}\mathbf{k}_1 + \dots + a_{s,s-1}\mathbf{k}_{s-1})), \\
 \mathbf{x}_1 &= \mathbf{x}_0 + h(b_1\mathbf{k}_1 + \dots + b_s\mathbf{k}_s),
 \end{aligned} \tag{4.15}$$

или

$$\begin{aligned}
 \mathbf{x}_1 &= \mathbf{x}_0 + h \sum_{j=1}^s b_j \mathbf{k}_j, \quad \mathbf{k}_1 = \mathbf{f}(t_0, \mathbf{x}_0), \\
 \mathbf{k}_i &= \mathbf{f}(t_0 + hc_i, \mathbf{x}_0 + h \sum_{j=1}^{i-1} a_{ij} \mathbf{k}_j) \quad (i = 2, \dots, s)
 \end{aligned}$$

называется s -этапным (s -стадийным) явным методом Рунге–Кутты для задачи (4.4). Явным метод называется потому, что в (4.15) каждая следующая величина \mathbf{k}_i явно выражается через предыдущие $\mathbf{k}_1, \dots, \mathbf{k}_{i-1}$. Метод также может быть переписан в эквивалентной форме

$$\begin{aligned}
 \mathbf{x}_1 &= \mathbf{x}_0 + hb_1\mathbf{f}(t_0, \mathbf{x}_0) + h \sum_{j=2}^s b_j \mathbf{f}(t_0 + hc_j, \mathbf{y}_j), \\
 \mathbf{y}_i &= \mathbf{x}_0 + h \sum_{j=1}^{i-1} a_{ij} \mathbf{f}(t_0 + hc_j, \mathbf{y}_j) \quad (i = 2, \dots, s).
 \end{aligned}$$

Коэффициенты a_{ij} , b_i и c_i определяются из так называемых условий порядка. Эти условия получаются из разложения приближенного решения (4.15) в ряд по степеням h и сравнения его с

разложением точного решения. При этом коэффициенты определяются таким образом, чтобы получить в разложениях как можно больше попарно одинаковых первых членов. Степень старшего из этих членов p задает порядок метода.

Например, разложения приближенного решения \mathbf{x}_1 двухэтапной схемы интегрирования и точного решения $\mathbf{x}(t_1)$ по малому параметру h дают соответственно

$$\mathbf{x}_1 = \mathbf{x}_0 + h(b_1 + b_2)\mathbf{f}(t_0, \mathbf{x}_0) + h^2[b_2c_2\mathbf{f}'_t(t_0, \mathbf{x}_0) + b_2a_{21}\mathbf{f}'_x(t_0, \mathbf{x}_0)\mathbf{f}(t_0, \mathbf{x}_0)] + O(h^3),$$

$$\mathbf{x}(t_1) = \mathbf{x}_0 + h\mathbf{f}(t_0, \mathbf{x}_0) + \frac{h^2}{2}[\mathbf{f}'_t(t_0, \mathbf{x}_0) + \mathbf{f}'_x(t_0, \mathbf{x}_0)\mathbf{f}(t_0, \mathbf{x}_0)] + O(h^3).$$

Тогда, сравнивая члены при одинаковых степенях h , получаем условия второго порядка

$$b_1 + b_2 = 1, \quad b_2c_2 = 1/2, \quad b_2a_{21} = 1/2. \quad (4.16)$$

Система уравнений (4.16) дает однопараметрическое семейство решений, которое можно представить в виде

$$b_1 = 1 - \alpha, \quad b_2 = \alpha, \quad a_{21} = c_2 = 1/(2\alpha),$$

где $\alpha \neq 0$ — свободный параметр. В частности, при $\alpha = 1/2$ имеем метод Хойна (явный метод трапеций)

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_0, \mathbf{x}_0), \\ \mathbf{k}_2 &= \mathbf{f}(t_0 + h, \mathbf{x}_0 + h\mathbf{k}_1), \\ \mathbf{x}_1 &= \mathbf{x}_0 + h(\mathbf{k}_1 + \mathbf{k}_2)/2, \end{aligned}$$

или

$$\mathbf{x}_1 = \mathbf{x}_0 + \frac{h}{2}[\mathbf{f}(t_0, \mathbf{x}_0) + \mathbf{f}(t_0 + h, \mathbf{x}_0 + h\mathbf{f}(t_0, \mathbf{x}_0))]. \quad (4.17)$$

Нетрудно вывести условия третьего порядка. Они будут

$$\sum_i b_i = 1, \quad 2 \sum_i b_i c_i = 1, \quad 3 \sum_i b_i c_i^2 = 1, \quad 6 \sum_{i,j} b_i a_{ij} c_j = 1.$$

Условия для более высоких порядков чрезвычайно громоздки, и поэтому мы не будем их рассматривать.

Очевидно, именно условия порядка задают связь между порядком метода p и числом этапов (числом обращения к функции правых частей) s . К сожалению, только до четвертого порядка $p = s$. Для методов более высокого порядка $p < s$.

Коэффициенты c_i обычно удовлетворяют условиям

$$c_2 = a_{21}, \quad c_3 = a_{31} + a_{32}, \quad \dots, \quad c_s = a_{s1} + \dots + a_{s,s-1},$$

или

$$c_i = \sum_j a_{ij}.$$

Кроме того, согласно определениям производной и метода Рунге–Кутты (4.15),

$$\mathbf{f}(t_0, \mathbf{x}_0) = \mathbf{x}'(t_0) = \lim_{h \rightarrow 0} \frac{\mathbf{x}_1 - \mathbf{x}_0}{h} = \mathbf{f}(t_0, \mathbf{x}_0) \sum_{j=1}^s b_j.$$

Следовательно,

$$\sum_j b_j = 1.$$

Метод Рунге–Кутты (4.15) имеет порядок p , если

$$\|\mathbf{x}(t_0 + h) - \mathbf{x}_1\| = O(h^{p+1}), \quad (4.18)$$

т.е. если ряды Тейлора для точного решения $\mathbf{x}(t_0 + h)$ и для приближенного \mathbf{x}_1 совпадают до члена h^p включительно.

Символически методы Рунге–Кутты (4.15) принято представлять в виде таблицы Бутчера

0					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
\vdots	\vdots	\vdots	\ddots		
c_s	a_{s1}	a_{s2}	\dots	$a_{s,s-1}$	
	b_1	b_2	\dots	b_{s-1}	b_s

В качестве примера рассмотрим явный 4-этапный классический метод Рунге–Кутты 4-го порядка:

$$\begin{aligned}
 \mathbf{k}_1 &= \mathbf{f}(t_0, \mathbf{x}_0), \\
 \mathbf{k}_2 &= \mathbf{f}(t_0 + h/2, \mathbf{x}_0 + h\mathbf{k}_1/2), \\
 \mathbf{k}_3 &= \mathbf{f}(t_0 + h/2, \mathbf{x}_0 + h\mathbf{k}_2/2), \\
 \mathbf{k}_4 &= \mathbf{f}(t_0 + h, \mathbf{x}_0 + h\mathbf{k}_3), \\
 \mathbf{x}_1 &= \mathbf{x}_0 + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4).
 \end{aligned}
 \tag{4.19}$$

Классический метод имеет табличный вид

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	2/6	2/6	1/6

4.3.2. Оценка погрешности и выбор шага

Локальная погрешность метода — это погрешность на одном шаге. Согласно (4.18), метод порядка p на любом n -ом шаге должен иметь локальную погрешность \mathbf{e}_n величины $\|\mathbf{e}_n\| = O(h_n^{p+1})$. Это означает, что существует такая постоянная C , что для любых h_n справедливо неравенство

$$\|\mathbf{e}_n\| \leq Ch_n^{p+1}. \tag{4.20}$$

Глобальной погрешностью называется погрешность численного решения после выполнения нескольких шагов.

Оценка глобальной погрешности. Одним из способов оценки глобальной погрешности является суммирование всех оценок для локальных погрешностей, перенесенных на конец интервала интегрирования (Hairer et al., 1993). Погрешность на некотором n -м шаге переносится посредством выполнения оставшихся $N - n$ шагов с использованием либо точных решений (рис. 4.1, а), либо приближенных (рис. 4.1, б).

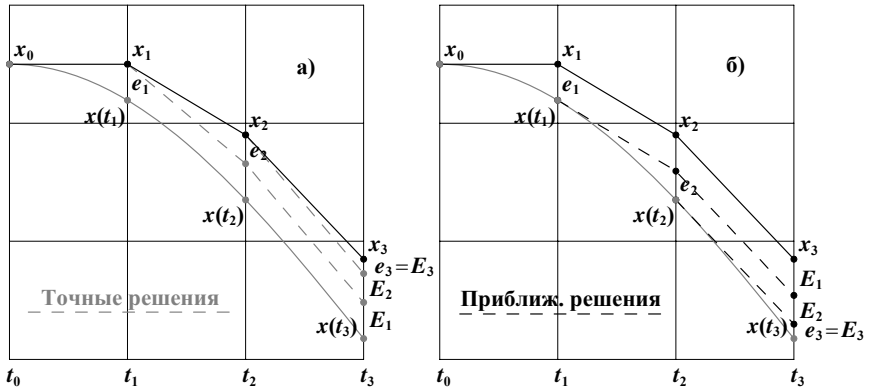


Рис. 4.1. Способы перенесения погрешностей

Пусть во втором способе переноса погрешностей (рис. 4.1, б) на некотором n -м шаге имеем два решения \mathbf{x}_n и \mathbf{z}_n , которые отличаются на локальную ошибку \mathbf{e}_n , т.е. $\mathbf{z}_n = \mathbf{x}_n + \mathbf{e}_n$. Для \mathbf{z}_n схема интегрирования запишется как

$$\begin{aligned} \mathbf{l}_1 &= \mathbf{f}(t_n, \mathbf{z}_n), \\ \mathbf{l}_2 &= \mathbf{f}(t_n + h_{n+1}c_2, \mathbf{z}_n + h_{n+1}a_{21}\mathbf{l}_1), \\ &\dots \\ \mathbf{z}_{n+1} &= \mathbf{z}_n + h_{n+1}(b_1\mathbf{l}_1 + \dots + b_s\mathbf{l}_s). \end{aligned}$$

Тогда для норм разностей будут справедливы неравенства

$$\begin{aligned} \|\mathbf{k}_1 - \mathbf{l}_1\| &\leq L\|\mathbf{x}_n - \mathbf{z}_n\|, \\ \|\mathbf{k}_2 - \mathbf{l}_2\| &\leq L(1 + |a_{21}|h_{n+1}L)\|\mathbf{x}_n - \mathbf{z}_n\|, \\ &\dots \\ \|\mathbf{x}_{n+1} - \mathbf{z}_{n+1}\| &\leq \|\mathbf{x}_n - \mathbf{z}_n\| + \\ &+ h_{n+1}(|b_1|\|\mathbf{k}_1 - \mathbf{l}_1\| + \dots + |b_s|\|\mathbf{k}_s - \mathbf{l}_s\|), \end{aligned} \quad (4.21)$$

где $L \geq 0$ — постоянная Липшица для \mathbf{f} . Если ввести постоянную

$$\Lambda = L \left(\sum_i |b_i| + hL \sum_{i,j} |b_i a_{ij}| + h^2 L^2 \sum_{i,j,k} |b_i a_{ij} a_{jk}| + \dots \right) \geq 0,$$

где $h = \max_{n=1, \dots, N} h_n$, последнее неравенство (4.21) можно переписать как

$$\|\mathbf{x}_{n+1} - \mathbf{z}_{n+1}\| \leq (1 + h_{n+1}\Lambda)\|\mathbf{x}_n - \mathbf{z}_n\| \leq \exp(h_{n+1}\Lambda)\|\mathbf{x}_n - \mathbf{z}_n\|.$$

Тогда в соответствии с (4.20) оценка для перенесенной погрешности на n -ом шаге будет

$$\|\mathbf{E}_n\| \leq \exp((t_N - t_n)\Lambda)\|\mathbf{e}_n\| \leq \exp((t_N - t_n)\Lambda)Ch^p h_n,$$

а оценка для глобальной погрешности

$$\begin{aligned} \|\mathbf{E}\| &\leq \sum_{n=1}^N \|\mathbf{E}_n\| \leq Ch^p \sum_{n=1}^N \exp((t_N - t_n)\Lambda)h_n \leq \\ &\leq Ch^p \int_{t_0}^{t_N} \exp((t_N - t)\Lambda)dt. \end{aligned}$$

Наконец, получаем

$$\|\mathbf{E}\| \leq h^p \frac{C}{\Lambda} (\exp((t_N - t_0)\Lambda) - 1) \quad \text{или} \quad \|\mathbf{E}\| = O(h^p). \quad (4.22)$$

Отсюда следует, что все явные методы Рунге–Кутты сходящиеся, т.е.

$$\lim_{h \rightarrow 0} \mathbf{x}_N = \mathbf{x}(t_N).$$

Оценку, подобную (4.22), для глобальной погрешности можно также получить и для неявных методов Рунге–Кутты, которые мы рассмотрим позже. Поэтому и все неявные методы — сходящиеся.

Практическая оценка локальной погрешности. Экстраполяция численного решения. Для практики оценка локальной погрешности необходима с целью выбора оптимального шага интегрирования: с одной стороны, достаточно малого для обеспечения

требуемой точности, но с другой — достаточно большого во избежание бесполезной вычислительной работы.

Локальная погрешность обычно оценивается путем сравнения двух численных решений: либо одного порядка по формулам Рунге (см. ниже) или Милна (подразд. 4.5.8), либо разных порядков, как во вложенных (подразд. 4.3.3) и экстраполяциянных (разд. 4.4) методах, а также в многошаговых методах Адамса–Мультона (подразд. 4.5.9). Рассмотрим пока один из способов, который использовал еще Рунге для оценки локальной точности.

Предположим, что мы используем некоторый метод Рунге–Кутты порядка p . Обозначим $\mathbf{x}_{n(h)}$ n -ое приближенное решение на сетке с постоянной величиной шага h . Разобьем временной отрезок $[t_0, t_0 + h]$ на N подотрезков и выполним пошаговое интегрирование с постоянным шагом h/N . В итоге получим приближенное решение $\mathbf{x}_{N(h/N)}$.

Допустим теперь, что некоторый n -й шаг выполняется с локальной ошибкой \mathbf{e}_n . Поскольку метод интегрирования имеет порядок p , то разложение локальной погрешности представимо в виде

$$\mathbf{e}_n = \mathbf{C}(t_{n-1})(h/N)^{p+1} + O(h^{p+2}) = \mathbf{C}(t_0)(h/N)^{p+1} + O(h^{p+2}), \quad (4.23)$$

где \mathbf{C} выражается через производные $(p+1)$ -порядка от \mathbf{x} , вычисленные на соответствующий момент независимой переменной t .

Оценим перенесенную погрешность \mathbf{E}_n на конечный момент $t_N = t_0 + h$, используя точные решения, т.е. способом, представленным на рис. 4.1, а. Рассмотрим на момент $t_n = t_0 + hn/N$ два решения $\mathbf{x}(t_n)$ и $\mathbf{z}(t_n) = \mathbf{x}(t_n) + \mathbf{e}_n$, отличающиеся на локальную ошибку. Разложение в ряд Тейлора в окрестности первого решения дает

$$\mathbf{x}(t_N) = \mathbf{x}(t_n) + \sum_{i=1}^{\infty} \frac{\mathbf{f}^{(i-1)}(t_n, \mathbf{x}(t_n))}{i!} (t_N - t_n)^i,$$

тогда как в окрестности второго —

$$\begin{aligned} \mathbf{z}(t_N) &= \mathbf{z}(t_n) + \sum_{i=1}^{\infty} \frac{\mathbf{f}^{(i-1)}(t_n, \mathbf{z}(t_n))}{i!} (t_N - t_n)^i = \\ &= \mathbf{z}(t_n) + \sum_{i=1}^{\infty} \frac{\mathbf{f}^{(i-1)}(t_n, \mathbf{x}(t_n))}{i!} (t_N - t_n)^i + O(h^{p+2}). \end{aligned} \quad (4.24)$$

Здесь использованы оценки $\mathbf{e}_n = O(h^{p+1})$ и $t_N - t_n = O(h)$. Следовательно,

$$\mathbf{E}_n = \mathbf{z}(t_N) - \mathbf{x}(t_N) = \mathbf{z}(t_n) - \mathbf{x}(t_n) + O(h^{p+2}) = \mathbf{e}_n + O(h^{p+2}).$$

Таким образом, для глобальной погрешности приближенного решения $\mathbf{x}_{N(h/N)}$ будем иметь

$$\mathbf{x}(t_0 + h) - \mathbf{x}_{N(h/N)} = \mathbf{E} = \sum_{n=1}^N \mathbf{E}_n = \sum_{n=1}^N \mathbf{e}_n + O(h^{p+2}) \quad (4.25)$$

или, учитывая (4.23),

$$\mathbf{x}(t_0 + h) - \mathbf{x}_{N(h/N)} = N\mathbf{C}(t_0)(h/N)^{p+1} + O(h^{p+2}). \quad (4.26)$$

Используя формулу (4.26), можно практически оценить локальную погрешность метода интегрирования. Действительно, согласно (4.26), после выполнения одного шага величины h погрешность приближенного решения \mathbf{e}_1 представима в виде ($N = 1$)

$$\mathbf{e}_1 = \mathbf{x}(t_0 + h) - \mathbf{x}_{1(h)} = \mathbf{C}(t_0)h^{p+1} + O(h^{p+2}). \quad (4.27)$$

Чтобы оценить неизвестную векторную величину $\mathbf{C}(t_0)$, уменьшим величину шага вдвое и выполним повторное интегрирование с двумя шагами ($N = 2$). Согласно (4.26), будем иметь

$$\mathbf{x}(t_0 + h) - \mathbf{x}_{2(h/2)} = 2\mathbf{C}(t_0)(h/2)^{p+1} + O(h^{p+2}). \quad (4.28)$$

Из (4.27) и (4.28) получаем константу $\mathbf{C}(t_0)$:

$$\mathbf{C}(t_0) = \frac{1}{h^{p+1}} \frac{\mathbf{x}_{2(h/2)} - \mathbf{x}_{1(h)}}{1 - (1/2)^p} + O(h);$$

а затем и погрешность \mathbf{e}_1 :

$$\mathbf{e}_1 = \mathbf{x}(t_0 + h) - \mathbf{x}_{1(h)} = \frac{\mathbf{x}_{2(h/2)} - \mathbf{x}_{1(h)}}{1 - (1/2)^p} + O(h^{p+2}). \quad (4.29)$$

Нетрудно видеть, что формула (4.29) позволяет по $\mathbf{x}_{2(h/2)}$ и $\mathbf{x}_{1(h)}$ получить экстраполяцию численного решения

$$\hat{\mathbf{x}}_1 = \mathbf{x}_{1(h)} + \frac{\mathbf{x}_{2(h/2)} - \mathbf{x}_{1(h)}}{1 - (1/2)^p}, \quad (4.30)$$

лучше представляющую точное решение $\mathbf{x}(t_0 + h)$. Действительно, согласно (4.29),

$$\mathbf{x}(t_0 + h) - \hat{\mathbf{x}}_1 = O(h^{p+2}),$$

т.е. решение $\hat{\mathbf{x}}_1$ (4.30) аппроксимирует точное с порядком $p + 1$, когда $\mathbf{x}_{2(h/2)}$ и $\mathbf{x}_{1(h)}$ — с порядком p .

Таким образом, формула (4.29) дает простой способ оценки погрешности, тогда как (4.30) позволяет дополнительно повысить точность численного интегрирования на один порядок.

Выведем теперь общую формулу экстраполяции. Рассмотрим два приближенных решения $\mathbf{x}_{N(h/N)}$ и $\mathbf{x}_{M(h/M)}$, аппроксимирующие точное решение $\mathbf{x}(t_0 + h)$. Тогда в соответствии с формулой (4.26) экстраполированное решение порядка $p + 1$ можно представить в виде

$$\hat{\mathbf{x}}_1 = \mathbf{x}_{N(h/N)} + \frac{\mathbf{x}_{M(h/M)} - \mathbf{x}_{N(h/N)}}{1 - (N/M)^p} \quad (\text{экстраполяция Ричардсона}). \quad (4.31)$$

Выбор шага. Возможность использования переменного шага необходима для обеспечения заданной величины локальной погрешности, что весьма важно при сложном поведении функции \mathbf{f} . Как правило, прибегают к следующему алгоритму выбора шага.

Пусть $\|\mathbf{e}\|_{tol} = \text{const}$ — допустимая величина локальной погрешности. Нам необходимо подобрать такой шаг интегрирования \bar{h} , который бы обеспечивал заданную величину локальной погрешности для некоторого метода Рунге–Кутты порядка p .

Согласно правилу Рунге, выполняем интегрирование на шаге h и на двух шагах $h/2$. В результате получаем два решения $\mathbf{x}_1(h)$ и $\mathbf{x}_2(h/2)$, по которым затем вычисляем величину локальной погрешности на шаге h , используя главный член оценки (4.29):

$$\|\mathbf{e}\|_{cal} = \frac{\|\mathbf{x}_2(h/2) - \mathbf{x}_1(h)\|}{1 - (1/2)^p}. \quad (4.32)$$

Тогда новый шаг \bar{h} определяем по формуле

$$\bar{h} = h \left(\frac{\|\mathbf{e}\|_{tol}}{\|\mathbf{e}\|_{cal}} \right)^{\frac{1}{p+1}}. \quad (4.33)$$

Она фактически получается из оценок

$$\|\mathbf{e}\|_{tol} = \|\mathbf{C}\|\bar{h}^{p+1} \quad \text{и} \quad \|\mathbf{e}\|_{cal} = \|\mathbf{C}\|h^{p+1}. \quad (4.34)$$

Далее обычно применяют экстраполяцию (4.30), с тем чтобы дополнительно повысить точность численного решения.

Следует заметить, что если поведение решения довольно гладкое, то оценку (4.33) можно применять не для текущего, а для следующего шага, т.е.

$$h_{n+1} = h_n \left(\frac{\|\mathbf{e}\|_{tol}}{\|\mathbf{e}\|_{cal}} \right)^{\frac{1}{p+1}}. \quad (4.35)$$

Формула (4.35) вполне приемлема, если компоненты вектора \mathbf{x} имеют одинаковую размерность и все они должны интегрироваться с одинаковой допустимой точностью $\|\mathbf{e}\|_{tol}$. Однако, если компоненты разной размерности, величина $\|\mathbf{e}\|_{tol}$ может быть допустимой только для одной группы компонент, тогда как для другой нет. В этом случае желательно для каждой компоненты x_i учитывать свою собственную задаваемую точность $(e_i)_{tol}$, и шаг выбирать таким образом, чтобы оцениваемая точность ни для одной из компонент не превышала задаваемую. Такая стратегия выбора шага возможна с использованием следующей формулы

$$h_{n+1} = h_n \min_{i=1, \dots, m} \left| \frac{(e_i)_{tol}}{(e_i)_{cal}} \right|^{\frac{1}{p+1}}. \quad (4.36)$$

Однако, на наш взгляд, наиболее естественным было бы обратиться к обобщенной формуле (4.35)

$$h_{n+1} = h_n \left(\sum_{i=1}^m \frac{(e_i)_{cal}^2}{(e_i)_{tol}^2} \right)^{-\frac{1}{2(p+1)}}. \quad (4.37)$$

Геометрический смысл (4.37) состоит в том, что шаг интегрирования выбирается такой величины, чтобы точка в m -мерном пространстве \mathbf{e} с координатами $(e_1)_{cal}, \dots, (e_m)_{cal}$ лежала на гиперэллипсоиде с полуосями $(e_1)_{tol}, \dots, (e_m)_{tol}$. Следовательно, ни одна из оценок точности для соответствующей компоненты вектора \mathbf{x} не будет превышать свою допустимую величину. Формула (4.35), между тем, является частным случаем (4.37), когда все полуоси равны $\|\mathbf{e}\|_{tol}$. В отличие же от (4.37), формула (4.36) удерживает точку пространства \mathbf{e} на прямоугольном гиперпараллелепипеде объема $2(e_1)_{tol} \times \dots \times 2(e_m)_{tol}$.

Во избежание значительного изменения оценочной величины $\|\mathbf{e}\|_{cal}$ на выбор шага накладываются превентивные ограничения. Допустим, r — отношение величин следующего и текущего шагов, т.е. $h_{n+1} = h_n r$. Тогда, например, чтобы ограничить изменение $\|\mathbf{e}\|_{cal}$ в пределах одного порядка, значение r в соответствии с (4.34) должно удовлетворять неравенствам

$$\frac{1}{\sigma} < r^{p+1} < \sigma, \quad \text{где } \sigma = \sqrt{10}. \quad (4.38)$$

Впрочем, на практике соблюдать условие (4.38) целесообразно только в начале интегрирования, при подборе первого шага h_1 . Далее можно учитывать только верхний предел: если $r^{p+1} > \sigma$, принимается $r = \sigma^{1/(p+1)}$. Кстати, такой алгоритм реализован в интеграторе Гаусса–Эверхарта (Авдюшев, 2010b), который будет рассмотрен в подразд. 4.3.7.

Разумеется, представленные формулы (4.35)–(4.38) для выбора шага применимы также и в тех случаях, когда локальная точность $\|\mathbf{e}\|_{cal}$ оценивается другими способами, не только по правилу Рунге (4.32).

4.3.3. Вложенные методы Рунге–Кутты

Вложенные методы Рунге–Кутты помимо основного решения \mathbf{x}_1 порядка p содержат также еще одно вспомогательное (вложенное) решение $\hat{\mathbf{x}}_1$ иного порядка q . Если $q > p$, по разности приближенных решений оценивается локальная ошибка $\|\mathbf{e}\|_{cal}$ основного решения; иначе ($q < p$) — вспомогательного. Действительно, допустим $q > p$, тогда

$$\mathbf{x}_1 = \mathbf{x}(t_1) + \mathbf{C}h^{p+1} + O(h^{p+2}), \quad \hat{\mathbf{x}}_1 = \mathbf{x}(t_1) + O(h^{q+1}).$$

Отсюда

$$\mathbf{x}_1 - \hat{\mathbf{x}}_1 = \mathbf{C}h^{p+1} + O(h^{p+2}) = \mathbf{x}_1 - \mathbf{x}(t_1) + O(h^{p+2}).$$

Таким образом, разность приближенных решений представляет локальную ошибку основного решения с точностью до малых порядка h^{p+2} , т.е. достаточно хорошо, поскольку сама ошибка — величина порядка h^{p+1} .

Как правило, $q = p + 1$ либо $q = p - 1$. Несмотря на то что при $q < p$ оценивается и обеспечивается точность не основного решения, а вспомогательного, для выбора шага оценка $\|\mathbf{e}\|_{cal}$ может быть вполне приемлема: по крайней мере, реальная точность основного решения будет не хуже задаваемой $\|\mathbf{e}\|_{tol}$.

Итак, если приближенные решения

$$\mathbf{x}_1 = \mathbf{x}_0 + h(b_1\mathbf{k}_1 + \dots + b_s\mathbf{k}_s) \quad \text{и} \quad \hat{\mathbf{x}}_1 = \mathbf{x}_0 + h(\hat{b}_1\mathbf{k}_1 + \dots + \hat{b}_s\mathbf{k}_s)$$

получены, то локальная ошибка (основного решения либо вспомогательного) оценивается как

$$\|\mathbf{e}\|_{cal} = \|\mathbf{x}_1 - \hat{\mathbf{x}}_1\| = h\|(b_1 - \hat{b}_1)\mathbf{k}_1 + \dots + (b_s - \hat{b}_s)\mathbf{k}_s\|.$$

Вложенные методы называют по фамилии автора с указанием порядков основного и вспомогательного решений p (q). На практике широко используются вложенные методы Мерсона, Инглэнда, Фельдберга, а также Дормана–Принса (Hairer et al., 1993).

Вложенные методы Рунге–Кутты имеют табличный вид

0					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
\vdots	\vdots	\vdots	\ddots		
c_s	a_{s1}	a_{s2}	\dots	$a_{s,s-1}$	
\mathbf{x}_1	b_1	b_2	\dots	b_{s-1}	b_s
$\hat{\mathbf{x}}_1$	\hat{b}_1	\hat{b}_2	\dots	\hat{b}_{s-1}	\hat{b}_s

В качестве примера приведем схему Мерсона 4 (5)²³:

0					
1/3	1/3				
1/3	1/6	1/6			
1/2	1/8	0	3/8		
1	1/2	0	-3/2	2	
\mathbf{x}_1	1/2	0	-3/2	2	0
$\hat{\mathbf{x}}_1$	1/6	0	0	2/3	1/6

Интересно, что классический метод Рунге–Кутты (4.19) можно также рассматривать как вложенный, если переписать его таблицу в виде

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
\mathbf{x}_1	1/6	2/6	2/6	1/6
$\hat{\mathbf{x}}_1$	0	0	1	0

Формально получаем метод Рунге–Кутты 4 (2) (Холл, Уатт, 1979).

²³Впрочем, следует заметить, что в методе Мерсона вспомогательное решение имеет пятый порядок только для линейных уравнений с постоянными коэффициентами (Hairer et al., 1993).

4.3.4. Неявные методы Рунге–Кутты

Несмотря на то что неявные методы Рунге–Кутты реализуются сложнее, нежели явные, на практике к ним прибегают гораздо чаще. Очевидная причина тому — это прежде всего практическая возможность конструировать схемы интегрирования ультравысоких порядков ($p > 10$), чего не могут позволить явные методы.

Метод

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_0 + hc_1, \mathbf{x}_0 + h(a_{11}\mathbf{k}_1 + \dots + a_{1s}\mathbf{k}_s)), \\ &\dots \\ \mathbf{k}_s &= \mathbf{f}(t_0 + hc_s, \mathbf{x}_0 + h(a_{s1}\mathbf{k}_1 + \dots + a_{ss}\mathbf{k}_s)), \\ \mathbf{x}_1 &= \mathbf{x}_0 + h(b_1\mathbf{k}_1 + \dots + b_s\mathbf{k}_s), \end{aligned} \quad (4.39)$$

или

$$\mathbf{x}_1 = \mathbf{x}_0 + h \sum_{j=1}^s b_j \mathbf{k}_j, \quad \mathbf{k}_i = \mathbf{f}(t_0 + hc_i, \mathbf{x}_0 + h \sum_{j=1}^s a_{ij} \mathbf{k}_j) \quad (i = 1, \dots, s),$$

называется s -этапным (s -стадийным) неявным методом Рунге–Кутты для задачи (4.4). Как и в явном случае, его можно переписать в эквивалентной форме

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0 + h \sum_{j=1}^s b_j \mathbf{f}(t_0 + hc_j, \mathbf{y}_j), \\ \mathbf{y}_i &= \mathbf{x}_0 + h \sum_{j=1}^s a_{ij} \mathbf{f}(t_0 + hc_j, \mathbf{y}_j) \quad (i = 1, \dots, s). \end{aligned}$$

Простые и хорошо известные примеры неявных методов Рунге–Кутты — это неявный метод Эйлера (первого порядка)

$$\mathbf{x}_1 = \mathbf{x}_0 + h\mathbf{f}(t_1, \mathbf{x}_1), \quad (4.40)$$

методы трапеций и средней точки (второго порядка):

$$\mathbf{x}_1 = \mathbf{x}_0 + \frac{h}{2} [\mathbf{f}(t_0, \mathbf{x}_0) + \mathbf{f}(t_1, \mathbf{x}_1)], \quad (4.41)$$

$$\mathbf{x}_1 = \mathbf{x}_0 + hf \left(t_0 + \frac{h}{2}, \frac{\mathbf{x}_0 + \mathbf{x}_1}{2} \right). \quad (4.42)$$

В соответствии с (4.39) неявные методы будут иметь табличный вид

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array}$$

Главным недостатком неявных методов Рунге–Кутты является то, что величины \mathbf{k}_i задаются в (4.39) неявным образом, и поэтому их нужно находить итерационным способом. Для этого, как правило, используют метод простых итераций:

$$\begin{aligned} \mathbf{k}_1^{l+1} &= \mathbf{f}(t_0 + hc_1, \mathbf{x}_0 + h(a_{11}\mathbf{k}_1^l + \dots + a_{1s}\mathbf{k}_s^l)), \\ &\dots \\ \mathbf{k}_s^{l+1} &= \mathbf{f}(t_0 + hc_s, \mathbf{x}_0 + h(a_{s1}\mathbf{k}_1^l + \dots + a_{ss}\mathbf{k}_s^l)). \end{aligned} \quad (4.43)$$

Введем матрицу $\mathbf{K} = (\mathbf{k}_1, \dots, \mathbf{k}_s)$ и норму $\|\mathbf{K}\| = \max_i \|\mathbf{k}_i\|$. Тогда (4.43) можно записать в виде $\mathbf{K}^{l+1} = \mathbf{F}(\mathbf{K}^l)$, где $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_s)$, а $\mathbf{f}_i(\mathbf{K}) = \mathbf{f}(t_0 + hc_i, \mathbf{x}_0 + h(a_{i1}\mathbf{k}_1 + \dots + a_{is}\mathbf{k}_s))$ ($i = 1, \dots, s$). Из условия Липшица (4.5) для двух произвольных \mathbf{K}_1 и \mathbf{K}_2 будет справедливо неравенство

$$\|\mathbf{F}(\mathbf{K}_1) - \mathbf{F}(\mathbf{K}_2)\| \leq hL \max_i \sum_j |a_{ij}| \|\mathbf{K}_1 - \mathbf{K}_2\|. \quad (4.44)$$

Для сходимости метода простых итераций отображение \mathbf{F} должно быть сжимающим, что, согласно (4.44), оказывается возможным при условии, когда

$$h < \frac{1}{L \max_i \sum_j |a_{ij}|}. \quad (4.45)$$

Очевидно, для нахождения \mathbf{k}_i необходимо, по меньшей мере, две итерации. Отсюда все неявные s -этапные методы требуют вычислений функций \mathbf{f} на шаге не меньше чем $2s$. (В данном случае

число s и число вычислений функций \mathbf{f} не одно и то же, как в явных методах.) Таким образом, неявные методы должны работать медленнее, нежели явные с тем же числом этапов s .

Несмотря на это, привлекательность неявных методов состоит в том, что при всех s существуют такие методы, которые имеют порядок $p = 2s$. Ж. Кунцман (Kuntzmann, 1961) и Дж. Бутчер (Butcher, 1964) показали, что такой порядок p достигается путем специального выбора коэффициентов c_i .

Рассмотрим два метода 8-го порядка: явный Дормана–Принса ($s = 13$) и неявный Кунцмана–Бутчера ($s = 4$) (Hairer et al., 1993). На практике начальные приближения \mathbf{k}_i^0 в неявных методах получают из соответствующих величин на предыдущем шаге. Тогда при хорошем выборе начальных приближений для сходимости итерационного процесса может потребоваться всего 2–3 итерации. Поэтому можно считать, что метод Кунцмана–Бутчера будет вычислять функцию \mathbf{f} 8–12 раз на шаге, тогда как метод Дормана–Принса — 13. Следовательно, благодаря своему замечательному свойству, неявный метод будет работать быстрее. Кроме того, необходимо заметить, что с повышением порядка p разность $s - p$ для явных методов увеличивается, поэтому указанное преимущество неявных методов будет только возрастать.

4.3.5. Коллокационные методы

Оказывается, что многие неявные методы Рунге–Кутты некоторым образом эквивалентны так называемым коллокационным методам (Hairer et al., 1993). Основная идея коллокационных методов для решения обыкновенных дифференциальных уравнений состоит в том, чтобы в качестве приближенного решения \mathbf{x} подобрать такой полином \mathbf{g} степени s , который бы удовлетворял условиям задачи Коши (4.4) в $s + 1$ точках:

$$\mathbf{g}(t_0) = \mathbf{x}_0, \quad \mathbf{g}'(t_0 + hc_i) = \mathbf{f}(t_0 + hc_i, \mathbf{g}(t_0 + hc_i)) \quad (i = 1, \dots, s), \quad (4.46)$$

где c_i — вещественные числа, выбираемые на отрезке $[0, 1]$. В соответствии с (4.46) приближенное решение уравнения на шаге h

представляется в виде $\mathbf{x}_1 = \mathbf{g}(t_0 + h)$. Например, для $s = 1$ полином должен иметь вид $\mathbf{g}(t) = \mathbf{x}_0 + (t - t_0)\mathbf{k}$ с линейным коэффициентом

$$\mathbf{k} = \mathbf{g}'(t_0 + c_1 h) = \mathbf{f}(t_0 + c_1 h, \mathbf{g}(t_0 + c_1 h)) = \mathbf{f}(t_0 + c_1 h, \mathbf{x}_0 + c_1 h \mathbf{k}).$$

Отсюда видно, что явный и неявный методы Эйлера, а также метод средней точки являются коллокационными с коэффициентами $c_1 = 0$, $c_1 = 1$ и $c_1 = 1/2$ соответственно.

Введем обозначение

$$\mathbf{k}_i = \mathbf{f}(t_0 + hc_i, \mathbf{g}(t_0 + hc_i)). \quad (4.47)$$

Тогда интерполирующая функция Лагранжа для $\mathbf{g}'(t)$ будет

$$\mathbf{g}'(t_0 + h\tau) = \sum_{j=1}^s \mathbf{k}_j l_j(\tau), \quad \text{где} \quad l_j(\tau) = \prod_{\substack{m=1 \\ m \neq j}}^s \frac{\tau - c_m}{c_j - c_m}, \quad (4.48)$$

а $\tau = (t - t_0)/h$. Отсюда, интегрируя (4.48) по τ и пользуясь записью (4.47), будем иметь²⁴

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0 + h \sum_{j=1}^s \mathbf{k}_j \int_0^1 l_j(\tau) d\tau = \mathbf{g}(t_0 + h), \quad \mathbf{k}_i = \mathbf{f}(t_0 + hc_i, \mathbf{y}_i), \\ \mathbf{y}_i &= \mathbf{x}_0 + h \sum_{j=1}^s \mathbf{k}_j \int_0^{c_i} l_j(\tau) d\tau = \mathbf{g}(t_0 + hc_i) \quad (i = 1, \dots, s). \end{aligned} \quad (4.49)$$

Нетрудно видеть, что формулы (4.49) совпадают с формулами неявного метода Рунге–Кутты (4.39), где в качестве коэффициентов a_{ij} и b_i выступают

$$a_{ij} = \int_0^{c_i} l_j(\tau) d\tau, \quad b_j = \int_0^1 l_j(\tau) d\tau \quad (i, j = 1, \dots, s). \quad (4.50)$$

²⁴Здесь можно обнаружить сходство коллокационных методов и квадратурных формул Ньютона–Котеса, применяемых для вычисления определенных интегралов.

Методы Рунге–Кутты с коэффициентами (4.50) называются коллокационными.

Далее, поскольку

$$\tau^{q-1} = \sum_{j=1}^s c_j^{q-1} l_j(\tau) \quad (q = 1, \dots, s)$$

(т.е. интерполяция Лагранжа точна для всех полиномов ниже s -го порядка), то коэффициенты (4.50) должны удовлетворять системам линейных уравнений

$$\sum_{j=1}^s a_{ij} c_j^{q-1} = \frac{c_i^q}{q}, \quad \sum_{j=1}^s b_j c_j^{q-1} = \frac{1}{q} \quad (i, q = 1, \dots, s). \quad (4.51)$$

Отметим, что коллокационные формулы очень удобны для определения приближенных значений \mathbf{k} следующего шага. Обозначая их как $\bar{\mathbf{k}}$, а величину следующего шага \bar{h} , согласно формулам (4.48), получаем экстраполированную оценку

$$\bar{\mathbf{k}}_i = \sum_{j=1}^s \mathbf{k}_j l_j(1 + c_i \bar{h}/h) \quad (i = 1, \dots, s).$$

Покажем, что коллокационный метод имеет, по крайней мере, порядок $p = s$ (Hairer et al., 1993). Согласно (4.47) и (4.48), коллокационный многочлен удовлетворяет уравнению

$$\mathbf{g}'(t_0 + h\tau) = \sum_{i=1}^s \mathbf{f}(t_0 + hc_i, \mathbf{g}(t_0 + hc_i)) l_i(\tau),$$

тогда как точное решение —

$$\mathbf{x}'(t_0 + h\tau) = \sum_{i=1}^s \mathbf{f}(t_0 + hc_i, \mathbf{x}(t_0 + hc_i)) l_i(\tau) + \mathbf{r}(\tau), \quad (4.52)$$

где $\mathbf{r}(\tau)$ — ошибка интерполяционной формулы Лагранжа, которая, как известно, может быть представлена в виде

$$\mathbf{r}(\tau) = h^s \frac{\mathbf{x}^{(s+1)}(t_0 + h\zeta)}{s!} \prod_{i=1}^s (\tau - c_i) = O(h^s), \quad \zeta = \zeta(\tau) \in [0, 1].$$

Интегрируя уравнения и вычитая их решения, будем иметь

$$\begin{aligned} \mathbf{x}(t_0 + h\tau) - \mathbf{g}(t_0 + h\tau) &= h \sum_{i=1}^s [\mathbf{f}(t_0 + hc_i, \mathbf{x}(t_0 + hc_i)) - \\ &\quad - \mathbf{f}(t_0 + hc_i, \mathbf{g}(t_0 + hc_i))] \int_0^\tau l_i(\xi) d\xi + O(h^{s+1}). \end{aligned} \quad (4.53)$$

Введем величину

$$B = \max_{\tau \in [0,1]} \sum_{i=1}^s \left| \int_0^\tau l_i(\xi) d\xi \right|.$$

Тогда, используя условие Липшица (4.5), для величины

$$M = \max_{\tau \in [0,1]} \|\mathbf{x}(t_0 + h\tau) - \mathbf{g}(t_0 + h\tau)\|$$

получаем оценку

$$M \leq hLBM + O(h^{s+1}) \quad \text{или} \quad M \leq O(h^{s+1}).$$

Она фактически говорит о том, что для любого $\tau \in [0, 1]$

$$\mathbf{x}(t_0 + h\tau) - \mathbf{g}(t_0 + h\tau) = O(h^{s+1}) \quad (4.54)$$

и, в частности, $\mathbf{x}(t_0 + h) - \mathbf{x}_1 = O(h^{s+1})$.

Итак, произвольное разбиение c_1, \dots, c_s дает коллокационный метод порядка $p = s$. Оставшиеся коэффициенты метода могут быть получены в результате решения систем линейных уравнений (4.51) относительно a_{ij} и b_j . Однако с практической точки зрения их удобнее вычислять из интегралов (4.50), используя простые рекуррентные формулы (Холл, Уатт, 1979).

Для каждого коэффициента a_{ij} имеем

$$\begin{aligned} \gamma_{0l} &= c_i^l / l! \quad (l = 1, \dots, s + 1), \\ \gamma_{kl} &= ((c_i - c_k)\gamma_{k-1,l} - l\gamma_{k-1,l+1}) / (c_j - c_k) \\ &\quad (k = 1, \dots, j - 1, j + 1, \dots, s; l = 1, \dots, s - k + 1), \end{aligned} \quad (4.55)$$

причем $\gamma_{j,l} = \gamma_{j-1,l}$ ($l = 1, \dots, s - j + 1$). Тогда $a_{ij} = \gamma_{s1}$. Для b_j используются те же формулы (4.55), но с заменой c_i на единицу.

4.3.6. Методы Гаусса

Как было показано, в общем случае метод (4.49) имеет порядок $p = s$. Однако если коэффициенты b_j и c_j удовлетворяют первым $2s$ уравнениям во второй системе (4.51) ($q = 1, \dots, 2s$) (т.е. если они являются соответственно весами и узлами квадратурной формулы Гаусса), то коллокационный метод будет иметь порядок $p = 2s$ (Butcher, 1964). Покажем это (Hairer et al., 2002).

Несмотря на то что коллокационный многочлен \mathbf{g} приближенно представляет точное решение \mathbf{x} , он является точным решением некоторого возмущенного уравнения

$$\mathbf{g}' = \mathbf{f}(t, \mathbf{g}) + \mathbf{p}(t).$$

Согласно (4.54), после линеаризации систему уравнений для возмущений можно представить в виде

$$\mathbf{g}' - \mathbf{x}' = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}(t)} (\mathbf{g} - \mathbf{x}) + \mathbf{p}(t) + O(h^{2s+2}).$$

Рассматривая ее как систему неоднородных линейных уравнений, применим для ее решения метод вариации произвольных постоянных:

$$\mathbf{x}_1 - \mathbf{x}(t_0 + h) = h \int_0^1 \Phi(t_0 + h) \Phi^{-1}(t_0 + h\tau) (\mathbf{p}(t_0 + h\tau) + O(h^{2s+2})) d\tau,$$

где $\Phi(t)$ — фундаментальная матрица системы однородных линейных уравнений. Квадратурная формула Гаусса для вычисления интеграла от функции $\Phi(t_0 + h) \Phi^{-1}(t_0 + h\tau) \mathbf{p}(t_0 + h\tau)$ дает нулевой результат, поскольку в точках коллокации $t_0 + hc_i$ возмущения $\mathbf{p}(t)$ равны нулю. Следовательно, интеграл можно рассматривать как ошибку квадратурной формулы, которая в данном случае представляет собой величину $O(h^{2s})$. Таким образом, $\mathbf{x}(t_0 + h) - \mathbf{x}_1 = O(h^{2s+1})$.

Известно, что узловые значения квадратурной формулы Гаусса на отрезке $[0, 1]$ являются корнями смещенного полинома Ле-

жандра, иначе говоря, они являются решениями уравнения

$$\frac{d^s}{d\tau^s}[\tau^s(\tau - 1)^s] = 0. \quad (4.56)$$

Поэтому для получения коллокационного метода порядка $p = 2s$ на практике удобно сначала вычислять c_i из уравнения (4.56), а затем — a_{ij} и b_i из формул (4.55).

Разбиение шага узловыми значениями c_i , удовлетворяющими уравнению (4.56), называется разбиением Гаусса–Лежандра. Если зафиксировать начальное узловое значение ($c_1 = 0$), либо конечное ($c_s = 1$), либо и то и другое ($c_1 = 0$ и $c_s = 1$), то получим другие известные разбиения Гаусса: левое и правое разбиения Радо и разбиение Лобатто, узловые значения которых соответственно удовлетворяют уравнениям

$$\begin{aligned} \frac{d^{s-1}}{d\tau^{s-1}}[\tau^s(\tau - 1)^{s-1}] &= 0, & \frac{d^{s-1}}{d\tau^{s-1}}[\tau^{s-1}(\tau - 1)^s] &= 0, \\ \frac{d^{s-2}}{d\tau^{s-2}}[\tau^{s-1}(\tau - 1)^{s-1}] &= 0. \end{aligned}$$

Однако любой коллокационный метод, построенный на одном из этих разбиений, будет иметь порядок ниже $2s$: узловые значения Радо дают метод порядка $p = 2s - 1$, а узловые значения Лобатто — $p = 2s - 2$. Методы, основанные на рассмотренных разбиениях, называют еще методами Гаусса.

Получим метод Кунцмана–Бутчера 4-го порядка ($s = 2$). Уравнение узлов (4.56) можно преобразовать к виду

$$6\tau^2 - 6\tau + 1 = 0, \quad \text{отсюда} \quad c_{1,2} = \frac{1}{2} \left(1 \mp \frac{1}{\sqrt{3}} \right).$$

Затем, используя формулы (4.50) с линейными функциями $l_1 = (\tau - c_2)/(c_1 - c_2)$ и $l_2 = (\tau - c_1)/(c_2 - c_1)$, находим коэффициенты

$$a_{11} = a_{22} = \frac{1}{4}, \quad a_{12,21} = \frac{1}{4} \left(1 \mp \frac{2}{\sqrt{3}} \right), \quad b_1 = b_2 = \frac{1}{2}.$$

Далее рассмотрим практическую реализацию коллокационных методов на примере интегратора Эверхарта, который широко применяется в небесной механике.

4.3.7. Интегратор Эверхарта

В 1973 г. Э. Эверхарт предложил интегратор (Everhart, 1973), разработанный им специально для численного исследования орбит, и продемонстрировал его высокую эффективность в задачах кометной динамики. По-видимому, обнаружив в дальнейшем принадлежность своего интегратора к семейству интеграторов типа Бутчера, Эверхарт акцентировал внимание на оригинально реализованном им алгоритме интегрирования и обобщил его для численного решения любых обыкновенных дифференциальных уравнений первого и второго порядков (Everhart, 1974; Everhart, 1985). Тем самым ему удалось расширить область применения своего интегратора, который тем не менее остается одним из самых популярных именно в решении задач небесной механики.

Основные формулы. Предположим, на шаге h мы решаем задачу Коши (4.4):

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}_0 = \mathbf{x}(t_0).$$

Представим приближенно правую часть уравнений \mathbf{f} в виде интерполяционного канонического многочлена степени $s-1$ на сетке c_1, \dots, c_s :

$$\mathbf{x}' = \mathbf{x}'_\tau/h = \mathbf{f} \approx \mathbf{g}'(t_0 + h\tau) = \sum_{j=1}^s \mathbf{a}_j \tau^{j-1}, \quad (4.57)$$

где коэффициенты \mathbf{a} пока не определены. Интегрируя (4.57) по τ , получаем приближенное решение

$$\mathbf{x}(t_0 + h\tau) \approx \mathbf{g}(t_0 + h\tau) = \mathbf{x}_0 + h \sum_{j=1}^s \frac{\mathbf{a}_j}{j} \tau^j. \quad (4.58)$$

Перепишем канонический многочлен (4.57) в виде интерполяционного многочлена Ньютона:

$$\mathbf{f} \approx \mathbf{g}'(t_0 + h\tau) = \boldsymbol{\alpha}_1 + \sum_{j=2}^s \boldsymbol{\alpha}_j \prod_{m=1}^{j-1} (\tau - c_m). \quad (4.59)$$

Равенство правой части уравнения и многочлена в (4.59) имеет место в точках коллокации, поэтому из соотношений

$$\mathbf{k}_1 = \mathbf{f}(t_0 + hc_1, \mathbf{g}(t_0 + hc_1)) = \mathbf{f}(t_0 + hc_1, \mathbf{y}_1) = \boldsymbol{\alpha}_1,$$

$$\begin{aligned} \mathbf{k}_i &= \mathbf{f}(t_0 + hc_i, \mathbf{g}(t_0 + hc_i)) = \mathbf{f}(t_0 + hc_i, \mathbf{y}_i) = \\ &= \boldsymbol{\alpha}_1 + \sum_{j=2}^i \boldsymbol{\alpha}_j \prod_{m=1}^{j-1} (c_i - c_m) \quad (i = 2, \dots, s) \end{aligned} \quad (4.60)$$

последовательно получаем разделенные разности $\boldsymbol{\alpha}$:

$$\begin{aligned} \boldsymbol{\alpha}_1 &= \mathbf{k}_1 \\ \boldsymbol{\alpha}_2 &= (\mathbf{k}_2 - \boldsymbol{\alpha}_1)/(c_2 - c_1), \\ \boldsymbol{\alpha}_3 &= ((\mathbf{k}_3 - \boldsymbol{\alpha}_1)/(c_3 - c_1) - \boldsymbol{\alpha}_2)/(c_3 - c_2), \\ &\dots \end{aligned} \quad (4.61)$$

В свою очередь, сравнивая коэффициенты многочленов (4.57) и (4.59) при одинаковых степенях τ , будем иметь

$$\begin{aligned} \mathbf{a}_1 &= \boldsymbol{\alpha}_1 + (-c_1)\boldsymbol{\alpha}_2 + (c_1c_2)\boldsymbol{\alpha}_3 + \dots + (-1)^{s-1}(c_1 \dots c_{s-1})\boldsymbol{\alpha}_s, \\ \mathbf{a}_2 &= \boldsymbol{\alpha}_2 + (-c_1 - c_2)\boldsymbol{\alpha}_3 + \dots, \\ &\dots \\ \mathbf{a}_s &= \boldsymbol{\alpha}_s. \end{aligned}$$

Эти соотношения, а также обратный переход от \mathbf{a} к $\boldsymbol{\alpha}$ можно представить как

$$\mathbf{a}_j = \sum_{i=j}^s c_{ij}\boldsymbol{\alpha}_i, \quad \boldsymbol{\alpha}_j = \sum_{i=j}^s d_{ij}\mathbf{a}_i \quad (j = 1, \dots, s). \quad (4.62)$$

Коэффициенты c_{ij} и d_{ij} являются числами Стирлинга, которые вычисляются по формулам

$$\begin{aligned} c_{ii} &= d_{ii} = 1, \quad c_{i0} = d_{i0} = 0 \quad (i > 0), \\ c_{ij} &= c_{i-1,j-1} - c_{i-1}c_{i-1,j}, \quad d_{ij} = d_{i-1,j-1} - c_jd_{i-1,j} \quad (i > j > 0). \end{aligned} \quad (4.63)$$

Интегрирование на шаге. Величины α (4.61) определяются по \mathbf{k} (4.60), которые, в свою очередь, вычисляются по приближенным решениям в точках коллокации \mathbf{y} . Согласно (4.58), эти решения будем задавать как

$$\mathbf{y}_i = \mathbf{g}(t_0 + hc_i) = \mathbf{x}_0 + h \sum_{j=1}^s \frac{\mathbf{a}_j}{j} c_i^j \quad (i = 1, \dots, s). \quad (4.64)$$

Формулы (4.64) представляют собой неявные уравнения относительно \mathbf{y} : $\mathbf{y} = \mathbf{y}(\mathbf{a}(\alpha(\mathbf{k}(\mathbf{y}))))$, поэтому они решаются итерационным способом. Для этого применяется метод простых итераций в модификации Зейделя.

Для получения начальных приближений $\bar{\mathbf{a}}$ и $\bar{\alpha}$ на следующем шаге \bar{h} используется информация о коэффициентах \mathbf{a} на текущем шаге h . Безразмерная независимая переменная следующего шага будет $\bar{\tau} = (t - t_1)/\bar{h}$, где $t_1 = t_0 + h$. Отсюда

$$\tau = r\bar{\tau} + 1, \quad (4.65)$$

где $r = \bar{h}/h$. Согласно (4.57),

$$\sum_{j=1}^s \mathbf{a}_j \tau^{j-1} = \sum_{j=1}^s \bar{\mathbf{a}}_j \bar{\tau}^{j-1}. \quad (4.66)$$

Подставляя (4.65) в (4.66) и приравнивая коэффициенты при одинаковых степенях $\bar{\tau}$, получаем

$$\bar{\mathbf{a}}_j = r^{j-1} \sum_{i=j}^s e_{ij} \mathbf{a}_i \quad (j = 1, \dots, s), \quad (4.67)$$

где e_{ij} — числа арифметического треугольника, вычисляемые по рекуррентным формулам

$$e_{ii} = e_{i1} = 1, \quad e_{ij} = e_{i-1,j-1} + e_{i-1,j} \quad (i > j > 1). \quad (4.68)$$

Кроме того, оценка $\bar{\mathbf{a}}$ уточняется путем внесения поправки $\Delta \mathbf{a}$, получаемой как разность между значениями \mathbf{a} после итераций и

оценкой $\bar{\mathbf{a}}$ на текущем шаге h . Далее пользуясь соотношениями (4.62) для перехода от \mathbf{a} к $\boldsymbol{\alpha}$, находим начальное приближение $\bar{\boldsymbol{\alpha}}$.

Каждая итерация выполняется следующим образом. Сначала определяется решение \mathbf{y}_1 , из которого по первой формуле (4.61) находится значение $\boldsymbol{\alpha}_1$, а затем по первой группе соотношений (4.62) уточняется \mathbf{a}_1 . Далее определяется \mathbf{y}_2 , по которому улучшается $\boldsymbol{\alpha}_2$ и соответственно \mathbf{a}_1 и \mathbf{a}_2 , и так до \mathbf{y}_s . Как правило, для получения достаточно хороших приближений необходимо всего лишь 2 итерации, очень редко — 3.

Как только величины \mathbf{a} получены, приближенное решение на шаге h ($\tau = 1$) будет

$$\mathbf{x}_1 = \mathbf{x}_0 + h \sum_{j=1}^s \frac{\mathbf{a}_j}{j}. \quad (4.69)$$

В начале интегрирования, на первом шаге, итерационный процесс буквально запускается с нуля, т.е. в качестве $\bar{\mathbf{a}}$ и $\bar{\boldsymbol{\alpha}}$ выбирают нулевые значения. Если начальный шаг достаточно большой, чтобы обеспечить заданную локальную точность, то его следует уменьшить. При оптимально выбранном шаге высокая точность приближений достигается уже на 4-й итерации.

В оригинальной версии интегратора Эверхарта используется левое разбиение Гаусса–Радо, а не Гаусса–Лежандра, поэтому аппроксимирующая схема имеет порядок $p = 2s - 1$. По-видимому, автор нашел это разбиение более удобным, поскольку оно не требует перевычислений коэффициентов в первой точке коллокации при решении нелинейных уравнений (4.64). Действительно, ведь в этом случае $c_1 = 0$ и $\mathbf{a}_1 = \boldsymbol{\alpha}_1 = \mathbf{k}_1 = \mathbf{f}(t_0, \mathbf{x}_0)$.

Выбор шага. В интеграторе Эверхарта контроль шага интегрирования осуществляется по величине последнего члена в (4.69). Пусть $\|\mathbf{e}\|_{tol}$ — заданная точность. Потребуем, чтобы на следующем шаге выполнялось равенство

$$\frac{\bar{h}}{s} \|\bar{\mathbf{a}}_s\| (= \|\mathbf{e}\|_{cal}) = \|\mathbf{e}\|_{tol}.$$

Отсюда, используя последнее соотношение (4.67) ($j = s$), получаем оценку, подобную (4.35):

$$\bar{h} = hr = h \left(\frac{s \|\mathbf{e}\|_{tol}}{\bar{h} \|\mathbf{a}_s\|} \right)^{1/s}. \quad (4.70)$$

Очевидно, при разбиениях Гаусса недостаток такой оценки состоит в том, что шаг по ней выбирается фактически как для решения порядка $s - 1$, поэтому, вообще говоря, она не обеспечивает сохранение требуемой локальной точности.

Во избежание слишком больших (и малых) локальных ошибок, на r следует наложить ограничение типа (4.38):

$$\frac{1}{\sigma} < r^s < \sigma. \quad (4.71)$$

Для того чтобы величина последнего члена в (4.69) была ограничена в пределах одного порядка, значение σ должно быть равно $\sqrt{10}$. Это следует из того факта, что

$$\bar{h} \|\bar{\mathbf{a}}_s\| \sim r^s.$$

Выполнение обоих неравенств проверяется лишь в начале интегрирования при выборе стартового шага: если (4.71) не выполняется, то интегрирование повторяется с новым шагом $\bar{h} = hr$ и так далее, пока не выполнится условие (4.71). Обычно для получения стартового шага требуется не более 4 итераций. В дальнейшем для ограничения r проверяется только правое неравенство: если неравенство не выполняется, то r принимает значение правого предела.

Начальное приближение стартового шага получается из оценки, подобной (4.70), для $s = 2$ (Авдюшев, 2010b):

$$\bar{h} = \sqrt{\frac{2h \|\mathbf{e}\|_{tol}}{\|\mathbf{k}_2 - \mathbf{k}_1\|}}, \quad \mathbf{k}_1 = \mathbf{f}(t_0, \mathbf{x}_0), \quad \mathbf{k}_2 = \mathbf{f}(t_0 + h, \mathbf{x}_0 + h\mathbf{k}_1),$$

где h — малая величина. Если h настолько мала, что в компьютерной арифметике $\mathbf{k}_1 = \mathbf{k}_2$, она увеличивается в 10 раз, и оценка повторяется снова.

Открытые вопросы и возможные решения. Смысл двойной полиномиальной интерполяции (4.57) и (4.59) вполне понятен: первый полином дает, с одной стороны, очень просто, с другой стороны, очень простую конструкцию приближенного решения как внутри (4.58), так и на конце шага (4.69), тогда как второй позволяет связать коэффициенты первого \mathbf{a} с правыми частями уравнения \mathbf{k} в узловых точках через разделенные разности $\boldsymbol{\alpha}$ посредством соотношений (4.62). Однако насколько насущной является необходимость прибегать к двойной интерполяции?

Даже при первом знакомстве с основными формулами интегратора видно, что, согласно (4.57) и условиям коллокации (4.46), величины \mathbf{a} могут быть выражены непосредственно через \mathbf{k} :

$$\mathbf{k}_i = \sum_{j=1}^s c_i^{j-1} \mathbf{a}_j, \quad \text{отсюда} \quad \mathbf{a}_i = \sum_{j=1}^s b_{ij} \mathbf{k}_j \quad (i = 1, \dots, s),$$

где b_{ij} — элементы матрицы, обратной к матрице Вандермонда с элементами c_i^{j-1} . Тогда величины $\boldsymbol{\alpha}$ становятся не нужны, поэтому итерационный процесс для решения нелинейных уравнений (4.64) можно было бы существенно упростить.

Кроме того, имеется эффективный способ составления обратной матрицы Вандермонда \mathbf{V} без численного обращения (Turner, 1966). Ее элементы могут быть определены из равенства интерполяционных полиномов, канонического и лагранжева:

$$\sum_{j=1}^s \mathbf{a}_j \tau^{j-1} = \sum_{j=1}^s \mathbf{k}_j l_j(\tau), \quad (4.72)$$

путем сопоставления их коэффициентов при одинаковых степенях τ . Получаемую матрицу \mathbf{V} удобно представить как произведение верхней \mathbf{U} и нижней-треугольной \mathbf{V} матриц: $\mathbf{V} = \mathbf{U}\mathbf{V}$. Элементы первой матрицы вычисляются по рекуррентным формулам, аналогичным (4.63):

$$\begin{aligned} u_{ii} &= 1, & u_{0j} &= 0 \quad (j > 0), \\ u_{ij} &= u_{i-1,j-1} - c_{j-1} u_{i,j-1} \quad (j > i > 0), \end{aligned}$$

тогда как элементы второй —

$$v_{11} = 1, \quad v_{ij} = \prod_{\substack{m=1 \\ m \neq j}}^i \frac{1}{c_j - c_m} \quad (j \leq i).$$

Почему Эверхарт не воспользовался этим подходом? Вероятно потому что обусловленность матрицы Вандермонда при увеличении степени интерполяции значительно ухудшается, в то время как обращение плохообусловленных матриц сопряжено с большими ошибками. Например, для интерполяции на разбиении Гаусса–Лежандра при $s = 8$ ($p = 16$) число обусловленности матрицы Вандермонда составляет порядка 10^5 , что может привести к потере вычислительной точности до пяти значащих цифр. Впрочем, при вычислении разделенных разностей α (в оригинальной версии интегратора с двойной интерполяцией) потеря точности не меньше (а возможно даже и больше).

Тогда возникает другой вопрос: если не каноническое представление (4.57) как одинарная интерполяция, почему не в форме Ньютона (4.59)? Очевидно, что конструкция решения в точках коллокации и на конце шага для ньютоновской интерполяции становится сложнее²⁵:

$$\mathbf{y}_i = \mathbf{x}_0 + \alpha_1 h c_i + h \sum_{j=2}^s \alpha_j \int_0^{c_i} \prod_{m=1}^{j-1} (\tau - c_m) d\tau \quad (i = 1, \dots, s),$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_1 h + h \sum_{j=2}^s \alpha_j \int_0^1 \prod_{m=1}^{j-1} (\tau - c_m) d\tau. \quad (4.73)$$

Тем не менее итерационный процесс столь же упрощается, как и при использовании канонического полинома, поскольку теперь отпадает необходимость вычислять величины \mathbf{a} .

Видимо, в данном случае автора смущали интегралы, которые должны вычисляться численно, и ему неизвестны были рекур-

²⁵Впрочем, она все же оказывается немного проще, нежели в коллокационных методах с интерполяцией Лагранжа (4.49).

рентный формулы типа (4.55). Хотя может быть и другое объяснение: например, усложнение определения начальных приближений $\bar{\mathbf{a}}$ для итерационного решения нелинейных уравнений.

Следует также отметить одну неприятную особенность в представлении решения (4.73), о чем уже вскользь было упомянуто. Вычисление разделенных разностей α предполагает вычитание близких величин, что вследствие вычислительных ошибок может приводить к низкой точности как самих разделенных разностей, так и составленного из них решения (4.73).

В смысле всего сказанного интерполяция Лагранжа для конструирования численного решения оказывается лучшим выбором, поскольку программная реализация коллокационного метода на ее основе не требует вычислений ни коэффициентов \mathbf{a} , ни разделенных разностей α . Однако, если бы Эверхарт выбрал именно эту интерполяцию, он неизменно получил бы легко узнаваемый неявный метод Рунге–Кутты, и вполне возможно, что такой удручающий факт склонил бы Эверхарта прекратить дальнейшие поиски оригинальных схем интегрирования и его интегратор в известной версии так бы никогда и не появился.

Между тем нельзя не отметить одну на первый взгляд незаметную, но важную деталь интегратора Эверхарта, которая значительно повышает эффективность программной реализации коллокационных методов. Дело в том, что начальные приближения $\bar{\mathbf{a}} + \Delta\mathbf{a}$ (иначе говоря, экстраполяция коэффициентов на следующий шаг $\bar{\mathbf{a}}$ с уточнением по информации об ошибках начальных коэффициентов $\Delta\mathbf{a}$ на текущем шаге) для итерационного решения нелинейных уравнений лучше с точки зрения сходимости, нежели аналогичные начальные приближения $\bar{\mathbf{k}} + \Delta\mathbf{k}$, тем, собственно, и хороша конструкция решения (4.58): здесь как раз Эверхарт не прогадал. Эта особенность интегратора Эверхарта позволяет получать достаточно хорошие результаты уже после двух итераций на шаге, в то время как другие интеграторы, основанные на интерполяции Лагранжа, могут потребовать больше итераций. Этот факт проверяется экспериментально.

Поскольку интегратор Эверхарта фактически основан на формулах коллокационных методов, он наследует как их достоинства, так и недостатки. Одна из главных проблем интегратора Эверхарта — это недооценивания локальной погрешности, о чем уже говорилось выше. Теоретически проблема разрешается очень просто, хотя практически это требует некосметической редакции программного кода.

Идея по сути состоит в том, чтобы оценивать локальную ошибку по старшему члену канонического полинома (4.58) более высокого порядка и вместо формулы (4.70) использовать

$$\bar{h} = hr = h \left(\frac{p \|\mathbf{e}\|_{tol}}{h \|\mathbf{a}_p\|} \right)^{1/p}.$$

Разумеется, можно также обратиться и к обобщенному варианту (4.37), полагая, что $(e_i)_{cal} = h(a_p)_i/p$. Но где взять коэффициент \mathbf{a}_p ? Правильно, нужно использовать значения правых частей на предыдущем шаге. Согласно (4.72) при $s = p$,

$$\mathbf{a}_p = \sum_{j=1}^p \mathbf{k}_j \prod_{\substack{m=1 \\ m \neq j}}^p \frac{1}{c_j - c_m} = \sum_{j=1}^p b_{pj} \mathbf{k}_j, \quad (4.74)$$

где \mathbf{k}_i для $i = s + 1, \dots, p$ соответствуют узловым значениям на предыдущем шаге $c_i = (c_{i-s} - 1)/r$. Кстати, эту формулу можно использовать и для любых других гауссовых интеграторов.

К сожалению, вследствие вычислительных ошибок формула (4.74) не всегда обеспечивает хорошую оценку \mathbf{a}_p . Ее точность падает с увеличением порядка p (степени интерполяции) и с уменьшением шага h (интервала интерполяции). Таким образом, уменьшение шага интегрирования приводит к понижению наибольшего эффективного порядка, для которого еще можно получить приемлемую оценку \mathbf{a}_p . Фактически формула работает только для методов невысоких порядков. В частности, при интегрировании уравнений орбитального движения в прямоугольных координатах в арифметике с двойной точностью они редко превышают восьми. В методах более высоких порядков вычислительные ошибки

могут быть настолько грубыми, что способны значительно увеличить оценку \mathbf{a}_p , а это, в свою очередь, провоцирует выбор малой величины шага, меньшей, чем требуется.

Коль скоро речь зашла об одном из применений численных решений предыдущего шага, уместно отметить другое их возможное назначение. Почему бы эти решения не подключить к построению схемы интегрирования, повышая тем самым ее порядок? Почему бы не решения предыдущего шага? И так далее. При таком подходе гауссовы разбиения теряют смысл: действительно, ведь тогда оказывается возможным получить схему интегрирования любого порядка даже на основе равномерного разбиения при том же количестве обращений к правым частям уравнений. На самом деле этот подход есть не что иное, как обобщение идеи построения многошаговых схем (см. подразд. 4.5.1), где каждое следующее решение (одно, а не несколько) определяется явно или неявно по уже полученным на предыдущих шагах (точнее, по соответствующим правым частям) как результат интегрирования интерполяционного полинома. Однако подводные камни здесь — это все те же пресловутые вычислительные ошибки, которые тем больше, чем выше степень интерполяции. В этом смысле примечательность гауссовых разбиений состоит в том, что они позволяют при высоком порядке метода интегрирования использовать интерполяцию низкой степени и тем самым уменьшить влияние вычислительных ошибок. Это в принципе объясняет тот факт, что в многошаговых схемах интегрирования, где практически невозможно использование гауссовых разбиений, вычислительные ошибки больше, нежели в гауссовых методах тех же порядков, что подтверждается результатами численного эксперимента в подразд. 4.7.

4.3.8. Порядок и шаг интегрирования при компьютерной реализации метода

Теоретически совместное увеличение порядка и уменьшение шага метода неограниченно повышают методическую точность численных результатов интегрирования. Однако при компьютерной

реализации в арифметике с определенной точностью вследствие ошибок округления существуют такие значения параметров интегрирования, которые дают предельно возможную наивысшую точность, когда методические ошибки становятся соизмеримыми с ошибками округления, и в этом случае не имеет смысла предпринимать какие-либо дальнейшие попытки получить более высокую точность численного интегрирования путем манипулирования порядком и шагом метода.

Получим оценки оптимальной пары порядок–шаг интегрирования применительно к численному решению дифференциальных уравнений круговой двумерной задачи двух тел:

$$\mathbf{r}' = \mathbf{v}, \quad \mathbf{v}' = -\frac{\mu}{|\mathbf{r}|^3} \mathbf{r}.$$

Здесь $\mathbf{r} = (r_1, r_2)^T$ и $\mathbf{v} = (v_1, v_2)^T$ — векторы положения и скорости соответственно, а μ — гравитационный параметр. Поскольку $|\mathbf{r}| = a = \text{const}$, будем полагать, что величина $\mu/|\mathbf{r}|^3 = \nu^2 = \text{const}$, т.е. решение $\mathbf{x} = (\mathbf{r}, \mathbf{v})^T$ описывается уравнениями гармонического осциллятора с частотой ν и может быть записано в виде

$$\begin{aligned} x_1 = r_1 &= a \cos \nu t, & x_2 = r_2 &= a \sin \nu t, \\ x_3 = v_1 &= -a\nu \sin \nu t, & x_4 = v_2 &= a\nu \cos \nu t. \end{aligned} \quad (4.75)$$

Оценим методическую ошибку $\|\mathbf{e}\|_M$ по главному члену погрешности (4.10):

$$\|\mathbf{e}\|_M = \frac{a\sqrt{1+\nu^2}(\nu h)^{p+1}}{(p+1)!}. \quad (4.76)$$

Здесь использована формула $\|\mathbf{x}^{(p+1)}\| = a\nu^{p+1}\sqrt{1+\nu^2}$. Согласно формулам методов Рунге–Кутты, ошибку округления $\|\mathbf{e}\|_R$ можно оценить как

$$\|\mathbf{e}\|_R = \varepsilon \|\mathbf{x}\| = \varepsilon a \sqrt{1+\nu^2}, \quad (4.77)$$

где ε — машинный эпсилон.

Очевидно, что не имеет смысла выбирать такие шаг и порядок интегрирования, при которых методическая ошибка будет меньше

ошибки округления. Из условия $\|\mathbf{e}\|_M = \|\mathbf{e}\|_R$ получим отношение между оптимальными параметрами интегрирования h и p :

$$h_\varphi^{p+1} = \varepsilon(p+1)!, \quad (4.78)$$

где $h_\varphi = \nu h$ — шаг по долготе $\varphi = \nu t$, соответствующий шагу h .

Отношение (4.78) дает нижнюю границу для шага h , в то время как для неявных методов имеет место верхняя граница, задаваемая условием (4.45). Если положить, что $\max_i \sum_j |a_{ij}| = 1$ (в действительности максимум близок к единице), то получим следующее ограничение на шаг интегрирования $h < 1/L$ или $h_\varphi < \nu/L$. Оценим постоянную Липшица L для исследуемой задачи.

Рассмотрим отношение

$$\frac{\|\delta\mathbf{f}\|^2}{\|\delta\mathbf{x}\|^2} = \frac{\nu^4|\delta\mathbf{r}|^2 + |\delta\mathbf{v}|^2}{|\delta\mathbf{r}|^2 + |\delta\mathbf{v}|^2},$$

где $\delta\mathbf{x}$, $\delta\mathbf{r}$ и $\delta\mathbf{v}$ — всевозможные разности векторов в соответствующих переменных. Принимая $|\delta\mathbf{r}| = \xi \cos \psi$ и $|\delta\mathbf{v}| = \xi \sin \psi$, где $\xi \geq 0$ и $0 \leq \psi \leq \pi/2$, будем иметь

$$\frac{\|\delta\mathbf{f}\|^2}{\|\delta\mathbf{x}\|^2} = (\nu^4 - 1) \cos^2 \psi + 1.$$

Отсюда нетрудно видеть, что все значения отношения лежат между 1 и ν^4 . Следовательно, согласно (4.5), в качестве постоянной Липшица можно выбрать $L = \max(1, \nu^2)$. Тогда получаем верхнюю границу шага

$$h_\varphi < \frac{\nu}{\max(1, \nu^2)} \leq 1.$$

Таким образом, в лучшем случае, а именно при $\nu = 1$, когда верхняя граница максимальна, шаг интегрирования h_φ должен удовлетворять неравенствам

$$\varepsilon(p+1)! < h_\varphi^{p+1} < 1.$$

Таблица 4.1. Оптимальные порядки p неявных методов Рунге–Кутты для различных ε

ε	$1.1 \cdot 10^{-7}$	$2.2 \cdot 10^{-16}$	$1.1 \cdot 10^{-19}$	$1.9 \cdot 10^{-34}$
p	9	16	19	29

Очевидно, условие

$$\varepsilon(p+1)! > 1$$

означает, что порядок метода завышен и использование такого метода при вычислениях в арифметике с точностью ε неразумно в том смысле, что ту же точность результатов интегрирования можно получить с использованием методов более низких порядков. Оптимальные порядки p неявных методов Рунге–Кутты для различных ε , соответствующих одинарной, двойной, расширенной и четверной точности, представлены в табл. 4.1. Хотя следует иметь в виду, что эти порядки получены для задачи с $\nu = 1$. В ином случае они могут быть меньше.

4.3.9. Орбитальная устойчивость методов Рунге–Кутты

Согласно решениям (4.75), дифференциальные уравнения для координат в круговом движении можно представить в виде

$$x'_1 = -\nu x_2, \quad x'_2 = \nu x_1. \quad (4.79)$$

Эта система эквивалентна дифференциальному уравнению

$$r' = i\nu r, \quad (4.80)$$

где r — уже комплексная переменная, аналог вектора положения \mathbf{r} с действительной частью r_1 и мнимой r_2 . При начальном условии $r_0 = r(t_0)$ точное решение линейного уравнения (4.80) будет

$$r(t) = e^{i\nu(t-t_0)} r_0.$$

Применяя метод Рунге–Кутты (4.39) к уравнению (4.80), получим последовательность приближенных решений

$$r_{n+1} = S(i\nu h)r_n.$$

Здесь

$$S(z) = 1 + z\mathbf{b} \cdot (\mathbf{I} - z\mathbf{A})^{-1}\mathbf{1}$$

— так называемая функция устойчивости (Деккер, Вервер, 1989); \mathbf{I} — единичная матрица размера $s \times s$; $\mathbf{1} = (1, \dots, 1)^T$ — s -мерный вектор с единичными компонентами;

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \dots & a_{ss} \end{pmatrix}; \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}.$$

Поскольку $|r(t)| = a = \text{const}$, желательно, чтобы метод удовлетворял этому условию, т.е. чтобы $|S(i\nu h)| = 1$, и тогда численные решения r_n будут всегда оставаться на круговой орбите. Поэтому назовем это свойство методов орбитальной устойчивостью, а если оно выполняется для любых h — абсолютной орбитальной устойчивостью.

Например, коллокационные методы на гауссовых разбиениях Лобатто и Лежандра — абсолютно орбитально устойчивые. В то же время на разбиениях Радо коллокационные методы — неустойчивые ни для каких шагов. Численные орбиты для таких методов (в частности, для явного и неявного методов Эйлера) представляют собой спиралевидные траектории: при левых разбиениях Радо вычисляемые положения отдаляются от центра, тогда как при правых — приближаются.

Разумеется, уравнения кругового движения в задаче двух тел принципиально сложнее уравнений (4.79), хотя бы потому что частота в первых рассматривается как функция координат: $\nu = \sqrt{\mu/|\mathbf{r}|^3}$, когда в последних — это постоянный параметр, $\nu = \sqrt{\mu/a^3}$. Тем не менее, если даже при интегрировании уравнений кругового орбитального движения орбитально устойчивый метод

не сохраняет радиус-вектор, в любом случае от него можно ожидать лучших результатов, нежели от неустойчивого метода.

4.4. Экстраполяционные методы

Экстраполяционные методы основаны на том факте (который мы не будем доказывать), что глобальная погрешность приближенного решения \mathbf{x}_N , полученного методом порядка p на отрезке $[t_0, t_0 + \Delta t]$ с постоянным шагом $h = \Delta t/N$, допускает разложение в степенной ряд по h (Gragg, 1964):

$$\mathbf{x}_N - \mathbf{x}(t_0 + \Delta t) = \sum_{i=p}^{\infty} \mathbf{e}_i h^i. \quad (4.81)$$

Формула (4.81) представляет приближенное решение как функцию величины шага $\mathbf{x}_N = \mathbf{x}_N(h)$ ²⁶. Для этой функции

$$\mathbf{x}_N(h) \rightarrow \mathbf{x}(t_0 + \Delta t) \quad \text{при} \quad h \rightarrow 0 \quad \text{и} \quad N \rightarrow \infty.$$

Теоретически это означает, что при достаточно малом шаге можно любым методом получить приближенное решение со сколь угодно высокой точностью. С практической точки зрения использование очень малых шагов неразумно, поскольку, во-первых, это увеличивает объем вычислений, что, во-вторых, сопряжено с большими ошибками округления. Тем не менее, используя идею экстраполяции, можно получить с высокой точностью предельное значение $\mathbf{x}_{\infty}(0)$, избегая при этом чрезмерно громоздких вычислений.

4.4.1. Общий подход

В соответствии с (4.81) представим интерполяцию приближенного решения в виде многочлена порядка $p + s - 2$:

$$\mathbf{g}(h) = \mathbf{a}_0 + \mathbf{a}_p h^p + \dots + \mathbf{a}_{p+s-2} h^{p+s-2}. \quad (4.82)$$

²⁶См. обозначение на с. 137.

Будем рассматривать значения величины шага h на полуотрезке $(0, H]$, где H — величина шага экстраполяционного метода. Выберем последовательность целых положительных чисел

$$N_1 < \dots < N_s \quad (4.83)$$

и определим соответствующую последовательность шагов $h_i = H/N_i$ ($i = 1, \dots, s$). Далее для каждого шага h_i вычислим решение $\mathbf{x}_{N_i(h_i)}$ на момент $t_0 + H$ каким-либо методом Рунге–Кутты порядка p . Тогда, требуя, чтобы $\mathbf{g}(h_i) = \mathbf{x}_{N_i(h_i)}$, получим следующую систему алгебраических уравнений:

$$\begin{aligned} \mathbf{g}(h_1) &= \mathbf{a}_0 + \mathbf{a}_p h_1^p + \dots + \mathbf{a}_{p+s-2} h_1^{p+s-2} = \mathbf{x}_{N_1(h_1)}, \\ \dots & \\ \mathbf{g}(h_s) &= \mathbf{a}_0 + \mathbf{a}_p h_s^p + \dots + \mathbf{a}_{p+s-2} h_s^{p+s-2} = \mathbf{x}_{N_s(h_s)}. \end{aligned} \quad (4.84)$$

После определения из системы (4.84) коэффициентов интерполирующего многочлена нам остается выполнить экстраполяцию при $h \rightarrow 0$, что, согласно (4.81) и (4.82), дает приближенное решение $\mathbf{a}_0 = \mathbf{g}(0)$ уже порядка $p + s - 1$.

При $s = 2$ интерполяционный многочлен приводится к виду

$$\mathbf{g}(h) = \mathbf{a}_0 + \mathbf{a}_p h^p,$$

коэффициенты которого будут определяться из системы

$$\begin{aligned} \mathbf{a}_0 + \mathbf{a}_p h_1^p &= \mathbf{x}_{N_1(h_1)}, \\ \mathbf{a}_0 + \mathbf{a}_p h_2^p &= \mathbf{x}_{N_2(h_2)}. \end{aligned}$$

Отсюда получаем экстраполяционную формулу (4.31):

$$\mathbf{a}_0(= \hat{\mathbf{x}}_1) = \mathbf{x}_{N_1(h_1)} + \frac{\mathbf{x}_{N_2(h_2)} - \mathbf{x}_{N_1(h_1)}}{1 - (N_1/N_2)^p}.$$

4.4.2. Алгоритм Эйткена–Невилла

При $p = 1$ (метод Эйлера) формулы (4.82) и (4.84) приводят к классической задаче интерполяции. В этом случае для упрощения вычислений интерполяционного многочлена удобно использовать схему Эйткена–Невилла.

Пусть $g_{i,j+1}(h)$ — интерполяционный многочлен степени j с $j + 1$ узлами интерполяции h_{i-j}, \dots, h_i для некоторой функции $g(h)$, в частности $g_{i,1}(h) = g(h_i)$. Тогда справедливы равенства

$$g_{i,j+1}(h) = \frac{g_{i,j}(h)(h - h_{i-j}) - g_{i-1,j}(h)(h - h_i)}{h_i - h_{i-j}},$$

а для экстраполяции при $h = 0$ имеем

$$g_{i,j+1} = \frac{g_{i-1,j}h_i - g_{i,j}h_{i-j}}{h_i - h_{i-j}}, \quad (4.85)$$

где $g_{i,j} = g_{i,j}(0)$. Фактически, вычисление экстраполяционного значения интерполяционного многочлена $g_{s,s}$ на сетке h_1, \dots, h_s сводится к последовательному вычислению по рекуррентной формуле (4.85) элементов таблицы

$$\begin{array}{cccc} g_{11} & & & \\ g_{21} & g_{22} & & \\ g_{31} & g_{32} & \dots & \\ \dots & \dots & \dots & \\ g_{s1} & g_{s2} & \dots & g_{ss} \end{array}$$

Из (4.85) получаем расчетную формулу для экстраполяции в векторном случае

$$\mathbf{g}_{i,j+1} = \mathbf{g}_{i,j} + \frac{\mathbf{g}_{i,j} - \mathbf{g}_{i-1,j}}{N_i/N_{i-j} - 1} \quad (j = 1, \dots, s-1; i = j+1, \dots, s), \quad (4.86)$$

где $\mathbf{g}_{i,1} = \mathbf{x}_{N_i(h_i)}$ ($i = 1, \dots, s$) вычисляются по схеме Эйлера:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{f}(t_n, \mathbf{x}_n) \quad (n = 0, \dots, N-1), \quad \mathbf{x}_{N(h)} = \mathbf{x}_N.$$

В итоге будем иметь экстраполированное решение \mathbf{g}_{ss} порядка s .

В (4.86) для получения узловых значений в качестве последовательности (4.83) обычно используют

$$1, 2, 4, 8, 16, 32, \dots \quad (\text{последовательность Ромберга})$$

либо

$$1, 2, 3, 4, 5, 6, \dots \quad (\text{гармоническая последовательность}).$$

4.4.3. Метод Грэгга

Если опорный метод порядка $p = 2q$ является симметричным (см. подразд. 4.6.4), то, как известно (Hairer et al., 1993), он будет иметь разложение глобальной ошибки по степеням h^2 , т.е.

$$\mathbf{x}_N - \mathbf{x}(t_0 + \Delta t) = \sum_{i=q}^{\infty} \mathbf{e}_i h^{2i}.$$

Следовательно, для экстраполяции нужно использовать интерполирующий полином вида

$$\mathbf{g}(h) = \mathbf{a}_0 + \mathbf{a}_q(h^2)^q + \dots + \mathbf{a}_{q+s-2}(h^2)^{q+s-2}. \quad (4.87)$$

Ввиду схожести полиномов (4.87) и (4.82) (при определенном s), вычисление их коэффициентов будет одинаково трудоемко в обоих случаях. Однако примечательно то, что в итоге интерполяция (4.87) дает экстраполированное решение \mathbf{a}_0 порядка $2(q + s - 1) = p + 2s - 2$, на $s - 1$ выше, нежели для (4.82).

У.Б. Грэгг (Gragg, 1965) предложил в качестве опорного использовать двушаговый симметричный метод второго порядка, основанный на правиле средней точки:

$$\mathbf{x}_1 = \mathbf{x}_0 + h\mathbf{f}(t_0, \mathbf{x}_0), \quad \mathbf{x}_{n+1} = \mathbf{x}_{n-1} + 2h\mathbf{f}(t_n, \mathbf{x}_n) \quad (n = 1, \dots, N),$$

$$\mathbf{x}_{N(h)} = \frac{\mathbf{x}_{N-1} + 2\mathbf{x}_N + \mathbf{x}_{N+1}}{4}. \quad (4.88)$$

Поскольку интерполирующий многочлен (4.87) имеет форму (4.82), для вычисления экстраполяции используется та же формула Эйткена–Невилла, где только h заменяется на h^2 , а именно:

$$\mathbf{g}_{i,j+1} = \mathbf{g}_{i,j} + \frac{\mathbf{g}_{i,j} - \mathbf{g}_{i-1,j}}{(N_i/N_{i-j})^2 - 1} \quad (j = 1, \dots, s-1; i = j+1, \dots, s). \quad (4.89)$$

Здесь $\mathbf{g}_{i,1} = \mathbf{x}_{N_i(h_i)}$ ($i = 1, \dots, s$) получаются из (4.88). При этом последовательности (4.83) выбираются с условием, что N_i должны быть четными. Например,

$$2, 4, 8, 16, 32, \dots$$

Метод, основанный на (4.88) и (4.89), называется методом Грэгга (или Грэгга–Булирша–Штера). Именно благодаря своим замечательным особенностям он единственный из всех экстраполяционных методов широко используется на практике.

4.4.4. Выбор шага

Шаг в экстраполяционных методах выбирается так же, как и во вложенных методах Рунге–Кутты. Второй индекс j в решении $\mathbf{g}_{i,j}$ определяет его порядок аппроксимации. В методе Грэгга решение $\mathbf{g}_{s,s}$ имеет (наивысший) порядок $2s$, а $\mathbf{g}_{s,s-1}$ — порядок $2s-2$. Поэтому для управления шагом естественно использовать оценку

$$\|\mathbf{e}\|_{cal} = \|\mathbf{g}_{s,s} - \mathbf{g}_{s,s-1}\|.$$

4.5. Многошаговые методы

До сих пор мы рассматривали одношаговые методы, в которых численное решение получают на основе только одного решения на предыдущем шаге. Многошаговые методы, в отличие от одношаговых, используют несколько решений, вычисленных на предыдущих шагах. В начале интегрирования, когда известно только одно решение (начальное условие), первые (стартовые) решения

для многошаговых методов обычно вычисляют с помощью одношаговых методов Рунге–Кутты²⁷. Затем на каждом следующем шаге выполняют многошаговую процедуру интегрирования.

Мы уже знакомы с одним из таких методов (4.88), который использовался в экстраполяционном методе Грэгга в качестве опорного. Это двушаговый симметричный метод второго порядка, где для вычисления начального второго решения используется метод Эйлера.

4.5.1. Методы Адамса

Многошаговые методы появились гораздо раньше, чем методы Рунге–Кутты. Впервые их получил Дж.К. Адамс еще в 1855 г. (Hairer et al., 1993). Тем же способом, что и Адамс, выведем первое семейство многошаговых методов (Адамса–Башфорта).

Предположим, для задачи Коши (4.4):

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0;$$

нам известны первые s решений $\mathbf{x}_0, \dots, \mathbf{x}_{s-1}$ на равномерной сетке t_0, \dots, t_{s-1} с шагом h . Тогда для шага s формально решение можно представить как

$$\mathbf{x}(t_s) = \mathbf{x}(t_{s-1}) + \int_{t_{s-1}}^{t_s} \mathbf{f}(t, \mathbf{x}(t)) dt. \quad (4.90)$$

Заменим подынтегральную функцию на ее интерполирующий полином Лагранжа (4.48), полученный по узловым значениям $\mathbf{k}_i = \mathbf{f}(t_i, \mathbf{x}_i)$ ($i = 0, \dots, s-1$):

$$\mathbf{g}(t) = \mathbf{g}(t_0 + h\tau) = \sum_{i=0}^{s-1} \mathbf{k}_i \prod_{\substack{j=0 \\ j \neq i}}^{s-1} \frac{\tau - j}{i - j}, \quad (4.91)$$

²⁷Хотя имеются многошаговые методы с возможностью самостарта (см. подразд. 4.5.9).

где, как и выше, $\tau = (t - t_0)/h$. Тогда численный аналог (4.90) будет задаваться формулой²⁸

$$\mathbf{x}_s = \mathbf{x}_{s-1} + \int_{t_{s-1}}^{t_s} \mathbf{g}(t) dt = \mathbf{x}_{s-1} + h \sum_{i=0}^{s-1} \mathbf{k}_i \int_{s-1}^s \prod_{\substack{j=0 \\ j \neq i}}^{s-1} \frac{\tau - j}{i - j} d\tau. \quad (4.92)$$

Согласно (4.92), для $s = 1, 2, 3$ получаем следующие явные формулы Адамса:

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0 + h\mathbf{k}_0 \quad (\text{метод Эйлера}), \\ \mathbf{x}_2 &= \mathbf{x}_1 + h \left[\frac{3}{2}\mathbf{k}_1 - \frac{1}{2}\mathbf{k}_0 \right], \\ \mathbf{x}_3 &= \mathbf{x}_2 + h \left[\frac{23}{12}\mathbf{k}_2 - \frac{16}{12}\mathbf{k}_1 + \frac{5}{12}\mathbf{k}_0 \right]. \end{aligned}$$

Формулы Адамса получаются при интегрировании интерполяционного многочлена (4.91) от t_{s-1} до t_s , т.е. вне интервала интерполяции. Однако, как мы знаем, вне этого интервала интерполяционный многочлен обычно дает довольно плохое приближение. Чтобы разрешить эту проблему, Адамс предложил для интерполяции использовать также значение \mathbf{k}_s . В итоге он получил неявные методы (Адамса–Мульттона). Приведем первые из них для $s = 0, 1$:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{x}_{-1} + h\mathbf{k}_0 \quad (\text{метод Эйлера}), \\ \mathbf{x}_1 &= \mathbf{x}_0 + h \left[\frac{1}{2}\mathbf{k}_1 + \frac{1}{2}\mathbf{k}_0 \right] \quad (\text{метод трапеций}). \end{aligned}$$

Итак, любой s -шаговый метод Адамса можно представить в общем виде

$$\mathbf{x}_s = \mathbf{x}_{s-1} + h \sum_{i=0}^s \beta_i \mathbf{k}_i, \quad (4.93)$$

²⁸Здесь, как и в случае коллокационных методов, можно увидеть сходство методов Адамса с квадратурными формулами Ньютона–Котеса.

где β_i — постоянные метода. Если $\beta_s = 0$, метод явный, иначе — неявный. Как и в коллокационных методах (см. подразд. 4.3.5), порядок метода Адамса определяется точностью интерполирующей формулы (4.91). В общем случае явный метод (4.93) имеет порядок $p = s$, неявный — $p = s + 1$.

Коэффициенты β_i выражаются через интегралы: в методах Адамса–Башфорта

$$\beta_i = \int_{s-1}^s \prod_{\substack{j=0 \\ j \neq i}}^{s-1} \frac{\tau - j}{i - j} d\tau \quad (i = 0, \dots, s-1);$$

в методах Адамса–Мультона

$$\beta_i = \int_{s-1}^s \prod_{\substack{j=0 \\ j \neq i}}^s \frac{\tau - j}{i - j} d\tau \quad (i = 0, \dots, s).$$

Для вычисления этих интегралов можно воспользоваться рекуррентными формулами, аналогичными (4.55). Так, для каждого коэффициента β_i в явных методах будем иметь

$$\begin{aligned} \gamma_{-1l} &= 1/l! \quad (l = 0, \dots, s+1), \\ \gamma_{kl} &= ((s-k)\gamma_{k-1,l} - l\gamma_{k-1,l+1})/(i-k) \quad (4.94) \\ &(k = 0, \dots, i-1, i+1, \dots, s; l = 0, \dots, s-k), \end{aligned}$$

причем $\gamma_{i,l} = \gamma_{i-1,l}$ ($l = 0, \dots, s-i$). Тогда $\beta_i = \gamma_{s-1,0}$. Эти же формулы применимы и к неявным методам, если увеличить правые пределы индексов k и l на единицу. Тогда $\beta_i = \gamma_{s,0}$.

4.5.2. Формулы дифференцирования

Формулы Адамса основаны на приближенном вычислении интеграла в (4.90), хотя многошаговые методы можно получить иным способом, а именно численно дифференцируя искомое решение.

Интерполяцию Лагранжа для \mathbf{x} по решениям $\mathbf{x}_0, \dots, \mathbf{x}_s$ формально можно представить в виде

$$\mathbf{g}(t) = \mathbf{g}(t_0 + h\tau) = \sum_{i=0}^s \mathbf{x}_i \prod_{\substack{j=0 \\ j \neq i}}^s \frac{\tau - j}{i - j}. \quad (4.95)$$

Последнее решение \mathbf{x}_s пока неизвестно и определим его так, чтобы многочлен \mathbf{g} удовлетворял дифференциальному уравнению, по крайней мере, в одной узловой точке:

$$\mathbf{g}'(t_r) = \mathbf{f}(t_r, \mathbf{x}_r) \quad \text{или} \quad \mathbf{g}'_r|_{\tau=r} = h\mathbf{k}_r. \quad (4.96)$$

Здесь индекс r принимает одно из значений $0, 1, \dots, s$. После взятия производной в (4.96) и подстановки t_r или r получаем схему многошагового интегрирования для \mathbf{x}_s . Если $r = s$, метод неявный (метод Гира), иначе — явный. В качестве примера приведем неявный метод для $s = 2$:

$$\frac{3}{2}\mathbf{x}_2 - 2\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_0 = h\mathbf{k}_2.$$

Таким образом, s -шаговый метод, полученный на основе дифференцирования, представим в виде

$$\sum_{i=0}^s \alpha_i \mathbf{x}_i = h\mathbf{k}_r, \quad \text{где} \quad \alpha_i = \sum_{\substack{k=0 \\ k \neq i}}^s \frac{1}{i-k} \prod_{\substack{j=0 \\ j \neq i, k}}^s \frac{r-j}{i-j}. \quad (4.97)$$

Порядок метода определяется точностью интерполирующей формулы (4.95). В общем случае метод (4.97) имеет порядок $p = s$.

4.5.3. Предиктор–корректор

Неявная схема интегрирования предполагает итерационное решение системы нелинейных уравнений относительно искомого решения \mathbf{x}_s . Для этого обычно применяют метод простых итераций. Легко показать, что в случае методов Адамса (4.93) и дифференцирования (4.97) итерации будут сходиться, если соответственно

$$h < \frac{1}{|\beta_s|L} \quad \text{и} \quad h < \frac{|\alpha_s|}{L}.$$

Итерационный процесс требует начального приближения решения нелинейной системы. Обычно его получают из явной схемы с точностью на порядок ниже. Явную схему называют предиктором (P), тогда как неявную — корректором (C), а итерационную процедуру в целом — предиктор–корректором. Кроме того,

как самостоятельный этап выделяют также вычисление функции правых частей уравнений \mathbf{k}_s (E). Символьно схема предиктор–корректор представляется как $P(EC)^l E$, если правая часть \mathbf{k}_s уточняется по исправленному решению корректора, или $P(EC)^l$ без уточнения правой части, где l — число исправлений решения в неявной схеме. Обычно $l = 2-4$.

Например, для метода трапеций (неявного метода Адамса второго порядка) в качестве предиктора используют метод Эйлера (явный метод Адамса первого порядка) и, следовательно, процедура предиктор–корректор $P(EC)^l$ имеет вид

$$P : \mathbf{x}_1^0 = \mathbf{x}_0 + h\mathbf{k}_0, \quad E : \mathbf{k}_1 = \mathbf{f}(\mathbf{x}_1^{l-1}),$$

$$C : \mathbf{x}_1^l = \mathbf{x}_0 + h \left[\frac{1}{2}\mathbf{k}_0 + \frac{1}{2}\mathbf{k}_1 \right].$$

Интересно, что первый шаг итерационной последовательности дает решение метода Хойна (4.17).

4.5.4. Линейные многошаговые методы

Таким образом, многошаговые методы можно представить в общем виде

$$\alpha_s \mathbf{x}_s + \dots + \alpha_0 \mathbf{x}_0 = h(\beta_s \mathbf{k}_s + \dots + \beta_0 \mathbf{k}_0). \quad (4.98)$$

При этом будем считать, что выполняются условия

$$\alpha_s \neq 0, \quad |\alpha_0| + |\beta_0| > 0.$$

Обычно полагают, что $\alpha_s = 1$. Если $\beta_s = 0$, метод (4.98) называется явным, иначе — неявным.

Методы вида (4.98) еще называют линейными многошаговыми, поскольку решения \mathbf{x} и функции \mathbf{k} входят в них линейным образом.

4.5.5. Порядок многошаговых методов

Многошаговый метод (4.98) имеет порядок p , если невязка ρ , которая получается после подстановки точного решения в (4.98), является величиной порядка h^{p+1} , т.е.

$$\rho \equiv \sum_{i=0}^s \alpha_i \mathbf{x}(t_i) - h \sum_{i=0}^s \beta_i \mathbf{x}'(t_i) = O(h^{p+1}). \quad (4.99)$$

Разложим решения $\mathbf{x}(t_i)$ в ряд Тейлора:

$$\mathbf{x}(t_i) = \mathbf{x}(t_0 + ih) = \sum_{j=0}^{p+1} \frac{(ih)^j}{j!} \mathbf{x}^{(j)}(t_0) + O(h^{p+2}). \quad (4.100)$$

Тогда для производных $\mathbf{x}'(t_i)$ будем иметь

$$\mathbf{x}'(t_i) = \mathbf{x}'(t_0 + ih) = \sum_{j=0}^p \frac{(ih)^j}{j!} \mathbf{x}^{(j+1)}(t_0) + O(h^{p+1}). \quad (4.101)$$

После подстановки (4.100) и (4.101) в (4.99) может быть получено разложение невязки по степеням h :

$$\begin{aligned} \rho = & \left(\sum_i \alpha_i \right) \mathbf{x}(t_0) + h \left(\sum_i i \alpha_i - \sum_i \beta_i \right) \mathbf{x}'(t_0) + \\ & + \frac{h^2}{2} \left(\sum_i i^2 \alpha_i - 2 \sum_i i \beta_i \right) \mathbf{x}''(t_0) + \\ & + \dots + \frac{h^p}{p!} \left(\sum_i i^p \alpha_i - p \sum_i i^{p-1} \beta_i \right) \mathbf{x}^{(p)}(t_0) + \\ & + \frac{h^{p+1}}{(p+1)!} \left(\sum_i i^{p+1} \alpha_i - (p+1) \sum_i i^p \beta_i \right) \mathbf{x}^{(p+1)}(t_0) + O(h^{p+2}). \end{aligned} \quad (4.102)$$

Отсюда, для того чтобы невязка ρ имела порядок h^{p+1} , все выражения в скобках первых $p + 1$ членов (4.102) должны равняться нулю. Таким образом, получаем условия порядка линейного многошагового метода

$$\begin{aligned} \sum_i \alpha_i &= 0, \\ \sum_i i\alpha_i - \sum_i \beta_i &= 0, \\ \dots & \\ \sum_i i^p \alpha_i - p \sum_i i^{p-1} \beta_i &= 0. \end{aligned} \tag{4.103}$$

Поскольку $p \geq 1$, то первые два условия должны выполняться всегда. Они образуют так называемое условие согласованности.

Найдем явный двушаговый метод максимально возможного порядка. Явный двушаговый метод можно представить в общем виде

$$\mathbf{x}_2 + \alpha_1 \mathbf{x}_1 + \alpha_0 \mathbf{x}_0 = h(\beta_1 \mathbf{k}_1 + \beta_0 \mathbf{k}_0) \quad (\alpha_2 = 1).$$

Из (4.103) выпишем первые четыре соотношения:

$$\begin{aligned} 1 + \alpha_1 + \alpha_0 &= 0, \\ 2 + \alpha_1 - \beta_0 - \beta_1 &= 0, \\ 4 + \alpha_1 - 2\beta_1 &= 0, \\ 8 + \alpha_1 - 3\beta_1 &= 0. \end{aligned}$$

Решая эту систему, получаем $\alpha_0 = -5$, $\alpha_1 = 4$, $\beta_0 = 2$, $\beta_1 = 4$. Пятое соотношение в (4.103)

$$16 + \alpha_1 - 4\beta_1 = 0$$

при найденных коэффициентах уже не выполняется. Поэтому двушаговый метод с полученными коэффициентами

$$\mathbf{x}_2 + 4\mathbf{x}_1 - 5\mathbf{x}_0 = h(4\mathbf{k}_1 + 2\mathbf{k}_0), \tag{4.104}$$

будет иметь третий порядок.

4.5.6. Устойчивость многошаговых методов

Ранее мы выяснили, что глобальная ошибка в методах Рунге–Кутты представляет собой величину порядка h^p : $\|\mathbf{E}\| = O(h^p)$. Это означает, что при уменьшении шага интегрирования глобальная точность численного решения повышается. Однако не для всех многошаговых методов это так, и уменьшение шага интегрирования приводит только к увеличению глобальной ошибки, несмотря на повышение локальной точности. Такая особенность многошаговых методов связана с понятием неустойчивости.

Важную роль в исследовании устойчивости схемы интегрирования играет поведение численного решения при $h \rightarrow 0$ и фиксированном $\Delta t = Nh$. Очевидно, при $h \rightarrow 0$ схема (4.98) для произвольного $(n + s)$ -го шага приводится к формуле

$$\alpha_s \mathbf{x}_{n+s} + \dots + \alpha_0 \mathbf{x}_n = \mathbf{0}. \quad (4.105)$$

Ее можно рассматривать как численную схему, примененную к решению уравнения

$$\mathbf{x}' = \mathbf{0}.$$

Представим решение как $\mathbf{x}_i = \mathbf{z}\zeta^i$ и подставим его в (4.105):

$$\mathbf{z}\zeta^n (\alpha_s \zeta^s + \dots + \alpha_0) = \mathbf{0}.$$

Отсюда ζ должно удовлетворять так называемому характеристическому уравнению

$$\alpha_s \zeta^s + \dots + \alpha_0 = 0. \quad (4.106)$$

Исследование характеристического уравнения усложняется, когда у него имеются кратные корни кратности m . В этом случае m решений для этого корня представляются в виде

$$\mathbf{x}_i = i^{j-1} \mathbf{z}\zeta^i \quad (j = 1, \dots, m). \quad (4.107)$$

Таким образом, если характеристический многочлен (4.106) имеет корни ζ_1, \dots, ζ_l кратностей m_1, \dots, m_l соответственно, то суперпозиция решений (4.107) будет давать общее решение уравнения

(4.105):

$$\mathbf{x}_i = \mathbf{p}_1(i)\zeta_1^i + \dots + \mathbf{p}_l(i)\zeta_l^i, \quad (4.108)$$

где $\mathbf{p}_j(i)$ — многочлены степеней $m_j - 1$, коэффициенты которых определяются из начальных условий. Поэтому для ограниченности решений \mathbf{x}_i необходимо потребовать, чтобы, во-первых, корни ζ лежали внутри единичной окружности ($|\zeta| < 1$) и, во-вторых, чтобы все корни на единичной окружности ($|\zeta| = 1$) были простыми. Многошаговые методы, удовлетворяющие этим требованиям, называются устойчивыми по Далквисту или D -устойчивыми (Hairer et al., 1993).

Исследуем на устойчивость метод (4.104). Характеристический многочлен для него запишется как

$$\zeta^2 + 4\zeta - 5 = 0.$$

Его корни $\zeta_1 = -5$, $\zeta_2 = 1$. Так как $|\zeta_1| > 1$, многошаговый метод (4.104) неустойчив. Между тем все методы Адамса (4.93) D -устойчивые, поскольку только один корень их характеристических многочленов равен единице, тогда как остальные — нулю.

4.5.7. Наивысший достижимый порядок для устойчивых методов

В условиях порядка (4.103) содержатся $2s + 1$ неизвестных ($\alpha_s = 1$). Для их определения необходимо $2s + 1$ условий. Следовательно, после определения неизвестных коэффициентов мы получим метод порядка $p = 2s$. Однако такие методы не имеют практической ценности, поскольку они неустойчивы. Г. Далквист (Dahlquist, 1956) показал, что порядок p устойчивого линейного s -шагового метода подчиняется ограничениям:

$$\begin{aligned} p &\leq s + 2 && \text{при четных } s, \\ p &\leq s + 1 && \text{при нечетных } s, \\ p &\leq s && \text{при } \beta_s/\alpha_s \leq 0. \end{aligned}$$

4.5.8. Практическая оценка локальной погрешности

Локальная погрешность многошагового метода может быть оценена в сопоставлении его решения на шаге с решением другого метода того же порядка. Согласно (4.99) и (4.102), для метода порядка p будем иметь

$$\begin{aligned} \mathbf{x}(t_s) &= \frac{1}{\alpha_s} \left(h \sum_{i=0}^s \beta_i \mathbf{x}'(t_i) - \sum_{i=0}^{s-1} \alpha_i \mathbf{x}(t_i) \right) + \frac{\boldsymbol{\rho}}{\alpha_s} = \\ &= \mathbf{x}_s + C_{p+1} \frac{h^{p+1}}{(p+1)!} \mathbf{x}^{(p+1)}(t_0) + O(h^{p+2}), \end{aligned} \quad (4.109)$$

где

$$C_{p+1} = \frac{1}{\alpha_s} \left(\sum_i i^{p+1} \alpha_i - (p+1) \sum_i i^p \beta_i \right).$$

Решение вспомогательного метода $\hat{\mathbf{x}}_s$ того же порядка также будет удовлетворять соотношению (4.109), но со своей константой \hat{C}_{p+1} . Отсюда, игнорируя члены порядка h^{p+2} , получаем оценку локальной погрешности (формулу Милна) (Milne, 1926)

$$\|\mathbf{e}\|_{cal} = \|\mathbf{x}_s - \mathbf{x}(t_s)\| = \frac{|C_{p+1}|}{|C_{p+1} - \hat{C}_{p+1}|} \|\mathbf{x}_s - \hat{\mathbf{x}}_s\|, \quad (4.110)$$

а также уточненное решение

$$\mathbf{x}_s - \frac{C_{p+1}}{C_{p+1} - \hat{C}_{p+1}} (\mathbf{x}_s - \hat{\mathbf{x}}_s) = \mathbf{x}(t_s) + O(h^{p+2}).$$

4.5.9. Выбор шага

Изменение величины шага для эффективного численного интегрирования в многошаговых методах (4.98) гораздо сложнее, чем в одношаговых. Это связано с тем, что при использовании многошаговой схемы интегрирования получение каждого следующего

решения на новом шаге предполагает перевычисление предыдущих решений на новой равномерной сетке с соответствующей величиной шага. Для этого, как правило, прибегают к интерполяционным многочленам. Главный недостаток такого подхода состоит в том, что вместе с решениями необходимо перевычислять и значения правых частей уравнений. Другой подход — это построение многошаговых схем интегрирования на неравномерной сетке.

Если в явном методе Адамса в качестве приближения подынтегральной функции выбрать полином Ньютона²⁹

$$\mathbf{g}(t) = \mathbf{k}_{s-1} + \mathbf{k}_{s-1,s-2}(t-t_{s-1}) + \dots + \mathbf{k}_{s-1,s-2,\dots,0}(t-t_{s-1}) \dots (t-t_1),$$

($\mathbf{k}_{s-1,\dots,s-1-i}$ — разделенная разность i -го порядка, определяемая на сетке $t_{s-1}, \dots, t_{s-1-i}$), то обобщение метода на случай переменного шага дает схему интегрирования

$$\mathbf{x}_s = \mathbf{x}_{s-1} + \sum_{i=0}^{s-1} \gamma_{i1} \mathbf{k}_{s-1,s-2,\dots,s-1-i}, \quad (4.111)$$

где интегралы

$$\gamma_{i1} = \int_{t_{s-1}}^{t_s} (t-t_{s-1}) \dots (t-t_{s-i}) dt$$

вычисляются рекуррентно по формулам (Холл, Уатт, 1979)

$$\gamma_{0j} = (t_s - t_{s-1})^j / j! \quad (j = 1, \dots, s+1),$$

$$\gamma_{ij} = (t_s - t_{s-i}) \gamma_{i-1,j} - j \gamma_{i-1,j+1} \quad (i = 1, \dots, s; j = 1, \dots, s-i+1).$$

На практике (4.111) удобно использовать в тандеме с неявными схемами интегрирования. В неявных методах с переменным шагом выбирается полином

$$\mathbf{g}(t) = \mathbf{k}_{s-1} + \mathbf{k}_{s-1,s-2}(t-t_{s-1}) + \dots + \mathbf{k}_{s,s-1,\dots,0}(t-t_{s-1}) \dots (t-t_0).$$

²⁹Для вычисления разделенных разностей здесь могут быть использованы рекуррентные формулы (4.61).

Тогда решение можно представить как

$$\mathbf{x}_s = \mathbf{x}_s^C = \mathbf{x}_s^P + \gamma_{s1} \mathbf{k}_{s,s-1,\dots,0}, \quad (4.112)$$

где \mathbf{x}_s^P — решение-предиктор (4.111).

Удобство тандема предиктор–корректор (4.111) и (4.112) в данном случае состоит в том, что корректирующий член в (4.112) можно использовать для контроля шага, поскольку он представляет собой оценку ошибки явного метода, т.е.

$$\|\mathbf{e}\|_{cal} = \|\gamma_{s1} \mathbf{k}_{s,s-1,\dots,0}\|.$$

Кроме того, многошаговые схемы интегрирования Адамса на неравномерной сетке позволяют отказаться от одношаговых схем, необходимых для получения начальных решений. Покажем, как это возможно.

В начале интегрирования вычислим правую часть уравнения $\mathbf{k}_0 = \mathbf{f}(t_0, \mathbf{x}_0)$ и выберем начальную величину шага $h_1 = \|\mathbf{e}\|_{tol}/\|\mathbf{k}_0\|$, которая формально, если можно так сказать, обеспечивает задаваемую точность как для метода нулевого порядка. Используя предиктор–корректор для $s = 1$ находим оценку локальной погрешности $\|\mathbf{e}\|_{cal}$ предиктора (метода Эйлера). Уточняя по этой оценке величину h_1 , с тем чтобы она соответствовала задаваемой точности предиктора, получим второе решение \mathbf{x}_1 и правую часть \mathbf{k}_1 . Этот этап можно представить в виде схемы

$$\begin{array}{ccccccc} \mathbf{x}_0 & \rightarrow & h_1 = \|\mathbf{e}\|_{tol}/\|\mathbf{k}_0\| & \rightarrow & P(EC)_{s=1}^I & \rightarrow & \mathbf{x}_1, \mathbf{k}_1 \\ & & & & \uparrow & & \downarrow \\ & & & & h_1 & \leftarrow & \|\mathbf{e}\|_{cal} \end{array} \quad (4.113)$$

Затем, используя предиктор–корректор для $s = 2$ с величиной шага $h_2 = h_1$, получаем \mathbf{x}_2 и \mathbf{k}_2 вместе с очередной оценкой $\|\mathbf{e}\|_{cal}$, которую уже применяем для определения следующего шага h_3 ; после, используя предиктор–корректор для $s = 3$, находим \mathbf{x}_3 , \mathbf{k}_3 и h_4 , и так далее, пока не наберем нужное количество правых

частей уравнения для заданного порядка метода:

$$\begin{aligned} \mathbf{k}_0, \mathbf{k}_1, \mathbf{x}_1, h_2 &\rightarrow P(EC)_{s=2}^l \rightarrow \mathbf{x}_2, \mathbf{k}_2 \rightarrow h_3, \\ \mathbf{k}_0, \mathbf{k}_1, \mathbf{k}_2, \mathbf{x}_2, h_3 &\rightarrow P(EC)_{s=3}^l \rightarrow \mathbf{x}_3, \mathbf{k}_3 \rightarrow h_4, \\ &\dots \end{aligned} \quad (4.114)$$

При этом, очевидно, локальные ошибки получаемых приближенных решений не будут превышать величину $\|\mathbf{e}\|_{tol}$, в то время как многошаговая схема при сборке с увеличением порядка будет плавно выходить на величину шага, обеспечивающую задаваемую локальную точность.

4.5.10. Порядок и шаг интегрирования при компьютерной реализации методов Адамса

Как и в случае неявных методов Рунге–Кутты (см. подразд. 4.3.8), шаг интегрирования h в зависимости от выбранного порядка p в многошаговых методах имеет ограничения: с одной стороны, он должен быть достаточно большим, чтобы обеспечивать методическую ошибку, не меньшую вычислительной; с другой стороны, он должен быть достаточно малым, чтобы схема интегрирования оставалась устойчивой. Оценим налагаемые на шаг пределы для методов Адамса–Мульттона, как наиболее используемых на практике, применительно к круговой задаче.

Характеристическое уравнение типа (4.106), соответствующее комплексному дифференциальному уравнению кругового движения (4.80), для метода Адамса–Мульттона (4.93) порядка $p = s + 1$ можно представить в виде

$$i\zeta^s - i\zeta^{s-1} + \nu h(\beta_s \zeta^s + \dots + \beta_0) = 0. \quad (4.115)$$

Как видно, корни характеристического уравнения (4.115) зависят от величины $h_\varphi = \nu h$. Между тем устойчивость схемы интегрирования имеет место, когда корни характеристического уравнения находятся внутри единичного круга на комплексной плоскости.

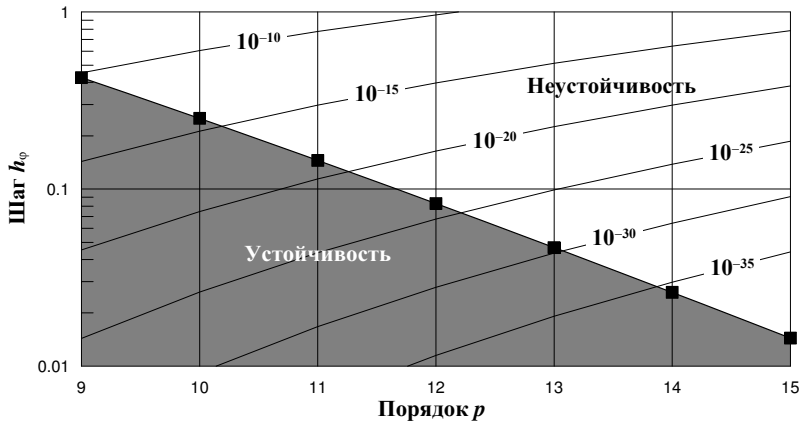


Рис. 4.2. Устойчивость методов Адамса–Мульттона в зависимости от порядка и шага интегрирования в круговой задаче. Кривыми линиями показаны уровни методической точности

Это условие фактически определяет верхние пределы допустимых значений h_ϕ . Некоторые из них для $p = 9\text{--}15$ показаны на рис. 4.2 (черными квадратиками).

Важно заметить, что с увеличением порядка метода верхние пределы стремительно уменьшаются. Отсюда, в частности, это приводит к тому, что при вычислениях в компьютерной арифметике с двойной точностью не имеет смысла использовать методы Адамса–Мульттона выше 10-го порядка, а в арифметике с четверной точностью — выше 13-го: более высокие порядки при устойчивом интегрировании обеспечивают излишне высокую методическую точность (см. кривые линии), которая превышает вычислительную. Скажем, если в арифметике с двойной точностью использовать метод 11-го порядка, то результаты, выше вычислительной точности, получить не удастся. Однако результаты той же точности можно получить методом 10-го порядка, причем с большим шагом интегрирования.

4.6. Геометрические методы

В последнее время в динамической астрономии возрос интерес к так называемым геометрическим методам численного интегрирования. Учитывая геометрические свойства дифференциальных уравнений, а точнее, их решений, эти методы позволяют качественно улучшить результаты численного интегрирования, в особенности, если оно выполняется на длительных (космогонических) интервалах времени.

Напомним сначала, что канонической системой уравнений называется система вида

$$\mathbf{x}' = \frac{\partial g}{\partial \mathbf{y}}, \quad \mathbf{y}' = -\frac{\partial g}{\partial \mathbf{x}}, \quad (4.116)$$

где \mathbf{x} и \mathbf{y} — канонические m -мерные векторы, а $g = g(\mathbf{x}, \mathbf{y})$ — гамильтониан, который при этом является интегралом канонической системы: $g(\mathbf{x}, \mathbf{y}) = \text{const}$. Отметим важное для нас свойство канонической системы: преобразование временного сдвига

$$\mathbf{x}(t) \xrightarrow{h} \bar{\mathbf{x}}(t) = \mathbf{x}(t+h), \quad \mathbf{y}(t) \xrightarrow{h} \bar{\mathbf{y}}(t) = \mathbf{y}(t+h) \quad (4.117)$$

является симплектическим (каноническим). Действительно, ведь новые переменные также удовлетворяют системе (4.116).

Кроме того, заметим, что все рассматриваемые нами дифференциальные уравнения обратимы по времени. Это означает, что выполненные последовательно прямое и обратное интегрирования задачи Коши с одним и тем же шагом h дают ее начальные условия:

$$\mathbf{x}(t_0) \xrightarrow{h} \mathbf{x}(t_0+h) \xrightarrow{-h} \mathbf{x}(t_0). \quad (4.118)$$

4.6.1. Уравнения гармонического осциллятора

Рассмотрим уравнение гармонического осциллятора

$$x'' = -x, \quad (4.119)$$

которое представимо в виде системы уравнений первого порядка

$$x' = y, \quad y' = -x. \quad (4.120)$$

Отметим важные свойства системы (4.120): 1) она имеет интеграл

$$g(x, y) = \frac{1}{2}(x^2 + y^2) = \text{const}; \quad (4.121)$$

2) система канонична с гамильтонианом g ; 3) она обратима по времени.

Таким образом, для интегрирования системы (4.120) целесообразно применять такие методы, которые бы учитывали названные свойства системы. Поскольку эти свойства тем или иным образом связаны с геометрией решений системы, то методы интегрирования, учитывающие хотя бы одно из них, обычно называют геометрическими (Hairer et al., 2002).

Интересно заметить, что свойства 1–3 характерны для многих дифференциальных уравнений небесной механики. В частности, уравнения кеплеровского движения (относительно центральной массы с гравитационным параметром μ) представимы в виде

$$\mathbf{x}' = \dot{\mathbf{x}} = \frac{\partial g}{\partial \dot{\mathbf{x}}}, \quad \dot{\mathbf{x}}' = -\mu \frac{\mathbf{x}}{|\mathbf{x}|^3} = -\frac{\partial g}{\partial \mathbf{x}}, \quad \text{где} \quad g = \frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|} = \text{const};$$

т.е. они каноничны, следовательно, имеют интеграл энергии, а также обратимы по времени.

4.6.2. Методы Эйлера

Применим теперь к уравнениям (4.120) явный и неявный методы Эйлера:

$$\begin{aligned} x_{n+1} &= x_n + hy_n, & y_{n+1} &= y_n - hx_n; \\ x_{n+1} &= x_n + hy_{n+1}, & y_{n+1} &= y_n - hx_{n+1}. \end{aligned} \quad (4.122)$$

Разрешение в неявной схеме $(n+1)$ -го решения относительно n -го дает

$$x_{n+1} = \frac{x_n + hy_n}{1 + h^2}, \quad y_{n+1} = \frac{y_n - hx_n}{1 + h^2}.$$

Отсюда видно, что результаты интегрирования, получаемые методами Эйлера на одном шаге, совпадают с точностью до коэффициента $1/(1+h^2)$. Подставляя решения в интегральное соотношение $g(x, y)$ (4.121), получаем для явного метода

$$g(x_{n+1}, y_{n+1}) = (1+h^2)g(x_n, y_n) = (1+h^2)^{n+1}g(x_0, y_0)$$

и для неявного

$$g(x_{n+1}, y_{n+1}) = \frac{g(x_n, y_n)}{1+h^2} = \frac{g(x_0, y_0)}{(1+h^2)^{n+1}}.$$

Как видно, при пошаговом интегрировании методами Эйлера интегральное соотношение не сохраняется: в явном случае оно увеличивается, тогда как в неявном уменьшается.

Далее, поскольку преобразование сдвига по времени (4.117) симплектично, то пошаговое отображение в соответствии со схемой

$$x_n \xrightarrow{h} x_{n+1}, \quad y_n \xrightarrow{h} y_{n+1} \quad (4.123)$$

также должно быть симплектическим. Проверим на симплектичность схемы Эйлера. Вычислим скобки Пуассона

$$\{x_{n+1}, y_{n+1}\} = \frac{\partial x_{n+1}}{\partial x_n} \frac{\partial y_{n+1}}{\partial y_n} - \frac{\partial x_{n+1}}{\partial y_n} \frac{\partial y_{n+1}}{\partial x_n}. \quad (4.124)$$

Тогда для явного и неявного методов получаем $\{x_{n+1}, y_{n+1}\} = 1+h^2$ и $\{x_{n+1}, y_{n+1}\} = 1/(1+h^2)$ соответственно. Вместе с тем известно, что преобразование (4.123) симплектично, если скобка Пуассона равна единице. Следовательно, методы Эйлера не симплектичны.

Обратимость по времени (4.118) уравнений (4.120) требует, чтобы схема интегрирования при ее последовательном использовании сначала с шагом h , а затем с шагом $-h$ давала исходные результаты. Однако, согласно схемам Эйлера (4.122), будем иметь в явном случае

$$x_n \xrightarrow{h} x_n + hy_n \xrightarrow{-h} (1+h^2)x_n, \quad y_n \xrightarrow{h} y_n - hx_n \xrightarrow{-h} (1+h^2)y_n$$

и в неявном случае

$$x_n \xrightarrow{h} \frac{x_n + hy_n}{1 + h^2} \xrightarrow{-h} \frac{x_n}{1 + h^2}, \quad y_n \xrightarrow{h} \frac{y_n - hx_n}{1 + h^2} \xrightarrow{-h} \frac{y_n}{1 + h^2}.$$

Следовательно, схемы Эйлера не обратимы по времени.

Таким образом, методы Эйлера не геометрические. Тем не менее они могут стать таковыми после введения ряда незначительных модификаций.

Простым способом сохранения интегрального соотношения является возвращение численного решения каждого шага на интегральную кривую, задаваемую уравнением $g(x, y) = g_0 = g(x_0, y_0)$. Как видно из (4.121), интегральные кривые уравнений гармонического осциллятора представляют собой концентрические окружности с центром в начале координат и радиусами

$$r_0 = \sqrt{2g_0} = \sqrt{x_0^2 + y_0^2},$$

определяемыми начальными условиями. Перемещения численных решений Δx_{n+1} и Δy_{n+1} будем осуществлять вдоль нормали к интегральной кривой в точке (x_{n+1}, y_{n+1}) , т.е. вдоль вектора положения этой точки в фазовом пространстве. Тогда из геометрических соображений нетрудно получить поправки к решениям:

$$\Delta x_{n+1} = -\frac{r_{n+1} - r_0}{r_{n+1}} x_{n+1}, \quad \Delta y_{n+1} = -\frac{r_{n+1} - r_0}{r_{n+1}} y_{n+1}, \quad (4.125)$$

где

$$r_{n+1} = \sqrt{2g_{n+1}} = \sqrt{x_{n+1}^2 + y_{n+1}^2}.$$

Интересно, что простой формальной заменой в схемах (4.122) первых (или вторых) частей для приближенных решений x_{n+1} (или y_{n+1}) удастся получить две явные симплектические схемы интегрирования

$$\begin{aligned} y_{n+1} &= y_n - hx_n, & x_{n+1} &= x_n + hy_{n+1}; \\ x_{n+1} &= x_n + hy_n, & y_{n+1} &= y_n - hx_{n+1}. \end{aligned} \quad (4.126)$$

Симплектичность модифицированных методов Эйлера несложно проверить, вычислив, как и ранее, скобки Пуассона (4.124), которые оказываются равными единице в обоих случаях.

Комбинируя схемы методов Эйлера, можно получить новую, удовлетворяющую условию обратимости по времени (4.118): применим сначала явную схему Эйлера с шагом $h/2$, а затем — неявную с тем же шагом, в итоге будем иметь схему метода трапеций³⁰

$$x_{n+1} = x_n + \frac{h}{2}(y_n + y_{n+1}), \quad y_{n+1} = y_n - \frac{h}{2}(x_n + x_{n+1}).$$

Поскольку метод трапеций симметричен, он будет удовлетворять условию обратимости (4.118)³¹ (см. подразд. 4.6.4).

4.6.3. Проекционный метод

Для сохранения интегралов систем дифференциальных уравнений часто применяют так называемый проекционный метод, который, по-видимому, впервые предложил П. Накози в 1971 г. (Nacozy, 1971)³². Заметим, что исследуемая ранее система (4.120) вместе с интегральным соотношением (4.121) являются частным случаем дифференциально-алгебраических систем уравнений, которые имеют вид

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad g(\mathbf{x}) = g(\mathbf{x}_0) = g_0 = \text{const},$$

где в качестве алгебраического уравнения обычно выступают интегралы системы дифференциальных уравнений либо налагаемые на нее связи. Рассмотрим проекционный метод на примере этой системы.

Идея метода состоит в том, чтобы после каждой стандартной процедуры интегрирования на шаге вносить в численное решение \mathbf{x}_{n+1} поправки $\Delta \mathbf{x}_{n+1}$ за отклонение $g(\mathbf{x}_{n+1})$ относительно g_0 ,

³⁰ Впрочем, здесь она является также и схемой метода средней точки.

³¹ Интересно заметить, что он, кроме того, сохраняет интеграл (4.121).

³² Мы уже рассматривали этот метод в подразд. 3.3.5.

что с геометрической точки зрения равносильно проекции решения \mathbf{x}_{n+1} в фазовом пространстве на гиперповерхность $g(\mathbf{x}) = g_0$, т.е. исправленное решение должно удовлетворять алгебраическому уравнению

$$g(\mathbf{x}_{n+1} + \Delta\mathbf{x}_{n+1}) = g_0.$$

С точностью до малых первого порядка имеем

$$g(\mathbf{x}_{n+1}) + \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_{n+1}) \cdot \Delta\mathbf{x}_{n+1} = g_0. \quad (4.127)$$

Чтобы обеспечить наименьшую величину поправки $\Delta\mathbf{x}_{n+1}$ в соответствии с (4.127), ее направляют вдоль нормали к поверхности $g(\mathbf{x}) = g(\mathbf{x}_{n+1})$ в точке \mathbf{x}_{n+1} , представляя, таким образом, в виде

$$\Delta\mathbf{x}_{n+1} = \alpha \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_{n+1}). \quad (4.128)$$

Коэффициент α определяется после подстановки (4.128) в (4.127). В итоге будем иметь поправку

$$\Delta\mathbf{x}_{n+1} = -\Delta g \left[\frac{\partial g}{\partial \mathbf{x}} \left(\frac{\partial g}{\partial \mathbf{x}} \cdot \frac{\partial g}{\partial \mathbf{x}} \right)^{-1} \right] (\mathbf{x}_{n+1}), \quad \Delta g = g(\mathbf{x}_{n+1}) - g_0. \quad (4.129)$$

Поскольку (4.129) получается из приближенного соотношения (4.127), то исправленное решение может еще не удовлетворять алгебраическому уравнению. В этом случае следует повторять процедуру уточнения решения (4.129) до тех пор, пока поправка $\Delta\mathbf{x}_{n+1}$ не станет достаточно малой.

Интересно заметить, что если мы выполним в (4.125) приближенную замену

$$r_{n+1} - r_0 \approx \frac{g_{n+1} - g_0}{r_{n+1}},$$

то получим коррекцию проекционного метода (4.129) применительно к дифференциально-алгебраической системе гармонического осциллятора (4.120) и (4.121).

4.6.4. Симплектические и симметричные методы

Простые симплектические методы. Симплектические схемы интегрирования строятся таким образом, чтобы отображение

$$\mathbf{x}_n \xrightarrow{h} \mathbf{x}_{n+1}, \quad \mathbf{y}_n \xrightarrow{h} \mathbf{y}_{n+1}$$

было каноническим. Тогда должна существовать производящая функция, например, $W = W(\mathbf{x}_n, \mathbf{y}_{n+1})$ либо $W = W(\mathbf{x}_{n+1}, \mathbf{y}_n)$, для которой

$$\mathbf{x}_{n+1} = \frac{\partial W}{\partial \mathbf{y}_{n+1}}, \quad \mathbf{y}_n = \frac{\partial W}{\partial \mathbf{x}_n}; \quad \text{либо} \quad \mathbf{x}_n = \frac{\partial W}{\partial \mathbf{y}_n}, \quad \mathbf{y}_{n+1} = \frac{\partial W}{\partial \mathbf{x}_{n+1}}.$$

Выберем в качестве таких функций (Yoshida, 1993)

$$W(\mathbf{x}_n, \mathbf{y}_{n+1}) = \mathbf{x}_n \cdot \mathbf{y}_{n+1} + hg(\mathbf{x}_n, \mathbf{y}_{n+1}),$$

$$W(\mathbf{x}_{n+1}, \mathbf{y}_n) = \mathbf{x}_{n+1} \cdot \mathbf{y}_n - hg(\mathbf{x}_{n+1}, \mathbf{y}_n),$$

тогда получаем симплектические схемы первого порядка

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{x}_n + h \frac{\partial g}{\partial \mathbf{y}_{n+1}}, & \mathbf{y}_{n+1} &= \mathbf{y}_n - h \frac{\partial g}{\partial \mathbf{x}_n}; \\ \mathbf{x}_{n+1} &= \mathbf{x}_n + h \frac{\partial g}{\partial \mathbf{y}_n}, & \mathbf{y}_{n+1} &= \mathbf{y}_n - h \frac{\partial g}{\partial \mathbf{x}_{n+1}}, \end{aligned} \quad (4.130)$$

в которых без труда узнаем первую и вторую симплектические схемы Эйлера (4.126) в случае интегрирования уравнений гармонического осциллятора, если вместо g подставить (4.121).

Для метода второго порядка выберем (Yoshida, 1993)

$$W(\mathbf{x}_n, \mathbf{y}_{n+1}) = \mathbf{x}_n \cdot \mathbf{y}_{n+1} + hg(\mathbf{x}_n, \mathbf{y}_{n+1}) + \frac{h^2}{2} \frac{\partial g}{\partial \mathbf{y}_{n+1}} \cdot \frac{\partial g}{\partial \mathbf{x}_n}.$$

Тогда данной производящей функции будет соответствовать неявный метод

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{x}_n + h \frac{\partial g}{\partial \mathbf{y}_{n+1}} + \frac{h^2}{2} \left[\frac{\partial^2 g}{\partial \mathbf{y}_{n+1}^2} \frac{\partial g}{\partial \mathbf{x}_n} + \left(\frac{\partial g}{\partial \mathbf{y}_{n+1}} \right)^T \frac{\partial^2 g}{\partial \mathbf{y}_{n+1} \partial \mathbf{x}_n} \right], \\ \mathbf{y}_{n+1} &= \mathbf{y}_n - h \frac{\partial g}{\partial \mathbf{x}_n} - \frac{h^2}{2} \left[\frac{\partial^2 g}{\partial \mathbf{y}_{n+1} \partial \mathbf{x}_n} \frac{\partial g}{\partial \mathbf{x}_n} + \left(\frac{\partial g}{\partial \mathbf{y}_{n+1}} \right)^T \frac{\partial^2 g}{\partial \mathbf{x}_n^2} \right]. \end{aligned}$$

Методы более высоких порядков можно конструировать подобным образом. Однако очевидно, что такой подход не столь эффективен, поскольку требует получения формул для частных производных высоких порядков от гамильтониана по каноническим переменным.

Рассмотрим теперь важный для классической механики случай, когда $g = \mathbf{y}^2/2 + U(\mathbf{x})$, где \mathbf{x} и \mathbf{y} — векторы положения и скорости соответственно. Симплектические схемы (4.130) тогда примут вид

$$\begin{aligned} \mathbf{y}_{n+1} &= \mathbf{y}_n - h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n), & \mathbf{x}_{n+1} &= \mathbf{x}_n + h\mathbf{y}_{n+1}; \\ \mathbf{x}_{n+1} &= \mathbf{x}_n + h\mathbf{y}_n, & \mathbf{y}_{n+1} &= \mathbf{y}_n - h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1}). \end{aligned} \quad (4.131)$$

Если за невозмущенное принять решение задачи при отсутствии сил ($U = 0$), то применение первой симплектической схемы можно рассматривать как коррекцию скорости за возмущение с последующим нахождением точного решения для положения и скорости невозмущенной задачи³³. Во второй схеме, наоборот, сначала находится точное решение для положения и скорости невозмущенной задачи, а затем корректируется скорость за возмущение. Такой взгляд позволяет сконструировать более точные симплектические схемы первого порядка в некоторых частных случаях.

Рассмотрим возмущенный гамильтониан $g = \mathbf{y}^2/2 + U_K(\mathbf{x}) + \mu U(\mathbf{x})$ с возмущением $\mu U(\mathbf{x})$ такой, что для невозмущенного гамильтониана $g_K = \mathbf{y}^2/2 + U_K(\mathbf{x})$ существует точное решение

$$\mathbf{x}_1 = \mathbf{x}(t_0 + h) = \Phi_h(\mathbf{x}_0, \mathbf{y}_0), \quad \mathbf{y}_1 = \mathbf{y}(t_0 + h) = \Psi_h(\mathbf{x}_0, \mathbf{y}_0); \quad (4.132)$$

например, решение задачи двух тел с кеплеровской энергией g_K . Здесь μ — малый параметр возмущающих сил с потенциальной функцией $\mu U(\mathbf{x})$. Тогда симплектические схемы первого порядка,

³³Положение в невозмущенной задаче со временем меняется линейно, а скорость — постоянна.

подобные (4.131), можно представить в виде

$$\begin{aligned} \mathbf{y}_{n+1} &= \Psi_h \left(\mathbf{x}_n, \mathbf{y}_n - \mu h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n) \right), \\ \mathbf{x}_{n+1} &= \Phi_h \left(\mathbf{x}_n, \mathbf{y}_n - \mu h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n) \right); \\ \mathbf{x}_{n+1} &= \Phi_h(\mathbf{x}_n, \mathbf{y}_n), \quad \mathbf{y}_{n+1} = \Psi_h(\mathbf{x}_n, \mathbf{y}_n) - \mu h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1}). \end{aligned}$$

Примечательно, что для этих схем величина локальной ошибки будет порядка μh^2 , т.е. она пропорциональна величине возмущающих сил. Следовательно, чем слабее возмущения, тем меньше методическая ошибка.

Простые симметричные методы. Условию обратимости (4.118) отвечают симметричные методы. Для их построения часто прибегают к комбинированию обычных методов Рунге–Кутты с их сопряженными методами.

Сопряженным к исходному методу Рунге–Кутты (4.39) называется метод, получаемый в результате замены в исходном методе (Hairer et al., 1993)

$$\mathbf{x}_0 \leftrightarrow \mathbf{x}_1, \quad h \leftrightarrow -h, \quad t_0 \leftrightarrow t_0 + h. \quad (4.133)$$

Очевидно, повторная замена дает исходный метод, иначе говоря, сопряженный к сопряженному методу является исходным.

Представим исходную схему интегрирования как отображение $\mathbf{x}_1 = \Phi_h(\mathbf{x}_0)$, а сопряженную — как $\mathbf{x}_1 = \Phi_h^*(\mathbf{x}_0)$. После замены (4.133) в исходной схеме будем иметь $\mathbf{x}_0 = \Phi_{-h}(\mathbf{x}_1)$. Отсюда $\mathbf{x}_1 = \Phi_{-h}^{-1}(\mathbf{x}_0)$. Следовательно, сопряженный метод фактически представляет обратное отображение с обратным шагом, т.е. $\Phi_h^* = \Phi_{-h}^{-1}$. Отсюда $\Phi_{-h}^* \circ \Phi_h = \Phi_h^{-1} \circ \Phi_h = \mathbf{I}$, где \mathbf{I} — тождественное преобразование, и это, в частности, означает, что прямое и обратное интегрирование симметричным методом ($\Phi_h^* = \Phi_h$) с шагом произвольной величины h также представляет собой тождественное преобразование.

Сопряженный метод также является методом Рунге–Кутты, коэффициенты которого a_{ij}^* , b_i^* и c_i^* связаны с коэффициентами исходного метода a_{ij} , b_i и c_i как

$$a_{ij}^* = b_{s+1-j} - a_{s+1-i, s+1-j}, \quad b_i^* = b_{s+1-i}, \quad c_i^* = 1 - c_{s+1-i} \quad (4.134)$$

($i, j = 1, \dots, s$). Действительно, выполнив замену (4.133) в методе (4.39), приводим его к сопряженному виду

$$\mathbf{x}_1 = \mathbf{x}_0 + h \sum_{j=1}^s b_j \mathbf{k}_j,$$

$$\mathbf{k}_i = \mathbf{f}(t_0 + h(1 - c_i), \mathbf{x}_0 + h \sum_{j=1}^s (b_j - a_{ij}) \mathbf{k}_j) \quad (i = 1, \dots, s).$$

Чтобы упорядочить в формуле \mathbf{k}_i по возрастанию коэффициентов c_i^* , нам остается произвести замену индексов $i \rightarrow s + 1 - i$ и $j \rightarrow s + 1 - j$. В итоге находим связь между коэффициентами (4.134).

Каков же порядок сопряженного метода? Чтобы выяснить это, оценим его локальную ошибку. Применим исходную схему интегрирования к точному решению $\mathbf{x}(t_0)$, а сопряженную схему с обратным шагом — к $\mathbf{x}(t_1)$:

$$\mathbf{x}_1 = \Phi_h(\mathbf{x}(t_0)), \quad \mathbf{x}_0^* = \Phi_{-h}^*(\mathbf{x}(t_1)).$$

Исходная схема представляет точное решение $\mathbf{x}(t_1)$ с ошибкой

$$\mathbf{e}_1 = \mathbf{x}(t_1) - \mathbf{x}_1 = \mathbf{C}(t_0)h^{p+1} + O(h^{p+2}).$$

Далее, поскольку, как и для любого метода Рунге–Кутты,

$$\Phi_{-h}^* = \mathbf{I} + O(h),$$

а $\Phi_{-h}^* = \Phi_h^{-1}$, то сопряженная схема будет представлять точное решение $\mathbf{x}(t_0)$ с ошибкой

$$\mathbf{e}_0^* = \mathbf{x}(t_0) - \mathbf{x}_0^* = \Phi_{-h}^*(\mathbf{x}_1) - \Phi_{-h}^*(\mathbf{x}(t_1)) = -\mathbf{C}(t_0)h^{p+1} + O(h^{p+2}).$$

Тогда для $\mathbf{x}_1^* = \Phi_h^*(\mathbf{x}(t_0))$ имеем (Hairer et al., 2002)

$$\mathbf{e}_1^* = (-1)^{p+2} \mathbf{C}(t_1) h^{p+1} + O(h^{p+2}) = (-1)^p \mathbf{C}(t_0) h^{p+1} + O(h^{p+2}).$$

Следовательно, сопряженный метод имеет тот же порядок, что и исходный метод.

Нетрудно видеть, что сопряженный к явному методу Эйлера будет неявный, и наоборот. Интересно, что и симплектические методы Эйлера (4.130) также взаимосопряженные. В то же время метод средней точки

$$\mathbf{x}_1 = \mathbf{x}_0 + h\mathbf{f}\left(t_0 + \frac{h}{2}, \frac{\mathbf{x}_0 + \mathbf{x}_1}{2}\right)$$

и метод трапеций

$$\mathbf{x}_1 = \mathbf{x}_0 + \frac{h}{2}[\mathbf{f}(t_0, \mathbf{x}_0) + \mathbf{f}(t_1, \mathbf{x}_1)]$$

остаются без изменения относительно преобразований (4.133). Такие методы, называются симметричными (см. подразд. 4.4.3).

Для симметричного метода $\Phi_h^* = \Phi_h$, поэтому $\mathbf{e}_1^* = \mathbf{e}_1$ и, в частности, $(-1)^p \mathbf{C}(t_0) = \mathbf{C}(t_0)$, что имеет место при четном p . Следовательно, любой симметричный метод должен иметь четный порядок.

Введение сопряженных методов позволяет путем комбинирования их с исходными очень просто конструировать симметричные методы. Действительно, последовательное использование на половине шага сначала сопряженного метода, а затем исходного: $\Phi_{h/2} \circ \Phi_{h/2}^*$, или наоборот: $\Phi_{h/2}^* \circ \Phi_{h/2}$, дает симметричный метод. Свойство симметричности проверяется следующим образом. Рассмотрим первый метод $\mathbf{x}_1 = \Phi_{h/2} \circ \Phi_{h/2}^*(\mathbf{x}_0)$. Выполним преобразование переменных (4.133) и разрешим сопряженный метод относительно решения \mathbf{x}_1 :

$$\mathbf{x}_0 = \Phi_{-h/2} \circ \Phi_{-h/2}^*(\mathbf{x}_1), \quad \text{тогда} \quad \mathbf{x}_1 = (\Phi_{-h/2}^*)^{-1} \circ \Phi_{-h/2}^{-1}(\mathbf{x}_0).$$

Учитывая, что $\Phi_{-h/2}^* = \Phi_{h/2}^{-1}$, снова получаем первый составной метод. Следовательно, он является симметричным. То же самое справедливо и для второго составного метода.

Например, если мы применим на полшага сначала явный метод Эйлера, а затем неявный, то получим метод трапеций; если в обратном порядке — метод средней точки. Следует заметить, что результирующие методы имеют уже второй порядок, тогда как их составляющие — первый. Это согласуется с уже полученным выше результатом, что симметричные методы должны иметь четный порядок. Если составляющие методы имеют первый порядок, то результирующие, конечно же, не должны иметь порядок ниже, в то время как следующий (четный) порядок — второй.

Методы Штермера–Верле. Скомпануем две схемы интегрирования как

$$\Phi_{h/2} \circ \Phi_{h/2}^* \quad \text{и} \quad \Phi_{h/2}^* \circ \Phi_{h/2},$$

где в качестве Φ выберем первую симплектическую схему Эйлера (4.130). Тогда составные схемы будут

$$\begin{aligned} \mathbf{x}_{n+1/2} &= \mathbf{x}_n + \frac{h}{2} \frac{\partial g}{\partial \mathbf{y}}(\mathbf{x}_{n+1/2}, \mathbf{y}_n), \\ \mathbf{y}_{n+1} &= \mathbf{y}_n - \frac{h}{2} \left[\frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_{n+1/2}, \mathbf{y}_n) + \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_{n+1/2}, \mathbf{y}_{n+1}) \right], \\ \mathbf{x}_{n+1} &= \mathbf{x}_{n+1/2} + \frac{h}{2} \frac{\partial g}{\partial \mathbf{y}}(\mathbf{x}_{n+1/2}, \mathbf{y}_{n+1}); \end{aligned} \quad (4.135)$$

$$\begin{aligned} \mathbf{y}_{n+1/2} &= \mathbf{y}_n - \frac{h}{2} \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_n, \mathbf{y}_{n+1/2}), \\ \mathbf{x}_{n+1} &= \mathbf{x}_n + \frac{h}{2} \left[\frac{\partial g}{\partial \mathbf{y}}(\mathbf{x}_n, \mathbf{y}_{n+1/2}) + \frac{\partial g}{\partial \mathbf{y}}(\mathbf{x}_{n+1}, \mathbf{y}_{n+1/2}) \right], \\ \mathbf{y}_{n+1} &= \mathbf{y}_{n+1/2} - \frac{h}{2} \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}_{n+1}, \mathbf{y}_{n+1/2}). \end{aligned} \quad (4.136)$$

Методы, определяемые схемами (4.135) и (4.136), называются методами Штермера–Верле (Hairer et al., 2002).

Для важного в классической механике случая $g = \mathbf{y}^2/2 + U(\mathbf{x})$ схемы переписутся как

$$\begin{aligned} \mathbf{x}_{n+1/2} &= \mathbf{x}_n + \frac{h}{2}\mathbf{y}_n, & \mathbf{y}_{n+1} &= \mathbf{y}_n - h\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1/2}), \\ \mathbf{x}_{n+1} &= \mathbf{x}_{n+1/2} + \frac{h}{2}\mathbf{y}_{n+1}; \end{aligned} \quad (4.137)$$

$$\begin{aligned} \mathbf{y}_{n+1/2} &= \mathbf{y}_n - \frac{h}{2}\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n), & \mathbf{x}_{n+1} &= \mathbf{x}_n + h\mathbf{y}_{n+1/2}, \\ \mathbf{y}_{n+1} &= \mathbf{y}_{n+1/2} - \frac{h}{2}\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1}). \end{aligned} \quad (4.138)$$

Используя третью формулу с заменой $n + 1$ на n в каждой схеме для подстановки \mathbf{x}_n и \mathbf{y}_n в первую формулу, получим

$$\begin{aligned} \mathbf{x}_{n+1/2} &= \mathbf{x}_{n-1/2} + h\mathbf{y}_n, & \mathbf{y}_{n+1} &= \mathbf{y}_n - h\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1/2}); \\ \mathbf{y}_{n+1/2} &= \mathbf{y}_{n-1/2} - h\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n), & \mathbf{x}_{n+1} &= \mathbf{x}_n + h\mathbf{y}_{n+1/2}. \end{aligned}$$

Схемы, записанные в такой форме, называются схемами прыгающих (друг через друга) лягушек по аналогии с чехардой, в которую «играют» канонические векторы \mathbf{x} и \mathbf{y} («лягушки»), вычисляемые поочередно с течением времени через каждые полшага.

Если вторую схему (4.138) записать для двух шагов, то удастся исключить переменные \mathbf{y} , и в результате можно получить многошаговую схему

$$\mathbf{x}_{n+1} - 2\mathbf{x}_n + \mathbf{x}_{n-1} = -h^2\frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_n),$$

которая применялась еще И. Ньютоном, Й.Ф. Энке и Ж.Л. Даламбером. Кстати, метод, основанный на этой схеме, иногда называют методом Энке (Hairer et al., 1993).

Методы Штермера–Верле (4.137) и (4.138) как составные реализуются на каждом шаге интегрирования в три этапа. Например,

в первой схеме (4.137) применительно к уравнениям с гамильтонианом $g = \mathbf{y}^2/2 + U(\mathbf{x})$ сначала находится точное решение для положения и скорости невозмущенной задачи ($U = 0$) на средний момент времени; затем корректируется скорость за возмущение; и, наконец, находится точное решение для положения и скорости невозмущенной задачи на конечный момент времени. Тогда для возмущенных уравнений с гамильтонианом $g = \mathbf{y}^2/2 + U_K(\mathbf{x}) + \mu U(\mathbf{x})$ и невозмущенными решениями (4.132) симплектическая схема, аналогичная (4.137), может быть представима в виде

$$\begin{aligned} \mathbf{x}_{n+1/2} &= \Phi_{h/2}(\mathbf{x}_n, \mathbf{y}_n), & \mathbf{y}_{n+1/2} &= \Psi_{h/2}(\mathbf{x}_n, \mathbf{y}_n) - \mu h \frac{\partial U}{\partial \mathbf{x}}(\mathbf{x}_{n+1/2}), \\ \mathbf{x}_{n+1} &= \Phi_{h/2}(\mathbf{x}_{n+1/2}, \mathbf{y}_{n+1/2}), & \mathbf{y}_{n+1} &= \Psi_{h/2}(\mathbf{x}_{n+1/2}, \mathbf{y}_{n+1/2}). \end{aligned} \quad (4.139)$$

Методы, основанные на схеме (4.139), впервые были предложены Дж. Уиздомом и М. Холманом (Wisdom, Holman, 1992). Примечательно, что симплектические методы Уиздома–Холмана имеют локальную ошибку порядка μh^3 . Следовательно, они могут быть весьма эффективны при интегрировании слабозвозмущенных орбит, когда параметр μ существенно мал.

Отметим важные свойства методов Штермера–Верле: 1) методы имеют второй порядок; 2) они симплектические и симметричные; 3) для разделенного гамильтониана $g(\mathbf{x}, \mathbf{y}) = g_1(\mathbf{x}) + g_2(\mathbf{y})$ методы явные. Таким образом, методы Штермера–Верле геометрические, однако для долгосрочного высокоточного интегрирования в задачах динамической астрономии они в оригинальном виде оказываются не востребованы, поскольку имеют низкий порядок. Тем не менее эти методы могут быть весьма полезными для получения составных геометрических методов высоких порядков.

Симплектические и симметричные методы высоких порядков. Составные методы. Интересный способ построения симплектических и симметричных методов высоких порядков — это последовательное использование на шаге простых схем интегрирования. Подобным образом мы уже получили методы второго по-

рядка Штермера–Верле, а также симметричные методы средней точки и трапеций, используя методы первого порядка.

Возьмем схему интегрирования Φ_h порядка p и сконструируем новую Ψ_h на основе последовательного интегрирования

$$\Psi_h = \Phi_{d_s h} \circ \dots \circ \Phi_{d_1 h}, \quad (4.140)$$

где

$$d_1 + \dots + d_s = 1. \quad (4.141)$$

Составная схема (4.140) предполагает пошаговое интегрирование с переменными шагами $d_n h$, поэтому для оценки ее ошибки можно использовать полученную выше формулу (4.25), где, однако, следует положить $N = s$ и $\mathbf{e}_n = \mathbf{C}(t_0)(d_n h)^{p+1} + O(h^{p+2})$:

$$\begin{aligned} \mathbf{x}(t_0 + h) - \mathbf{x}_1 &= \mathbf{x}(t_0 + h) - \Phi_{d_s h} \circ \dots \circ \Phi_{d_1 h}(\mathbf{x}_0) = \\ &= \mathbf{C}(t_0)(d_1^{p+1} + \dots + d_s^{p+1})h^{p+1} + O(h^{p+2}). \end{aligned}$$

Если же теперь потребовать, чтобы

$$d_1^{p+1} + \dots + d_s^{p+1} = 0, \quad (4.142)$$

то составная схема будет иметь, по крайней мере, порядок $p + 1$. Очевидно, что условие (4.142) выполнимо для четных p . Однако тогда, если составная схема оказывается симметричной, она будет иметь порядок уже $p + 2$. Эту особенность подметил Х. Йошида и он предложил оригинальный подход для построения составных симметричных методов высоких (четных) порядков, которые к тому же являются симплектическими (Yoshida, 1990).

Методы Йошиды строятся на основе составной схемы вида

$$\Psi_h = \Phi_{d_1 h} \circ \Phi_{d_0 h} \circ \Phi_{d_1 h}. \quad (4.143)$$

В этом случае условия (4.141) и (4.142) переписутся как

$$d_0 + 2d_1 = 1, \quad d_0^{p+1} + 2d_1^{p+1} = 0. \quad (4.144)$$

Тогда, если метод Φ_h — симметричный четного порядка $p = 2q$, то составной метод Ψ_h будет также симметричным и иметь порядок $p = 2q + 2$.

Теперь допустим, мы имеем симметричный метод второго порядка $\Phi_h^{(2)}$, скажем, Штермера–Верле. Тогда, согласно (4.143) и (4.144), получаем метод четвертого порядка

$$\Phi_h^{(4)} = \Phi_{d_1 h}^{(2)} \circ \Phi_{d_0 h}^{(2)} \circ \Phi_{d_1 h}^{(2)},$$

где

$$d_0 = -\frac{\sqrt[3]{2}}{2 - \sqrt[3]{2}}, \quad d_1 = \frac{1}{2 - \sqrt[3]{2}}.$$

Применяя (4.143) к $\Phi_h^{(4)}$, получаем метод шестого порядка $\Phi_h^{(6)}$ с соответствующими коэффициентами d_0 и d_1 , и так далее. Очевидно, если исходный метод $\Phi_h^{(2)}$ симплектический (например, тот же Штермера–Верле), то и составной метод на его основе также будет симплектическим.

Заметим, что очередное повышение порядка по схеме (4.143) утраивает применение исходного метода $\Phi_h^{(2)}$, так что в составном методе порядка $p = 2q$ число использования исходного метода будет равно $s = 3^{q-1}$. Например, $s = 27$ уже при порядке $p = 8$. В связи с этим для методов высоких порядков возникает проблема уменьшения числа s при сохранении порядка. Разрешая эту проблему, У. Кахану и Р. Ли (Kahan, Li, 1997) удалось получить составной симметричный метод восьмого порядка (отличный от метода Йошиды) с числом $s = 17$.

Отметим очевидные свойства составных методов вида

$$\Psi_h = \Phi_{d_{(s-1)/2} h} \circ \dots \circ \Phi_{d_1 h} \circ \Phi_{d_0 h} \circ \Phi_{d_1 h} \circ \dots \circ \Phi_{d_{(s-1)/2} h},$$

такие, как: 1) методу Ψ_h передаются свойства симметричности либо симплектичности метода Φ_h ; 2) если метод Φ_h сохраняет какой-либо инвариант системы уравнений, то и составной метод Ψ_h также будет сохранять этот инвариант.

Методы Гаусса. В 1988 г. И.М. Санц-Серна (Sanz-Serna, 1988) и Ф.М. Ласани (Lasagni, 1988) обнаружили, что если коэффициенты неявного метода Рунге–Кутты удовлетворяют условиям

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0 \quad (i, j = 1, \dots, s), \quad (4.145)$$

то метод является симплектическим. Нетрудно убедиться, что, согласно (4.145), среди одноэтапных методов Рунге–Кутты симплектическим будет только метод средней точки с коэффициентами $a_{11} = 1/2$ и $b_1 = 1$. В общем случае условиям (4.145) оказываются удовлетворяют методы Гаусса на разбиениях Лежандра (см. разд. 4.3.6), которые, кроме того, являются симметричными (Hairer et al., 2002).

Многошаговые методы. Эффективными геометрическими методами высоких порядков среди многошаговых являются симметричные методы Куинлэна–Тремейна (Quinlan, Tremaine, 1990), которые применяются для интегрирования уравнений второго порядка $\mathbf{x}'' = \mathbf{f}(\mathbf{x})$ и имеют вид

$$\sum_{i=0}^s \alpha_i \mathbf{x}_i = h^2 \sum_{i=0}^s \beta_i \mathbf{f}(\mathbf{x}_i),$$

где $\alpha_i = \alpha_{s-i}$ и $\beta_i = \beta_{s-i}$ ($i = 0, \dots, s$). Авторы получили методы вплоть до 14-го порядка и продемонстрировали их высокую эффективность на примере планетной задачи.

Особенности в использовании симплектических методов. Очевидно, вдоль точного решения канонической системы ее гамильтониан сохраняется, тогда как вдоль приближенного решения нет. Тем не менее в силу симплектичности схем интегрирования они будут сохранять другой, близкий к оригинальному гамильтониан. Так, нетрудно убедиться, что методы Эйлера (4.126) не сохраняют интеграл гармонического осциллятора (4.121), однако их приближенные решения на каждом шаге удовлетворяют другим интегральным соотношениям

$$\bar{g}(x, y) = g(x, y) - \frac{h}{2}xy = \text{const} \quad \text{и} \quad \bar{g}(x, y) = g(x, y) + \frac{h}{2}xy = \text{const} \quad (4.146)$$

соответственно для первой и второй схем (4.126). Функции \bar{g} еще называют модифицированными гамильтонианами (Hairer et al., 2002). Соотношения (4.146) описывают эллипсы, большие полуоси которых направлены вдоль прямых $y = x$ и $y = -x$. Таким образом, приближенные решения x_{n+1} и y_{n+1} (4.126), согласно (4.146), будут лежать на замкнутых и компактных (при малых h) кривых, поэтому ошибка в g будет ограничена:

$$\Delta g = g(x_{n+1}, y_{n+1}) - g(x_0, y_0) = \pm \frac{h}{2}(x_{n+1}y_{n+1} - x_0y_0),$$

хотя при использовании обычных (негеометрических) методов ее величина имеет тенденцию к неограниченному росту.

В случае применения методов Эйлера (4.130) для интегрирования канонической системы с произвольным гамильтонианом g известно, что модифицированный гамильтониан будет иметь вид (Yoshida, 1993)

$$\bar{g} = g + \sum_{i=1}^{\infty} (\mp h)^i g_i = g + O(h), \quad g_1 = \frac{1}{2} \frac{\partial g}{\partial \mathbf{x}} \cdot \frac{\partial g}{\partial \mathbf{y}}, \quad (4.147)$$

$$g_2 = \frac{1}{12} \left[\left(\frac{\partial g}{\partial \mathbf{y}} \right)^T \frac{\partial^2 g}{\partial \mathbf{x}^2} \left(\frac{\partial g}{\partial \mathbf{y}} \right) + \left(\frac{\partial g}{\partial \mathbf{x}} \right)^T \frac{\partial^2 g}{\partial \mathbf{y}^2} \left(\frac{\partial g}{\partial \mathbf{x}} \right) \right], \dots$$

В степенном ряду (4.147) знак минус при h соответствует первой схеме Эйлера (4.130), тогда как плюс — второй. Если используются симплектические методы p -го порядка, то модифицированный гамильтониан формально можно представить как

$$\bar{g} = g + O(h^p).$$

Таким образом, симплектический метод будет сохранять модифицированный гамильтониан \bar{g} , близкий к оригинальному g , причем близость гамильтонианов определяется не только величиной шага интегрирования h , но и порядком метода p . Это, в свою очередь, обеспечивает ограниченность ошибки Δg .

Следует заметить, что интегральное соотношение $\bar{g} = \text{const}$ явно зависит от h , поэтому оно будет сохраняться только при постоянном шаге. Если шаг переменный, то модифицированный гамильтониан уже не будет интегралом и в этом случае ограниченность ошибки Δg не гарантируется. Так, при симплектическом интегрировании орбиты задачи двух тел с переменным шагом гамильтониан задачи (кеплеровская энергия) не сохраняется и ошибка в интегральном соотношении неограниченно возрастает. В связи с этим для эффективного использования симплектических методов необходимо выполнять интегрирование с постоянным шагом.

Впечатляющими оказываются результаты при использовании симплектической схемы средней точки, которая также является симметричной. Применительно к уравнениям гармонического осциллятора она точно сохраняет гамильтониан системы. Следует заметить, что в случае линейных систем дифференциальных уравнений схема средней точки совпадает с симметричной схемой трапеций, которая, вообще говоря, не симплектична. Вместе с тем во многих последних работах по геометрическим интеграторам (Hairer et al., 2002) показывается, что симметричные методы так же, как и симплектические, позволяют ограничить ошибки в гамильтониане при интегрировании канонических систем.

4.7. Сравнительный анализ эффективности методов

Об эффективности методов интегрирования можно говорить лишь в отношении к конкретным дифференциальным уравнениям. Естественно, нас интересуют уравнения орбитального движения, и в данном разделе мы покажем, насколько эффективно использование методов интегрирования применительно к уравнениям в прямоугольных координатах, которые в общем случае, как известно, являются далеко не самой лучшей формализацией орбит.

4.7.1. Численный эксперимент

Методы численного интегрирования тестировались применительно к дифференциальным уравнениям плоского движения вида (3.70):

$$\frac{d^2\mathbf{x}}{dt^2} = -\frac{\mathbf{x}}{|\mathbf{x}|^3} - \mu_P \frac{\mathbf{x} - \mathbf{x}_P}{|\mathbf{x} - \mathbf{x}_P|^3} - \mu_P \frac{\mathbf{x}_P}{|\mathbf{x}_P|^3},$$

где, напомним, μ_P и \mathbf{x}_P — соответственно гравитационный параметр и вектор положения возмущающего тела, вычисляемый по формулам кругового движения с орбитальным радиусом a_P .

Рассматривались пять безразмерных задач: две невозмущенные — круговая и эллиптическая с эксцентриситетом $e = 0.7$, и три возмущенные, где возмущения вызывались в одной задаче притяжением от внутреннего быстрого тела (\mathbf{P}_1 : $a_P = 0.14$, $\mu_P = 1.7 \cdot 10^{-7}$), в двух других — от внешнего (\mathbf{P}_2 : $a_P = 2.08$, $\mu_P = 10^{-3}$), причем в первых двух интегрировались почти круговые орбиты, тогда как в последней (сильновозмущенной) — эллиптическая ($e \approx 0.7$), где регулярно имели место сближения с возмущающим телом порядка 0.4. В слабовозмущенных задачах параметры были подобраны так, чтобы они приблизительно соответствовали орбитальной динамике астероида главного пояса в рамках тройных систем астероид–Солнце–Меркурий и астероид–Солнце–Юпитер. Первая из возмущенных задач короткопериодическая, где частоты исследуемого и возмущаемого тел относятся как $n/n_P \approx 1/20$, тогда как другие — резонансные с соотношением тех же частот $n/n_P = 3/1$, что, кстати, и приводит к регулярным сближениям. Конфигурация орбит, а также начальные положения исследуемого (точки) и возмущающих тел (кружочки) показаны на рис. 4.3.

Точность интегрирования $|\Delta\mathbf{x}|$ оценивалась на конце временного интервала, за который объект выполняет 1000 об., путем сравнения полученных результатов с начальными данными в невозмущенной задаче³⁴ и с результатами высокоточного интегри-

³⁴Здесь учитывается свойство периодичности кеплеровского движения.

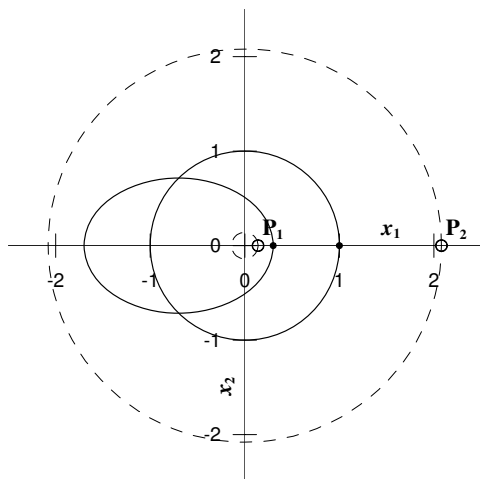


Рис. 4.3. Конфигурация орбит

рования, полученными при использовании интегратора Гаусса–Эверхарта 15-го порядка, в возмущенных задачах. В качестве показателя быстродействия взято количество обращений к процедуре правых частей уравнений NF . Все вычисления выполнялись в арифметике с двойной точностью.

4.7.2. Выбранные методы

Для тестирования выберем наиболее ярких представителей из семейств методов 8-го порядка (табл. 4.2): вложенный метод Дормана–Принса 8(7); метод Гаусса–Эверхарта (неявный коллокационный метод Рунге–Кутты типа Бутчера на разбиении Гаусса–Лобатто); экстраполяционный метод Грэгга; многошаговый неявный метод Адамса–Мульттона с предикторной явной схемой Адамса–Башфорта 7-го порядка; составной геометрический метод Кахана–Ли и многошаговый явный симметричный метод Куинлэна–Тремейна для уравнений второго порядка. В табл. 4.2 даются ссылки

Таблица 4.2. Методы интегрирования 8-го порядка

Метод	Sn	Стр.	Источник
Дормана–Принса	DP	142	(Hairer et al., 1993)
Гаусса–Эверхарта	GE	152	(Авдюшев, 2010b)
Грэгга	GBS	169	(Hairer et al., 1993)
Адамса–Мультона	AM	180	(Холл, Уатт, 1979)
Кахана–Ли	KL	200	(Hairer et al., 2002)
Куинлэна–Тремейна	QT	201	(Quinlan, Tremaine, 1990)

на страницы в главе и литературу, где описаны теории и алгоритмы методов.

Отметим некоторые особенности в компьютерной реализации методов. В неявных методах Гаусса–Эверхарта и Адамса–Мультона для численного решения нелинейных уравнений задавалось только две итерации, что было вполне достаточно для сходимости итерационного процесса. Многошаговое интегрирование методом Куинлэна–Тремейна выполнялось с постоянным шагом: ввиду специфики схемы интегрирования использование переменного шага просто невозможно. Между тем стартовые решения в методе Куинлэна–Тремейна вычислялись путем численного интегрирования классическим методом Рунге–Кутты (4.19) с уменьшенным на порядок шагом, тогда как в методе Адамса–Мультона — в соответствии с самостартующей схемой (4.113) и (4.114). В методе Кахана–Ли в качестве опорного Φ был взят первый метод Штермера–Верле (4.135), а шаг выбирался пропорционально радиус-вектору.

4.7.3. Численные результаты

На рис. 4.4–4.8 представлены характеристики эффективности методов, точность–быстродействие. Каждая из них получена путем многократного численного интегрирования уравнений соответствующим методом при различных значениях задаваемой ло-

кальной точности (и при различных величинах постоянного шага для метода Куинлэна–Тремейна).

При интегрировании круговой (рис. 4.4) и почти круговых (рис. 4.6 и 4.7) орбит впечатляющую эффективность с точки зрения соотношения точности и быстродействия демонстрирует многошаговый метод Куинлэна–Тремейна (**QT**). В данном случае его высокая эффективность достигается главным образом за счет того, что метод является геометрическим и, кроме того, требует всего одного вычисления функции правых частей уравнений на шаге, поскольку метод явный. Неявный же метод Адамса–Мультонна (**AM**) требует двух вычислений функции правых частей, и он не геометрический, поэтому его эффективность заметно ниже. Несмотря на то, что метод Кахана–Ли (**KL**) геометрический (симметричный и симплектический), он дает результаты, близкие к результатам метода Адамса–Мультонна. Объясняется это прежде всего тем, что составной метод является 17-этапным, и это, конечно, существенно сказывается на его быстродействии. Заметно ниже эффективность у методов Гаусса–Эверхарта (**GE**), Грэгга–Булирша–Штера (**GBS**) и Дормана–Принса (**DP**). Причем последний демонстрирует наименьшую эффективность во всех задачах. Тем не менее нельзя не заметить, что предельно высокая точность для этих методов значительно выше, нежели для других, на 1–3 порядка. По-видимому, специфика многошаговых и составных методов такова, что ошибки округления для них накапливаются существенно интенсивнее.

В эллиптическом случае (рис. 4.5 и 4.8) картина резко меняется. Прежде всего из-за регулярных сближений с центральным телом эффективность методов существенно падает. Причем абсолютно теряются явные преимущества многошаговых методов: их эффективность становится даже ниже, чем у метода Гаусса–Эверхарта (**GE**).

Хотелось бы обратить внимание на пологое поведение характеристик для метода Грэгга–Булирша–Штера (**GBS**) при больших NF (т.е. при малых h). Оно вызвано тем, что компьютерная реализация метода (Hairer et al., 1993) при достаточно малой вели-

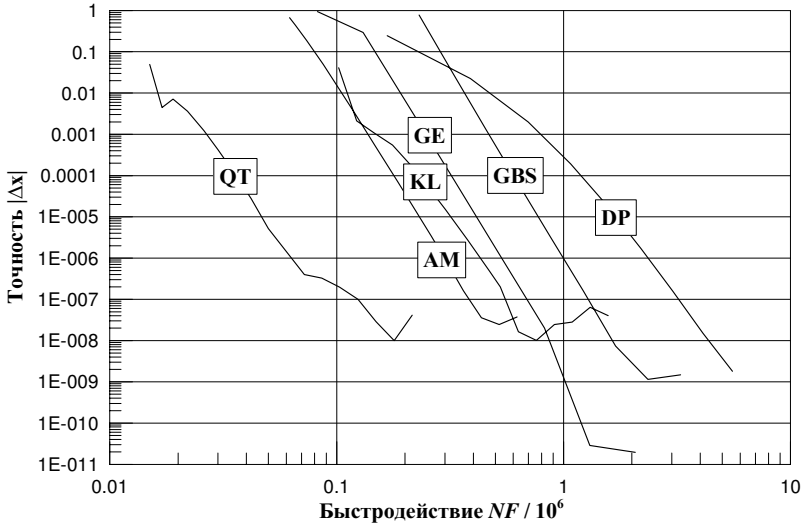


Рис. 4.4. Круговая задача

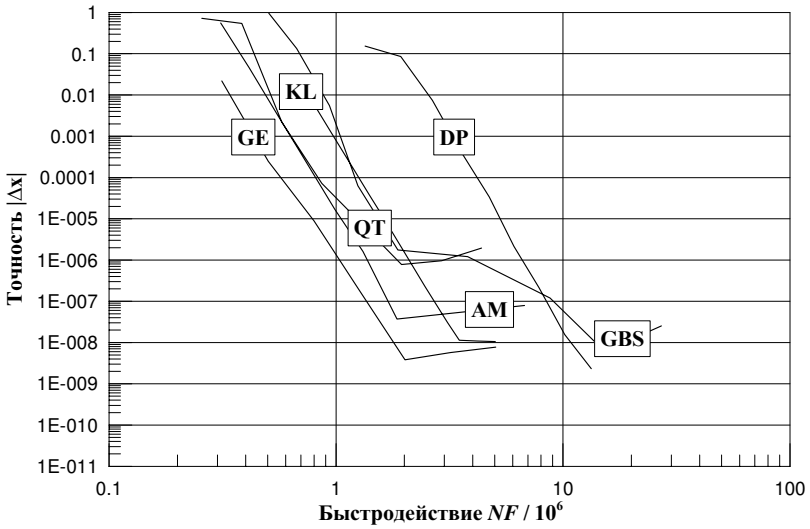


Рис. 4.5. Эллиптическая задача ($e = 0.7$)

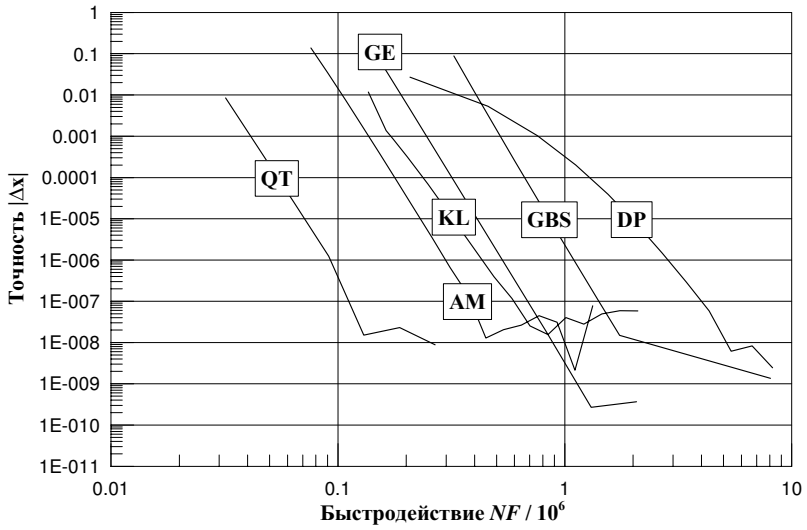


Рис. 4.6. Резонансная возмущенная задача ($n/n_P = 3/1$)

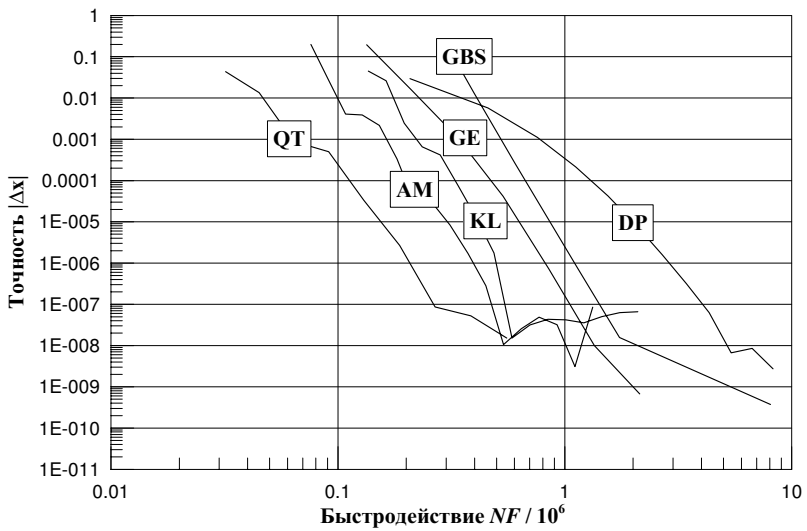


Рис. 4.7. Короткопериодическая возмущенная задача ($n/n_P \approx 1/20$)

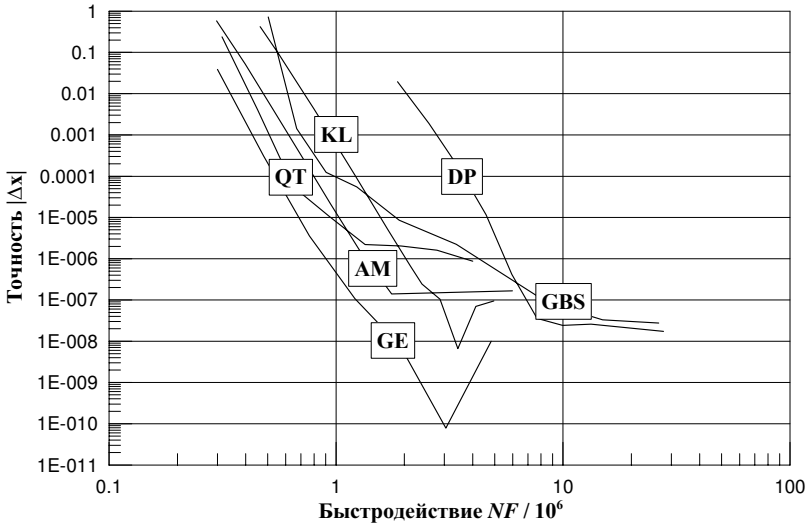


Рис. 4.8. Резонансная сильновозмущенная задача ($n/n_P = 3/1$)

чине шага интегрирования дает решение более низкого (четного) порядка от второго до шестого, т.е. таблица приближенных решений вычисляется не до восьмого порядка. Особенность такой реализации, по мнению автора, должна оптимизировать численный процесс, но, как мы видим, в данном случае это не так.

Слабые резонансные возмущения, когда интегрируется почти круговая орбита, не существенно влияют на эффективность методов (рис. 4.6). С другой стороны, при короткопериодических возмущениях характеристики становятся заметно теснее (рис. 4.7). Впрочем, в данном случае показатель короткопериодичности (3.73) для восьмого порядка не столь большой, чтобы нивелировать эффективность методов, и составляет величину $\nu \approx 2.7$.

Таким образом, для высокоточного моделирования (сложного) орбитального движения, безусловно, следует рекомендовать неявные методы Рунге–Кутты типа Бутчера, в частности интегратор Гаусса–Эверхарта (**GE**), который мы, кстати, использова-

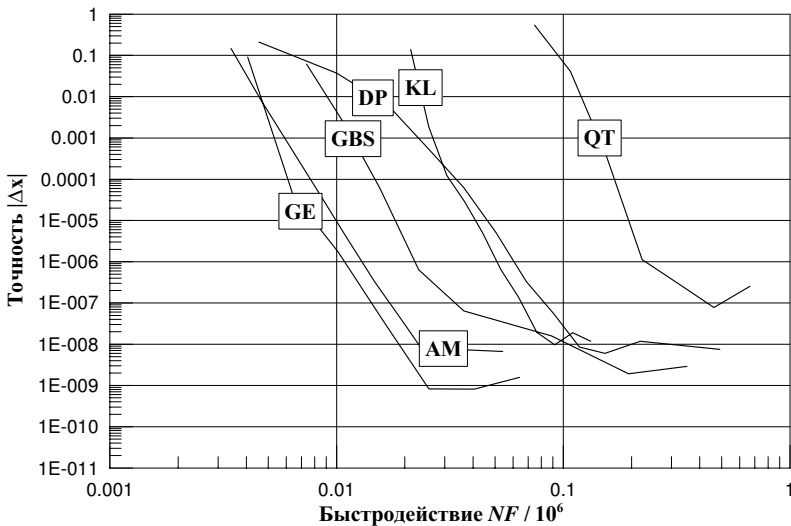


Рис. 4.9. Ограниченная задача трех тел

ли для получения всех численных результатов, представленных в других главах. При невысоких требованиях к точности, в особенности, когда моделируется почти круговая орбита, естественно следует обращаться либо к симметричному многошаговому методу Куинлэна–Тремейна (**QT**), если интегрируются динамические уравнения только второго порядка, либо к методу Адамса–Мультонна (**AM**), когда уравнения не допускают приведения ко второму порядку.

Наконец, эффективность методов была также исследована применительно к моделированию движения ограниченной задачи трех тел, сформулированной на с. 52 (см. рис. 3.5). Характеристики точность–быстродействие, представленные на рис. 4.9, выявляют двух безусловных фаворитов: это — неявные методы Рунге–Кутты (**GE**) и Адамса–Мультонна (**AM**). Другие методы оказываются более медленными: метод Грэгга–Булирша–Штера (**GBS**) — в два раза; методы Дормана–Принса (**DP**) и Кахана–Ли (**KL**) —

почти в четыре раза. Наименее эффективен многошаговый метод Куинлэна–Тремейна (**QT**), что объясняется невозможностью выполнять интегрирование с переменным шагом. Следует отметить, что в методе Кахана–Ли (**KL**) шаг интегрирования выбирался пропорционально величине $f = (\dot{\mathbf{x}}^2 + \ddot{\mathbf{x}}^2)^{-3/8}$, т.е. так чтобы соответствующее фиктивное время (см. разд. 3.2) менялось приблизительно равномерно.

Выдавить пасту из тюбика — это прямая задача,
а загнать ее затем снова в тюбик — обратная.
А.М. Черепашук

Глава 5. ОПРЕДЕЛЕНИЕ ОРБИТ ИЗ НАБЛЮДЕНИЙ

Одним из основных и важных этапов моделирования является определение орбитальных параметров (или просто определение орбиты), т.е. решение обратной задачи. Действительно, ведь если значения параметров неизвестны, математическая модель с практической точки зрения становится не более чем бесполезной грудой формул. Орбитальные параметры определяются из наблюдений небесного тела, как правило, в рамках задачи наименьших квадратов, что позволяет из семейства орбит, описываемых математической моделью, выделить единственную, которая наилучшим образом соответствует реальному движению исследуемого объекта, по крайней мере, представляемому его наблюдениями.

В данной главе формулируется обратная задача орбитальной динамики, а также излагаются широкоиспользуемые на практике численные методы ее решения и приводятся результаты сравнительного анализа их эффективности применительно к определению орбит близких спутников. В силу нелинейности обратной задачи формально ее решение, вообще говоря, не является единственным. В конце главы эта проблема исследуется на примере определения спутниковых орбит.

5.1. Обратная задача орбитальной динамики

Обратная задача орбитальной динамики состоит в определении орбитальных параметров из астрометрических наблюдений в рамках выбранной модели. Пусть имеем J наблюдаемых положений \mathbf{r}_i^O в L -мерном пространстве на моменты времени t_i ($i = 1, \dots, J$).

Требуется по наблюдениям \mathbf{p}_i^O определить K орбитальных параметров \mathbf{q} модели $\mathbf{p}^C = \mathbf{p}^C(t, \mathbf{q})$.

Обычно определение \mathbf{q} сводится к минимизации суммы квадратов невязок, т.е. целевой функции

$$S(\mathbf{q}) = \sum_{i=1}^J \|\mathbf{p}_i^O - \mathbf{p}_i^C\|^2, \quad (5.1)$$

где $\mathbf{p}_i^C = \mathbf{p}^C(t_i, \mathbf{q})$. Поскольку норма в (5.1) евклидова, то целевую функцию можно переписать в виде

$$S(\mathbf{q}) = \|\mathbf{p}^O - \mathbf{p}^C\|^2, \quad (5.2)$$

где $\mathbf{p}^O = (\mathbf{p}_1^O, \dots, \mathbf{p}_J^O)^T$ и $\mathbf{p}^C = (\mathbf{p}_1^C, \dots, \mathbf{p}_J^C)^T$ — N -мерные векторы измерений и их модельных представлений; $N = JL$.

Если наблюдения разнородны или неравноточны, то целесообразно ввести весовую симметричную (как правило, диагональную) матрицу \mathbf{W} размера $N \times N$. Тогда целевую функцию (5.2) формально можно записать как

$$S(\mathbf{q}) = \|(\mathbf{W}^{1/2})^T(\mathbf{p}^O - \mathbf{p}^C)\|^2, \quad (5.3)$$

где $\mathbf{W}^{1/2}$ — такая матрица размера $N \times N$ (скажем, матрица-множитель разложения Холецкого), для которой $\mathbf{W}^{1/2}(\mathbf{W}^{1/2})^T = \mathbf{W}$. Фактически весовая матрица подбирается таким образом, чтобы линейно преобразованные наблюдения $(\mathbf{W}^{1/2})^T \mathbf{p}^O$ стали равноточными (см. также разд. 6.4). Например, при обработке астрометрических наблюдений ($L = 2$) в сферических координатах α^O (прямое восхождение) и δ^O (склонение), полученных с равноточных телескопов, обычно используют весовую матрицу $\mathbf{W} = \text{diag}(w_1, \dots, w_{2J})$, где $w_{2i-1} = \cos^2 \delta_i^O$ и $w_{2i} = 1$ ($i = 1, \dots, J$): известно, что с увеличением склонения δ^O дисперсия ошибок измерения для α^O обратно пропорциональна $\cos^2 \delta^O$, для чего и вводятся соответствующие компенсирующие веса.

Минимум функции (5.3) находится из необходимого условия экстремума:

$$\frac{\partial S}{\partial \mathbf{q}} = -2(\mathbf{p}^O - \mathbf{p}^C)^T \mathbf{W} \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} = \mathbf{0}. \quad (5.4)$$

Система уравнений (5.4) нелинейна, и поэтому на практике для ее решения прибегают к итерационным методам, которые будут изложены ниже. Решение системы (5.4), минимизирующее также целевую функцию, будем обозначать $\hat{\mathbf{q}}$.

Справедливости ради хотелось бы упомянуть еще один принципиально иной, хотя и не так часто используемый способ определения орбит, а именно основанный на принципе наименьших модулей. Примечательной особенностью этого способа является то, что если ошибки наблюдений распределены заметно не по нормальному закону, то оценки $\hat{\mathbf{q}}$ в рамках задачи наименьших модулей могут быть эффективнее оценок задачи наименьших квадратов (Мудров, Кушко, 1971), т.е. можно ожидать более высокую точность определяемых орбитальных параметров.

Целевая функция задачи наименьших модулей имеет вид

$$S(\mathbf{q}) = \sum_{i=1}^N |p_i^O - p_i^C|.$$

В то же время нетрудно видеть, что формально ее можно привести к целевой функции взвешенной задачи наименьших квадратов (5.3) с диагональной весовой матрицей \mathbf{W} , где в качестве весов будут выступать обратные величины невязок $w_i = 1/|p_i^O - p_i^C|$.

Таким образом, для численного решения задачи наименьших модулей могут применяться те же итерационные методы, что и для задачи наименьших квадратов. Однако ввиду зависимости весов от параметров \mathbf{q} задача наименьших модулей решается поэтапно (Мудров, Кушко, 1971), а именно на каждом следующем этапе используется весовая матрица, определяемая по оценкам параметров (точнее, по невязкам) предыдущего этапа, которые рассматриваются как постоянные, а очередные приближенные оценки находятся в результате итерационного решения взвешенной задачи наименьших квадратов, и так до сходимости оценок по достижении требуемой точности. При этом на первом этапе принимается единичная весовая матрица. Очевидно, реализация такого подхода может быть существенно более трудоемкой, нежели в случае определения орбит в рамках задачи наименьших квадратов.

5.2. Модельные представления наблюдений

Наблюдения небесных тел могут быть представлены в самых разнообразных величинах. Однако наиболее часто наблюдаются угловые координаты объекта (прямое восхождение α и склонение δ) на небесной сфере, отнесенной к наземному наблюдателю (топоцентру). Приведем основные формулы для модельного представления этих наблюдений.

Координатное преобразование \mathbf{T} (1.1) можно представить как последовательность преобразований

$$\mathbf{T} = \mathbf{T}_{\alpha,\delta} \circ \mathbf{T}_{top} \circ \mathbf{T}_{geo}, \quad (5.5)$$

где \mathbf{T}_{geo} и \mathbf{T}_{top} — переходы соответственно от координатной системы \mathbf{x} к геоцентрической, отнесенной на стандартную эпоху, скажем, J2000.0, и от гео- к топоцентрической системе; $\mathbf{T}_{\alpha,\delta}$ — получение угловых координат относительно наблюдателя. Преобразование перехода к топоцентру $\mathbf{T}_{top} \circ \mathbf{T}_{geo}$ формально можно записать в виде

$$\mathbf{T}_{top} \circ \mathbf{T}_{geo}(\mathbf{x}) = \mathbf{x} - \mathbf{x}_{EO} - \mathbf{x}_{TE} = \mathbf{x}_T, \quad (5.6)$$

где \mathbf{x}_{EO} — положение Земли в пространстве \mathbf{x} ; \mathbf{x}_{TE} — геоцентрическое положение наблюдателя, вычисляемое по его сферическим координатам: геоцентрическому расстоянию b , широте φ и местному звездному времени s , как

$$x_{TE1} = b \cos \varphi \cos s, \quad x_{TE2} = b \cos \varphi \sin s, \quad x_{TE3} = b \sin \varphi,$$

$$s = \lambda + 2\pi UT1 + 2\pi(24110^s \cdot 54841 + 8640184^s \cdot 812866T + 0^s \cdot 093104T^2 - 6^s \cdot 2 \cdot 10^{-6}T^3)/86400^s. \quad (5.7)$$

Здесь λ — долгота наблюдателя; T — юлианские столетия от стандартной эпохи J2000.0: $T = (UT1 - 2451545)/36525$, $UT1$ — среднее солнечное время в сутках. Наконец, угловые координаты получаются по формулам

$$\alpha^C = \arctg \frac{x_{T2}}{x_{T1}}, \quad \delta^C = \arcsin \frac{x_{T3}}{|\mathbf{x}_T|}.$$

Наземные наблюдения $\mathbf{p}_i^O = (\alpha_i^O, \delta_i^O)^T$, как правило, отнесены к моментам t_i^O всемирного координированного времени (аппроксимация UT1), поэтому при использовании численной модели (1.1) требуется предварительный переход к эфемеридному времени t_i^* . Кроме того, для получения видимых положений небесного тела $\mathbf{p}_i^C = (\alpha_i^C, \delta_i^C)^T$ относительно земного наблюдателя необходимо также учитывать эффект запаздывания света, который выражается посредством уравнения

$$t_i = t_i^* - |\mathbf{x}_{TO}(t_i^*) - \mathbf{x}(t_i)|/c, \quad (5.8)$$

где $\mathbf{x}_{TO}(t_i^*)$ — положение топоцентра в пространстве \mathbf{x} на момент эфемеридного времени t_i^* ; c — скорость света. Уравнение (5.8) решается методом простых итераций, что предполагает многократное численное интегрирование орбиты (обычно 2–3 итерации) для получения $\mathbf{x}(t_i)$ на каждое новое приближение эфемеридного времени. В качестве начального приближения выбирается $t_i = t_i^*$. Кстати, именно полученный в результате решения уравнения вектор положения \mathbf{x} должен использоваться в (5.6).

5.3. Предварительное определение орбиты

Поскольку нелинейные уравнения (5.4) решаются итерационными методами, необходимо иметь начальное приближение \mathbf{q}_0 . Этап его нахождения при решении обратной задачи в теоретической астрономии называется предварительным определением орбиты.

Начальные параметры (составляющие вектор динамического состояния) определяются, как правило, по трем угловым наблюдениям ($N = 6$) в рамках модели кеплерового движения либо методом Гаусса, либо Лапласа (Escobal, 1965). Это равнозначно интерполированию наблюдательных данных формулами задачи двух тел, что фактически сводится к решению уравнений (5.4) при $N = K = 6$, где, однако, $\mathbf{p}^C = \mathbf{p}^C(t, \mathbf{q})$ — уже кеплерова модель орбиты³⁵.

³⁵Впрочем, некоторые авторы (Нумеров, 1923; Herrick, 1952; Шефер, 2010) предлагают более точные модели, представляющие возмущенную орбиту.

При численном определении сложных орбит метод Лапласа предпочтительнее, поскольку он, в отличие от метода Гаусса, непосредственно основан на дифференциальных уравнениях орбитального движения, и поэтому позволяет выполнять поиск начальных параметров вне семейства кеплеровских орбит. Это, например, особенно важно, когда определяется сильновозмущенная орбита астероида по наблюдениям, выполненным во время тесного сближения объекта с Землей (или с какой-либо другой планетой).

Кроме того, в методе Лапласа фактически снимаются ограничения на количество используемых наблюдений: их может быть и больше трех. Интерполяция величин, связанных с угловыми измерениями, в оригинальном методе Лапласа на три момента времени дает возможность вычислить необходимые их скорость и ускорение на орбитальной дуге. Однако динамические параметры можно вычислять точнее, если использовать дополнительные наблюдения, покрывающие ту же дугу орбиты. В этом случае движение объекта на небесной сфере представляется уже не интерполяцией, а полиномиальной (квадратичной) аппроксимацией, получаемой методом наименьших квадратов. Причем привлекаются те наблюдения, которые обеспечивают хорошее согласие с их представлениями полиномиальной аппроксимацией. Рассмотрим обобщенный метод Лапласа подробнее³⁶.

Вектор положения наблюдаемого небесного тела относительно наземного наблюдателя можно представить как $\mathbf{x}_T = \rho \mathbf{l}$, где $\rho = |\mathbf{x}_T|$ — расстояние от наблюдателя до объекта, а \mathbf{l} — единичный вектор:

$$l_1 = \cos \alpha \cos \delta, \quad l_2 = \sin \alpha \cos \delta, \quad l_3 = \sin \delta.$$

Тогда, согласно (5.6),

$$\mathbf{x} = \rho \mathbf{l} + \mathbf{x}_{EO} + \mathbf{x}_{TE}, \quad (5.9)$$

$$\dot{\mathbf{x}} = \dot{\rho} \mathbf{l} + \rho \dot{\mathbf{l}} + \dot{\mathbf{x}}_{EO} + \dot{\mathbf{x}}_{TE}, \quad (5.10)$$

$$\ddot{\mathbf{x}} = \ddot{\rho} \mathbf{l} + 2\dot{\rho} \dot{\mathbf{l}} + \rho \ddot{\mathbf{l}} + \ddot{\mathbf{x}}_{EO} + \ddot{\mathbf{x}}_{TE} = \mathcal{P}(t, \mathbf{x}, \dot{\mathbf{x}}). \quad (5.11)$$

³⁶Оригинальный метод Лапласа представлен в Приложении.

При наблюдениях на каждый момент времени t_i будем иметь соответствующий вектор \mathbf{l}_i . Построим для единичного вектора \mathbf{l} квадратичную аппроксимирующую функцию

$$\mathbf{g}(t) = \mathbf{g}_0 + \mathbf{g}_1(t - t_0) + \mathbf{g}_2(t - t_0)^2 \quad (5.12)$$

по тем векторам \mathbf{l}_i , для которых бы среднеквадратическая ошибка не превышала заданную предельную величину, соответствующую, например, дисперсии ошибок астрометрических наблюдений. Такой набор единичных векторов обязательно найдется, по крайней мере, потому, что уже любые три взятых вектора обеспечивают нулевую среднеквадратическую ошибку, когда аппроксимация становится интерполяцией как в оригинальном методе Лапласа. Здесь в качестве начальной эпохи t_0 выбирается момент в середине временного интервала аппроксимации. Аппроксимирующая функция определяется из принципа наименьших квадратов, согласно которому коэффициенты полинома \mathbf{g} должны удовлетворять нормальной системе линейных уравнений

$$\begin{aligned} \mathbf{g}_0 \sum_i 1 + \mathbf{g}_1 \sum_i \Delta t_i + \mathbf{g}_2 \sum_i \Delta t_i^2 &= \sum_i \mathbf{l}_i \\ \mathbf{g}_0 \sum_i \Delta t_i + \mathbf{g}_1 \sum_i \Delta t_i^2 + \mathbf{g}_2 \sum_i \Delta t_i^3 &= \sum_i \mathbf{l}_i \Delta t_i \\ \mathbf{g}_0 \sum_i \Delta t_i^2 + \mathbf{g}_1 \sum_i \Delta t_i^3 + \mathbf{g}_2 \sum_i \Delta t_i^4 &= \sum_i \mathbf{l}_i \Delta t_i^2 \end{aligned}$$

где $\Delta t_i = t_i - t_0$, а индекс i принимает значения, соответствующие выбранным для аппроксимации моментам времени. Эта система может быть решена методом Крамера, к которому мы также обратимся ниже применительно к другой системе.

Для момента t_0 представим соотношение (5.11) в виде

$$\ddot{\rho}_0 \mathbf{l}_0 + 2\dot{\rho}_0 \dot{\mathbf{l}}_0 + \rho_0 \ddot{\mathbf{l}}_0 = \mathcal{P}(t_0, \mathbf{x}_0, \dot{\mathbf{x}}_0) - (\ddot{\mathbf{x}}_{EO})_0 - (\ddot{\mathbf{x}}_{TE})_0 \equiv \mathbf{a}_0, \quad (5.13)$$

помечая нулем начальные значения величин. Здесь в качестве единичного вектора и его производных принимаются их приближенные аналоги: $\mathbf{l}_0 = \mathbf{g}_0$, $\dot{\mathbf{l}}_0 = \mathbf{g}_1$, $\ddot{\mathbf{l}}_0 = 2\mathbf{g}_2$. Вторые производные $(\ddot{\mathbf{x}}_{EO})_0$ и $(\ddot{\mathbf{x}}_{TE})_0$, впрочем, как и первые $(\dot{\mathbf{x}}_{EO})_0$ и $(\dot{\mathbf{x}}_{TE})_0$, которые

понадобятся позже, получаются путем дифференцирования (возможно даже численно) аналитических представлений \mathbf{x}_{EO} (например, эфемериды DE405) и \mathbf{x}_{TE} (5.7).

Далее, рассматривая (5.13) как систему линейных уравнений относительно $\ddot{\rho}_0$, $\dot{\rho}_0$ и ρ_0 , найдем ее решение методом Крамера. Для ρ_0 и $\dot{\rho}_0$ будем иметь

$$\rho_0 = \frac{|\mathbf{I}\mathbf{I}\mathbf{a}|_0}{|\mathbf{I}\mathbf{I}\mathbf{I}|_0}, \quad \dot{\rho}_0 = \frac{|\mathbf{I}\mathbf{a}\mathbf{I}|_0}{2|\mathbf{I}\mathbf{I}\mathbf{I}|_0}. \quad (5.14)$$

Здесь запись типа $|\mathbf{abc}|$ означает определитель матрицы, составленной из векторов \mathbf{a} , \mathbf{b} и \mathbf{c} . В действительности же пока получены только нелинейные уравнения для расстояния и ее производной, поскольку $\mathbf{a} = \mathbf{a}(\rho, \dot{\rho})$. Впрочем, в подавляющем большинстве случаев учитываемые силы \mathcal{P} либо не зависят от скоростей, либо, если даже какие-нибудь и зависят, то настолько слабы, что на этапе предварительного определения орбиты вполне могут быть отброшены, поэтому $\mathbf{a} = \mathbf{a}(\rho)$. Тогда остается лишь одно нелинейное уравнение для ρ_0 . Решая его каким-либо численным методом, найдем ρ_0 , а затем — $\dot{\rho}_0$ из второго соотношения (5.14), используя уже вычисленное значение \mathbf{a}_0 . Если необходимо учитывать эффект запаздывания света, то силы \mathcal{P} следует вычислять для опережающего положения объекта на момент $t_0 - \Delta t$, где $\Delta t = \rho_0/c$, т.е. $\mathcal{P}(t_0 - \Delta t, \mathbf{x}_0)$. Заметим, что положение \mathbf{x}_0 здесь, как и скорость $\dot{\mathbf{x}}_0$, определяемые по (5.9) и (5.10), уже автоматически отнесены на эпоху $t_0 - \Delta t$, а не на t_0 . В результате нахождения ρ_0 и $\dot{\rho}_0$ будем иметь также и параметры \mathbf{x}_0 и $\dot{\mathbf{x}}_0$, из которых получаем начальное приближение вектора динамического состояния \mathbf{q}_0 .

5.4. Изохронные производные

Решение уравнений (5.4) предполагает вычисление так называемых изохронных производных $\partial \mathbf{p}^C / \partial \mathbf{q}$. Согласно (1.1), для каждого момента времени t_i их можно представить как

$$\left. \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right|_{t_i} = \left. \frac{\partial \mathbf{T}}{\partial \mathbf{x}} \right|_{t_i} \left. \frac{\partial \mathbf{x}}{\partial \mathbf{q}} \right|_{t_i}. \quad (5.15)$$

Здесь размеры матриц в левой части $L \times K$, а в правой $L \times 3$ и $3 \times K$ соответственно. При обработке астрометрических наблюдений ($L = 2$) производные $\partial \mathbf{T} / \partial \mathbf{x}$ в динамической модели (1.1) определяются аналитически из дифференциальных соотношений

$$\begin{aligned} \rho^C d\alpha^C \cos \delta^C &= -dx_1 \sin \alpha^C + dx_2 \cos \alpha^C, \\ \rho^C d\delta^C &= -dx_1 \cos \alpha^C \sin \delta^C - dx_2 \sin \alpha^C \sin \delta^C + dx_3 \cos \delta^C, \end{aligned}$$

тогда как производные $\partial \mathbf{x} / \partial \mathbf{q}$ в соответствии с (1.3) находятся численно из дифференциальных уравнений в вариациях

$$\frac{d^2}{dt^2} \frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \frac{\partial \mathcal{P}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{q}} + \frac{\partial \mathcal{P}}{\partial \dot{\mathbf{x}}} \frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{q}} \quad (5.16)$$

с начальными условиями

$$\left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right)_0 = \mathbf{E}, \quad \left(\frac{\partial \mathbf{x}}{\partial \dot{\mathbf{x}}_0} \right)_0 = \mathbf{0}, \quad \frac{d}{dt} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right)_0 = \mathbf{0}, \quad \frac{d}{dt} \left(\frac{\partial \mathbf{x}}{\partial \dot{\mathbf{x}}_0} \right)_0 = \mathbf{E}.$$

Здесь $\mathbf{T} = (\alpha^C, \delta^C)^T$, $\rho^C = |\mathbf{x}_T|$ — топоцентрическое расстояние космического объекта; \mathbf{E} — единичная матрица размера 3×3 . Уравнения (5.16) интегрируются численно совместно с уравнениями движения (1.3). Таким образом, при моделировании интегрируется система 42-го порядка.

Для решения системы уравнений (5.16) требуются производные $\partial \mathcal{P} / \partial \mathbf{x}$ и $\partial \mathcal{P} / \partial \dot{\mathbf{x}}$. В общем случае они представляют собой весьма сложные выражения. Однако обычно для минимизации целевой функции достаточно знать лишь производные от основных силовых факторов, а именно от гравитационной силы протяженного (центрального) тела с точностью до зональных гармоник низкого порядка (2.5) и от силы притяжения точечной массы (2.7).

Производные от гравитационной силы протяженного тела — это вторые производные от потенциала V (2.1), и они вычисляются по формулам

$$\frac{\partial \mathcal{P}_V}{\partial \mathbf{x}} = \frac{\partial^2 V}{\partial \mathbf{x}^2} = -\mu \sum_n C_n^0 b^n \frac{\partial^2}{\partial \mathbf{x}^2} \frac{L_n(\sin \varphi)}{|\mathbf{x}|^{n+1}},$$

где

$$\begin{aligned} & \frac{\partial^2 L_n(\sin \varphi)}{\partial \mathbf{x}^2} \frac{1}{|\mathbf{x}|^{n+1}} = \\ & = \frac{1}{|\mathbf{x}|^{n+1}} \left[L_n''(\sin \varphi) \left(\frac{\partial}{\partial \mathbf{x}} \sin \varphi \right)^T \frac{\partial}{\partial \mathbf{x}} \sin \varphi + L_n'(\sin \varphi) \frac{\partial^2}{\partial \mathbf{x}^2} \sin \varphi \right] - \\ & - \frac{n+1}{|\mathbf{x}|^{n+3}} \left[L_n'(\sin \varphi) \left(\left(\frac{\partial}{\partial \mathbf{x}} \sin \varphi \right)^T \mathbf{x}^T + \mathbf{x} \frac{\partial}{\partial \mathbf{x}} \sin \varphi \right) + \right. \\ & \quad \left. + L_n(\sin \varphi) \left(\mathbf{E} - (n+3) \frac{\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^2} \right) \right], \\ & \frac{\partial^2}{\partial \mathbf{x}^2} \sin \varphi = -\frac{1}{|\mathbf{x}|^2} \left[\frac{\mathbf{a}\mathbf{x}^T + \mathbf{x}\mathbf{a}^T}{|\mathbf{x}|} + \sin \varphi \left(\mathbf{E} - \frac{3\mathbf{x}\mathbf{x}^T}{|\mathbf{x}|^2} \right) \right]. \end{aligned}$$

Дифференцирование гравитационной силы точечной массы по координатам дает

$$\frac{\partial \mathcal{P}_P}{\partial \mathbf{x}} = -\mu_P \frac{1}{|\mathbf{x}_P - \mathbf{x}|^3} \left[\mathbf{E} - 3 \frac{(\mathbf{x}_P - \mathbf{x})(\mathbf{x}_P - \mathbf{x})^T}{|\mathbf{x}_P - \mathbf{x}|^2} \right],$$

где μ_P и \mathbf{x}_P — гравитационный параметр и вектор положения материальной точки.

Заметим, что здесь запись типа $\mathbf{a}\mathbf{b}^T$ представляет собой тензорное произведение векторов, в отличие от скалярного $\mathbf{a}^T \mathbf{b} = \mathbf{a} \cdot \mathbf{b}$.

Если же взятие производных от силы \mathcal{P} по \mathbf{x} и $\dot{\mathbf{x}}$ в (5.16) проблематично, то можно воспользоваться численным дифференцированием для вычисления производных $\partial \mathbf{x} / \partial \mathbf{q}$. Для этого, как правило, прибегают к разделенным разностям первого порядка:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{q}} \approx \left(\frac{\Delta \mathbf{x}}{\Delta q_1}, \dots, \frac{\Delta \mathbf{x}}{\Delta q_K} \right),$$

где

$$\frac{\Delta \mathbf{x}}{\Delta q_i} \equiv \frac{\mathbf{x}(t, q_1, \dots, q_i + \Delta q_i, \dots, q_K) - \mathbf{x}(t, q_1, \dots, q_i, \dots, q_K)}{\Delta q_i}. \quad (5.17)$$

Таким образом, чтобы получить матрицу производных $\partial \mathbf{x} / \partial \mathbf{q}$ на заданный момент времени, необходимо $K + 1$ раз численно проинтегрировать орбиту с различными наборами параметров: один набор с исходными параметрами и K с проварьированными.

Разделенные разности (5.17) представляют производные с методической ошибкой $O(\Delta q_i)$. Однако точность численного дифференцирования можно повысить до порядка $O(\Delta q_i^2)$, если использовать разделенную разность вида

$$\frac{\Delta \mathbf{x}}{\Delta q_i} \equiv \frac{\mathbf{x}(t, q_1, \dots, q_i + \Delta q_i, \dots, q_K) - \mathbf{x}(t, q_1, \dots, q_i - \Delta q_i, \dots, q_K)}{2\Delta q_i}. \quad (5.18)$$

Как видно, в отличие от (5.17), формула (5.18) предполагает численное интегрирование орбиты большее число раз, а именно $2K$ (орбита с исходными параметрами не требуется).

Очевидный недостаток численного дифференцирования для вычисления изохронных производных — это многократное численное интегрирование орбиты. Кроме того, следует заметить, что каждая вариация для соответствующего параметра должна выбираться индивидуально путем ее подбора по достижении наивысшей точности разделенной разности, т.е. когда методические и вычислительные ошибки станут соизмеримы.

5.5. Методы решения обратной задачи

Численный метод Гаусса–Ньютона (или метод дифференциального исправления орбиты) — наиболее используемый для решения обратных задач в небесной механике, поскольку в большинстве случаев, когда начальные приближения достаточно хорошие, он прекрасно работает и позволяет находить решение за несколько итераций. Между тем при грубых начальных приближениях метод может расходиться, и тогда возникает необходимость прибегать к помощи других альтернативных методов.

Так как целевая функция (5.3) является некоторой мерой оптимальности, фактически методы решения обратных задач (определения орбит), минимизирующие целевую функцию, — это хо-

рошо известные методы оптимизации (Аттетков и др., 2001), которые успешно развиваются вне зависимости от небесной механики. Существует множество самых разнообразных методов, однако мы рассмотрим только весьма ограниченный набор из них, который, на наш взгляд, реально может быть востребован на практике именно для численного решения обратных задач орбитальной динамики.

5.5.1. Метод градиентного спуска

Достаточно простую, но не лучшую итерационную схему для минимизации целевой функции (5.3) дает метод градиентного спуска или, что то же самое, метод простых итерций по отношению к системе (5.4). Согласно этой схеме, поправка к каждому приближению \mathbf{q} вычисляется как

$$\Delta \mathbf{q} = -h \left(\frac{\partial S}{\partial \mathbf{q}} \right)^T. \quad (5.19)$$

Здесь h — неопределенный параметр, шаг спуска. Он обычно выбирается так, чтобы минимизировать целевую функцию вдоль градиента $\partial S / \partial \mathbf{q}$ (метод наискорейшего градиентного спуска), что обеспечивает релаксационную последовательность приближений, т.е. когда целевая функция с каждым следующим приближением убывает. Однако даже при оптимальном выборе шага метод на практике очень медленно сходится, в особенности когда поведение целевой функции достаточно сложное относительно определяемых параметров \mathbf{q} (например, в овражных задачах).

5.5.2. Метод Ньютона

Для решения системы уравнений (5.4) чаще всего прибегают к итерационным методам типа Ньютона (Гилл и др., 1985), где привлекаются вторые производные $\partial^2 S / \partial \mathbf{q}^2$ или их приближения в качестве дополнительной информации, уточняющей локальное поведение целевой функции. Согласно итерационной схеме Ньюто-

на, поправка к текущему приближению \mathbf{q} определяется как

$$\Delta \mathbf{q} = - \left(\frac{\partial^2 S}{\partial \mathbf{q}^2} \right)^{-1} \left(\frac{\partial S}{\partial \mathbf{q}} \right)^T, \quad (5.20)$$

где $\partial^2 S / \partial \mathbf{q}^2$ — матрица Гессе:

$$\frac{\partial^2 S}{\partial \mathbf{q}^2} = 2 \left(\frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right)^T \mathbf{W} \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} - 2(\mathbf{p}^O - \mathbf{p}^C)^T \mathbf{W} \frac{\partial^2 \mathbf{p}^C}{\partial \mathbf{q}^2}. \quad (5.21)$$

5.5.3. Метод Гаусса–Ньютона

Ввиду проблематичности вычисления вторых производных в (5.21) на практике их опускают и, таким образом, используют упрощенную схему так называемого метода Гаусса–Ньютона, который относится к широкому классу квазиньютоновских методов.

Введем матрицы

$$\mathbf{A}_{N \times K} = (\mathbf{W}^{1/2})^T \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \quad \text{и} \quad \mathbf{B}_{N \times 1} = (\mathbf{W}^{1/2})^T (\mathbf{p}^O - \mathbf{p}^C). \quad (5.22)$$

Тогда поправку (5.20) можно переписать как

$$\Delta \mathbf{q} = -\mathbf{Q}^{-1} \mathbf{G}, \quad (5.23)$$

где $\mathbf{Q} = \mathbf{A}^T \mathbf{A}$ — нормальная матрица, которая в минимуме целевой функции $S(\mathbf{q})$ при достаточно малых невязках \mathbf{B} близка к деленной на два матрице Гессе, а $\mathbf{G} = -\mathbf{A}^T \mathbf{B}$ — деленный на два градиент функции S по \mathbf{q} .

Несмотря на то, что метод Гаусса–Ньютона при достаточно малых невязках обеспечивает быструю сходимость итерационного процесса вблизи минимума целевой функции, тем не менее область сходимости метода в нелинейных обратных задачах порой чрезвычайно мала. Для разрешения этой трудности на практике, как правило, прибегают либо к модифицированным методам Гаусса–Ньютона (Ortega, Reinboldt, 2000), либо к составным методам.

5.5.4. Демпфированный метод Гаусса–Ньютона

Демпфирование является одной из модификаций метода Гаусса–Ньютона, которая состоит в уменьшении поправки (5.23) путем введения так называемого демпфирующего множителя h :

$$\Delta \mathbf{q} = -h\mathbf{Q}^{-1}\mathbf{G}, \quad \text{где } 0 < h < 1. \quad (5.24)$$

Примечательно, что при определенном выборе h сходимость итерационного процесса может быть достигнута практически всегда (Ortega, Reinboldt, 2000). Однако очевидный недостаток такого подхода — это значительное понижение скорости сходимости, в особенности когда начальное приближение достаточно далеко от минимума целевой функции. Кроме того, следует заметить, что эффективность метода непосредственно зависит от того, насколько удачно будут выбираться множители h на каждом шаге итерационного процесса.

5.5.5. Метод Левенберга–Марквардта

Альтернативная модификация метода Гаусса–Ньютона реализована в методе Левенберга–Марквардта, который в силу специфики его итерационной схемы наследует взаимодополняющие достоинства методов градиентного спуска (5.19) и Гаусса–Ньютона (5.23): первый, в отличие от второго, может быть эффективным в области, далекой от минимума целевой функции, тогда как преимущество второго над первым проявляется именно в окрестности минимума. Учитывая эти особенности, К. Левенберг (Levenberg, 1944) предложил следующую итерационную схему:

$$\Delta \mathbf{q} = -(\mathbf{Q} + h\mathbf{E})^{-1}\mathbf{G}, \quad \text{где } 0 < h < \infty, \quad (5.25)$$

а \mathbf{E} — единичная матрица. Как видно, при $h \rightarrow 0$ метод Левенберга вырождается в метод Гаусса–Ньютона, тогда как при $h \rightarrow \infty$ — в метод градиентного спуска с шагом $1/h$. Таким образом, итерационный процесс по схеме (5.25) реализуется с выбором h от бесконечно больших значений до нуля.

Чтобы улучшить свойства схемы (5.25), Д. Марквардт (Marquardt, 1963) предложил вместо единичной матрицы использовать диагональную \mathbf{Q}^* , составленную из диагональных элементов матрицы \mathbf{Q} :

$$\Delta \mathbf{q} = -(\mathbf{Q} + h\mathbf{Q}^*)^{-1} \mathbf{G}. \quad (5.26)$$

К сожалению, как показывает практика, метод Левенберга–Марквардта часто не сходится (Демиденко, 1989)³⁷. Причина этому — постоянство длины шага. Поэтому, как правило, на практике используют так называемый модифицированный метод Левенберга–Марквардта³⁸ (Bard, 1974) со схемой

$$\Delta \mathbf{q} = -h(\mathbf{Q} + h^*\mathbf{Q}^*)^{-1} \mathbf{G}, \quad \text{где } 0 < h \leq 1, \quad 0 < h^* < \infty. \quad (5.27)$$

Очевидно, как и в демпфированном методе Гаусса–Ньютона, главным недостатком этого метода является неясная стратегия выбора параметров h и h^* , хотя при ее удачной выработке в некоторых случаях можно добиться весьма высокой скорости сходимости итерационного процесса.

5.5.6. Овражные методы

Среди всех обратных задач особого внимания заслуживают овражные задачи. Формально овражность обнаруживается в значениях (неотрицательных) собственных чисел нормальной матрицы \mathbf{Q} , характеризующей локальное поведение целевой функции около решения $\hat{\mathbf{q}}$:

$$S(\mathbf{q}) \approx S(\hat{\mathbf{q}}) + (\mathbf{q} - \hat{\mathbf{q}})^T \mathbf{Q}(\hat{\mathbf{q}})(\mathbf{q} - \hat{\mathbf{q}}), \quad (5.28)$$

где, очевидно, в квадратичной форме линейный член равен нулю, поскольку $\mathbf{G}(\hat{\mathbf{q}}) = \mathbf{0}$. Из (5.28), в частности, видно, что целевая функция в окрестности минимума ведет себя как многомерный эллиптический параболоид, а ее поверхности постоянных значений

³⁷Мы также столкнулись с этой проблемой, о чем будет сказано ниже.

³⁸В (Ortega, Reinboldt, 2000) его называют модифицированным методом Гаусса–Ньютона.

близки к эллипсоидам. Число обусловленности $\text{cond} \mathbf{Q}$ как отношение максимального собственного числа к минимальному определяет степень овражности целевой функции. Большое число обусловленности сигнализирует о сильной овражности. Фактически, это означает, что существуют такие направления в параметрическом пространстве, вдоль которых целевая функция меняется чрезвычайно медленно. В данном случае целесообразно анализировать направления, задаваемые собственными векторами \mathbf{Q} , которые ортогональны друг другу ввиду симметричности нормальной матрицы. Число направлений медленного изменения целевой функции называется размерностью оврага (Гельфанд, Цетлин, 1962). Естественно, что оно формально должно определяться количеством существенно малых собственных чисел нормальной матрицы в сравнении с другими ее собственными числами.

Направления медленного изменения целевой функции задают так называемый хребет оврага. Если хребет не плоский, т.е. его нельзя натянуть на некоторую прямоугольную систему координат, по крайней мере, в окрестности минимума целевой функции, то возникают серьезные сложности в итерационном поиске решения обратной задачи, особенно в окрестности хребта, которому, кстати, принадлежит и само решение.

Можно выделить две типичные овражные обратные задачи, довольно часто встречающиеся в небесной механике. Первая задача (единичной размерности) возникает при обработке астрометрических наблюдений, распределенных на очень короткой орбитальной дуге. В данном случае направление одномерного хребта оврага в параметрическом пространстве определяется главным образом топоцентрическим вектором положения, вдоль которого отсутствует информация о движении наблюдаемого объекта. На практике такие задачи обычно решаются с помощью демпфированной схемы Гаусса–Ньютона (5.24). Основные затраты компьютерного времени в решении обратной задачи приходятся на интегрирование орбиты, необходимое для представления наблюдений. Однако, поскольку орбита интегрируется на очень коротком интервале времени, весь вычислительный процесс выполняется до-

статочно быстро, даже если минимизация целевой функции предполагает большое количество итераций (100 и более).

Серьезные трудности возникают в задачах второго типа, а именно при определении орбит внутренних спутников планет из астрометрических наблюдений, рассредоточенных на длительном интервале времени. Размерность оврагов в таких обратных задачах, как правило, на единицу меньше числа определяемых орбитальных параметров, т.е. равна 5, а целевая функция очень чувствительна вдоль направления, задаваемого частотной переменной типа среднего движения, и чем больше период наблюдений, тем выше чувствительность целевой функции (Авдюшев, Баньщикова, 2008). Численное интегрирование орбиты близкого спутника для представления его наблюдений (на одной итерации) может потребовать до нескольких часов компьютерного времени и поэтому использование демпфированного метода Гаусса–Ньютона практически нецелесообразно. Впрочем, другие методы, представленные выше, также оказываются неэффективными. В данном случае необходимы специально разработанные методы для минимизации овражных целевых функций.

Хотелось бы, конечно, отметить, что проблема овражности может быть разрешена путем выбора подходящих определяемых орбитальных параметров. Однако, если теоретически такое возможно, то практически эта задача трудно решается и ее, как правило, исследователь даже не ставит перед собой. В то же время полезно знать, что степень овражности имеет связь с начальной эпохой t_0 , на которую отнесены орбитальные параметры \mathbf{q} . В работах (Черницов, 1975; Авдюшев, Баньщикова, 2008) показано, что степень овражности минимальна для эпохи, представляющей собой среднее арифметическое всех моментов наблюдений. Таким образом, начальная эпоха должна лежать, по крайней мере, внутри периода наблюдений, тогда как вынесение ее далеко за пределы периода будет приводить к усилению овражности целевой функции.

При решении овражных задач хорошо зарекомендовал себя метод Гельфанда–Цетлина (Гельфанд, Цетлин, 1962). Идея метода состоит в поиске минимума целевой функции вдоль ее оврага.

Сначала выбираются два близких приближения, которые затем спускаются на дно оврага методом наискорейшего градиентного спуска. С этой задачей метод справляется достаточно быстро, несмотря на то что внутри оврага работает чрезвычайно неэффективно, поскольку его итерационный процесс почти заклинивается. Далее вдоль направления, задаваемого двумя полученными овражными приближениями, определяется третье приближение, которое, вообще говоря, не является овражным, и поэтому после (как и начальные приближения) оно спускается на дно оврага. Таким же способом находится каждое следующее овражное приближение по двум предыдущим, и так до сходимости итерационного процесса. Простота метода Гельфанда–Цетлина, конечно же, подкупает. Между тем и главный недостаток его вполне очевиден, а именно отсутствие универсальной стратегии выбора длины овражного шага, которая непосредственно зависит от решаемой обратной задачи.

Впрочем, авторы овражного метода указывают, что при решении конкретной задачи обычно удается разработать специальные и эффективные методы, основанные на некоторых априорных знаниях о поведении целевой функции. Так, в обратных задачах динамики близких спутников известно, что вдоль оврага частотная переменная почти постоянна и, следовательно, эту информацию необходимо использовать для выполнения овражных шагов.

С целью достижения быстрой сходимости итерационного процесса можно совместно с методом Гаусса–Ньютона последовательно использовать другие, а именно метод градиентного спуска и так называемый проекционный метод (Himmelblau, 1972).

Представим схему градиентного спуска (5.19) в виде

$$\Delta \mathbf{q} = - \frac{\mathbf{G} \cdot \mathbf{G}}{\mathbf{Q} \mathbf{G} \cdot \mathbf{G}} \mathbf{G}. \quad (5.29)$$

В случае, когда целевая функция описывается квадратичной формой, она соответствует наискорейшему градиентному спуску (Аттетков и др., 2001). В процессе использования схемы (5.29) приближенное решение достаточно быстро (за несколько итераций)

сваливается на дно оврага целевой функции, но в дальнейшем невероятно медленно сходится к ее минимуму. На этом этапе можно было бы воспользоваться методом Гаусса–Ньютона. Однако, как показывает практика, если приближенное решение оказывается довольно далеко от минимума целевой функции, схема (5.23) будет давать такие поправки, которые способны выкинуть решение из области сходимости метода.

Впрочем, как отмечалось выше, специфика рассматриваемых обратных задач такова, что вдоль оврага целевой функции частотная переменная почти постоянна, поэтому схема градиентного спуска позволяет очень быстро получить уже достаточно хорошую ее оценку. Следовательно, для дальнейшей минимизации целевой функции необходимо учитывать эту особенность и выполнять итерационный процесс с налагаемыми ограничениями, с тем чтобы сохранялось значение частотной переменной.

В качестве такого ограничения примем полную энергию $H(\mathbf{q})$, которая непосредственно связана с частотной переменной, а итерационный процесс будем выполнять по схеме Гаусса–Ньютона с поправкой каждого приближения за его отклонение от энергетической поверхности в фазовом пространстве \mathbf{q} . Этот прием обеспечивает быстрый спуск к минимуму целевой функции вдоль оврага.

В общем случае налагаемое ограничение определяет множество поправок, однако из них целесообразно использовать лишь ортогональную проекцию приближенного решения (5.23) на поверхность $H(\mathbf{q})$. Алгоритмически это возможно по следующей итерационной схеме:

$$\Delta \mathbf{q}_{i+1} = \Delta \mathbf{q}_i - \frac{H(\mathbf{q} + \Delta \mathbf{q}_i) - H(\mathbf{q})}{\mathbf{G}_H \cdot \mathbf{G}_H} \mathbf{G}_H \quad (i = 0, 1, \dots), \quad (5.30)$$

где $\mathbf{G}_H = (\partial H / \partial \mathbf{q})^T$ — градиент H по \mathbf{q} , а $\Delta \mathbf{q}_0$ определяется из (5.23). Как только приближенное решение становится достаточно близким к минимуму, что определяется малостью величины $\Delta \mathbf{q}_0$, заканчиваем процесс минимизации S по схеме Гаусса–Ньютона (5.23) без ограничений.

5.6. Сравнительный анализ эффективности методов

В данном разделе приводятся результаты численного эксперимента по исследованию эффективности рассмотренных методов применительно к решению одной сложной овражной задачи, в которой требовалось по астрометрическим наблюдениям, покрывающим длительный период времени, определить вектор начального динамического состояния $\mathbf{q} = (\mathbf{x}_0, \dot{\mathbf{x}}_0)^T$ Адрастеи, внутреннего спутника Юпитера.

Движение спутника моделировалось на основе формул задачи двух тел (см. Приложение). Модель представляла движение спутника Адрастеи в пространстве угловых координат. Из наземных наблюдений методом Лапласа (Escobal, 1965) (см. также Приложение) предварительно были получены орбитальные параметры спутника, которые соответствовали почти круговой орбите с большой полуосью $\bar{a} = 8.68 \cdot 10^{-4}$ а.е. и эксцентриситетом $\bar{e} = 0.0161$. На основе этих параметров моделировались (точные) наблюдаемые положения $\mathbf{p}^O = (\alpha^O, \delta^O)^T$ в моменты реальных наблюдений ($J = 90, N = 180$), рассредоточенных на концах интервала времени 12 лет как две почти равновесные группы, при этом начальная эпоха t_0 была взята в середине временного интервала.

Сходимость методов исследовалась для двух довольно грубых начальных приближений $(\mathbf{q}_0)_1$ и $(\mathbf{q}_0)_2$, соответствующих орбитальным параметрам $\alpha_1 = -10^{-5}$, $e_1 = 0.1$; и $\alpha_2 = 10^{-5}$, $e_2 = 0.1$, где $\alpha = (a - \bar{a})/\bar{a}$. Итерационный процесс для каждого метода завершался по достижении точности в координатах 10^{-10} а.е. Мониторинг итерационного процесса проводился в плоскости (α, e) . Результаты приведены на рис. 5.1–5.3. Здесь также представлены графики изолиний S как функции α и e при истинных значениях других орбитальных параметров.

В первую очередь необходимо сказать, что итерационный процесс ни для метода наискорейшего градиентного спуска (5.29), ни для метода Гаусса–Ньютона (5.23) не достигал сходимости. Несмотря на то что первые приближения метода градиентного спуска быстро сваливались на дно оврага, далее следовало зацик-

Таблица 5.1. Характеристики быстродействия методов

Метод	Начальное приближение			
	$(\mathbf{q}_0)_1$		$(\mathbf{q}_0)_2$	
	<i>NS</i>	<i>NF</i>	<i>NS</i>	<i>NF</i>
Гаусса–Ньютона	Не сходится		Не сходится	
Гаусса–Ньютона ($h = 10^{-3}$)	11127	11127	10590	10590
Гаусса–Ньютона ($h \neq \text{const}$)	6693	7155	5952	6362
Левенберга–Марквардта	1123	1155	Не сходится	
Левенберга–Марквардта–Гаусса	816	853	816	853
Составной	14	14	14	14

ливание итерационного процесса. В то же время метод Гаусса–Ньютона давал настолько большие поправки, что приближения вовсе не соответствовали объекту Солнечной системы.

Демпфирование же метода Гаусса–Ньютона (5.24) обеспечивало сходимости итераций только при множителях порядка $h = 0.001$ и меньше. На рис. 5.1 показаны траектории приближений метода при различных h . Из рисунка, в частности, видно, что для $h = 0.001$ последовательные приближения с начальным $(\mathbf{q}_0)_1 = (\alpha_1, e_1)^T$ определяются весьма непредсказуемым образом, хотя итерационный процесс все же сходится. В то же время существенное уменьшение множителя ($h = 0.0001$) как для $(\mathbf{q}_0)_1$, так и для $(\mathbf{q}_0)_2$ приводит к желаемой релаксационной последовательности приближений.

Несмотря на столь простой подход в решении проблемы сходимости, он, как уже отмечалось, все же неприемлем для определения орбит реальных близких спутников, поскольку в силу специфики обратной задачи сопряжен с очень низкой скоростью сходимости. Так, при использовании схемы с множителем $h = 0.001$ для сходимости итерационного процесса (с принятой координатной точностью 10^{-10} а.е.) нам потребовалось выполнить более 10000 итераций (*NS* в табл. 5.1).

Впрочем, скорость сходимости можно повысить, если использовать переменный множитель h . После многочисленных экспе-

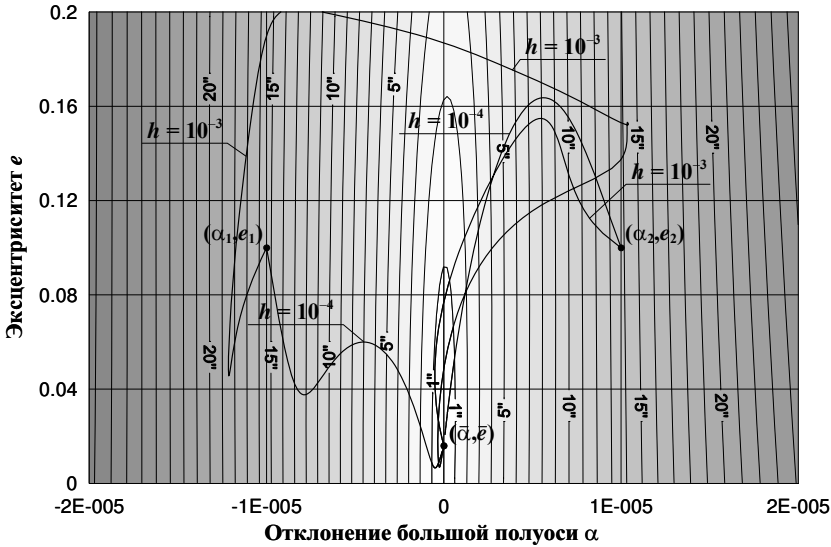


Рис. 5.1. Сходимость демпфированного метода Гаусса–Ньютона с постоянным h (точные наблюдения, $N = 90$)

риментов была выработана следующая стратегия выбора шага, обеспечивающая наименьшее число итераций NS . В качестве коэффициента увеличения h на каждом следующем шаге принималась величина 1.05. Если значение целевой функции на текущем шаге становилось больше, нежели на предыдущем, для обеспечения релаксационной последовательности приближений поправка (5.24) перевычислялась с уменьшенным вдвое множителем h . Такие ситуации часто имели место, и поэтому число вычислений целевой функции NF оказалось заметно больше, нежели число шагов NS . При этом начальное значение множителя было равно 10^{-3} . Таким образом, использование переменного множителя h позволило повысить быстродействие вычислительного процесса приблизительно всего лишь в 1.5 раза (табл. 5.1).

Вообще низкая эффективность демпфированного метода Гаусса–Ньютона объясняется тем, что при грубом начальном прибли-

жении ньютоновское направление не является оптимальным для поиска минимума целевой функции (см. рис. 5.2). В этом смысле, принимая во внимание вышеуказанные достоинства метода Левенберга–Марквардта (5.26), можно полагать, что его эффективность будет гораздо выше. Как и для демпфированного метода Гаусса–Ньютона, для метода Левенберга–Марквардта выбирались такие коэффициенты увеличения и уменьшения, а также начальное значение множителя h , которые обеспечивали наилучшую скорость сходимости. Интересно заметить, что они оказались обратными к соответствующим параметрам в демпфированном методе. Причем сначала параметры в методе Левенберга–Марквардта подбирались применительно к начальному приближению $(\mathbf{q}_0)_1$. Таким образом, удалось повысить скорость итерационного процесса почти в 7 раз (табл. 5.1).

В данном случае повышение эффективности вычислений оказывается возможным главным образом за счет того, что на первых итерациях для поиска минимума целевой функции выбирается направление градиентного спуска (рис. 5.2). Далее же, при движении вдоль оврага, скорость итерационного процесса становится приблизительно такой же, как и для метода Гаусса–Ньютона с переменным демпфирующим множителем.

К сожалению, какие бы параметры не подбирались в случае начального приближения $(\mathbf{q}_0)_2$, итерационный процесс не сходился и закидывался на дне оврага в окрестности точки $\alpha = 0$, $e = 0.1$. Для разрешения этой трудности использовался модифицированный метод Левенберга–Марквардта (5.27), который фактически является гибридом, полученным из методов Гаусса–Ньютона и Левенберга–Марквардта. С учетом указанной выше связи между параметрами этих методов стратегия выбора параметров h и h^* в (5.27) принималась в соответствии с соотношением $h^* = c/h$, где $c = 10^{-8}$ — коэффициент, подобранный опытным путем. Таким образом, с помощью модифицированного метода удалось добиться сходимости итерационного процесса для $(\mathbf{q}_0)_2$. Кроме того, если сравнивать результаты с полученными выше для $(\mathbf{q}_0)_1$, можно сказать, что модификация позволяет заметно повы-

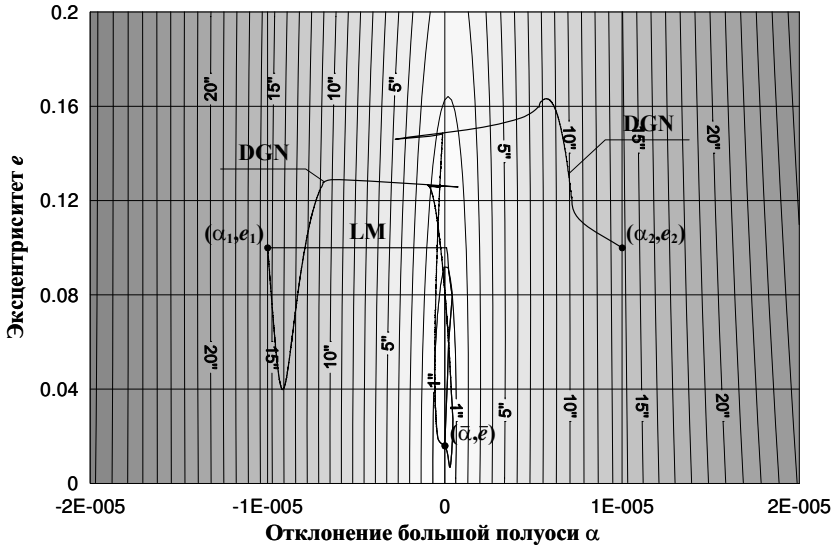


Рис. 5.2. Сходимость демпфированного метода Гаусса–Ньютона с переменным h (DGN) и метода Левенберга–Марквардта (LM)

суть быстрдействие вычислений (табл. 5.1). Тем не менее даже в этом случае скорость сходимости весьма низкая.

Впечатляющие результаты получаются при использовании составного метода со схемами (5.23), (5.29) и (5.30) (Авдюшев, Банныцкова, 2008): для сходимости в обоих случаях начальных приближений потребовалось всего 14 итераций. Характер сходимости метода представлен на рис. 5.3 (Complex). Горизонтальные траектории здесь соответствуют градиентному спуску (5.29), тогда как вертикальные — схеме метода Гаусса–Ньютона совместно с проекционной схемой (5.30), где в качестве H была взята кеплеровская энергия. При этом схема градиентного спуска использовалась, когда изменение среднеквадратической величины $\sigma = \sqrt{S/N}$ после выполнения очередного шага становилось меньше $0.001''$, что рассматривалось как признак попадания приближения в овраг. Далее, применялась схема Гаусса–Ньютона. Если поправки по коор-

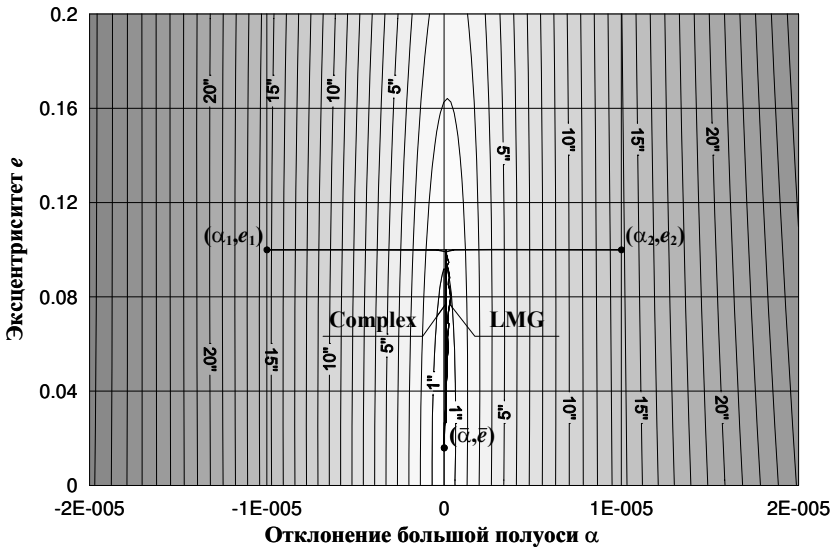


Рис. 5.3. Сходимость модифицированного метода Гаусса–Ньютона (Левенберга–Марквардта–Гаусса) (LMG) и составного метода (Complex)

динамам превышали величину 10^{-8} а.е., применялась коррекция решения по проекционной схеме (5.30). Ввиду высокой эффективности составного подхода именно его следует рекомендовать для определения орбит близких спутников планет.

5.7. Проблема неоднозначного определения орбит

Главным образом ввиду того, что решаемая задача нелинейна, она может иметь множество решений, иначе говоря, целевая функция может иметь множество минимумов. Формально в качестве нужного решения, представляющего наиболее точные значения орбитальных параметров, выбирают, очевидно, то, которое обеспечивает наименьшее значение целевой функции.

Стоит отметить, что в большинстве случаев (например, при определении астероидных орбит) проблема выбора нужного решения совершенно не возникает: необходимые для итерационного

метода начальные приближения, получаемые из предварительного определения орбит, часто настолько точны, что итерационный процесс быстро сходится к нужному решению, доставляющему абсолютный минимум целевой функции.

Тем не менее имеется ряд задач, когда такой формальный выбор решения не является безусловным, поскольку имеются другие альтернативные решения, для которых целевая функция принимает значения, очень близкие к значению абсолютного минимума. Сомнения здесь вызываются, прежде всего, тем обстоятельством, что наблюдения могут содержать ошибки, при которых решение, доставляющее абсолютный минимум целевой функции, т.е. наилучшим образом представляющее наблюдения, вполне возможно будет далеко не самым лучшим с точки зрения прогнозирования движения небесного тела в будущем. В работе (Авдюшев, Баньщикова, 2008), например, показано, что проблема неоднозначного определения орбит может иметь место в задачах динамики близких спутников планет при малочисленном составе наблюдений, моменты которых рассредоточены в нескольких группах на достаточно большом интервале времени порядка 10000 оборотов объекта и более.

5.7.1. Наблюдения на длительном временном интервале

Пусть в некотором пространстве \mathbf{p} известны N положений небесного объекта \mathbf{p}_i^O в моменты времени t_i ($i = 1, \dots, N$), при этом орбитальное движение объекта описывается моделью

$$\mathbf{p}^C = \mathbf{p}^C(t, \mathbf{q}, E), \quad E = n(t - t_0) + E_0, \quad (5.31)$$

где \mathbf{q} — орбитальные параметры, однозначно определяющие траекторию небесного тела; E — аномалия (быстрая переменная); n — частота; E_0 — аномалия E в эпоху t_0 . При этом \mathbf{p}^C 2π -периодична по E , т.е. $\mathbf{p}^C(t, \mathbf{q}, E) = \mathbf{p}^C(t, \mathbf{q}, E + 2\pi)$.

Положим пока, что параметры \mathbf{q} не зависят от n , и рассмотрим квадраты расстояний $\rho_i^2 = \rho^2(\mathbf{p}_i^O, \mathbf{p}_i^C)$ между положениями \mathbf{p}_i^O и \mathbf{p}_i^C как функции n : $\rho_i^2 = \rho_i^2(n)$. Нетрудно видеть, что функции

ρ_i^2 периодичны по n с соответствующими периодами $2\pi/(t_i - t_0)$. В силу этого, очевидно, среднеквадратическая величина

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N \rho_i^2, \quad (5.32)$$

рассматриваемая в данном случае как целевая функция, будет иметь бесконечное множество минимумов вдоль n . Следовательно, минимизация σ^2 по n будет давать множество решений, из которых только одно соответствует истинной частоте \bar{n} . Причем на равномерной последовательности моментов наблюдений t_i могут иметь место такие частоты $n \neq \bar{n}$, при которых орбитальная модель будет точно представлять наблюдения, как и при \bar{n} . Это явление хорошо известно в спектральном анализе временных рядов под названием «алиасинг» (Витязев, 2001).

Орбиты близких спутников могут быть также представлены в виде (5.31), где, однако, некоторые из параметров \mathbf{q} (такие, как, например, большая полуось или фокальный параметр) непосредственно связаны с частотой n . Поэтому, вообще говоря, ρ_i^2 оказываются непериодичными по n . Тем не менее при достаточно больших величинах $t_i - t_0$ функции ρ_i^2 в окрестности истинной частоты \bar{n} будут очень близки к рассмотренным выше периодическим составляющим, и, следовательно, проблема множества минимумов для σ^2 здесь также будет иметь место.

Характерная особенность в поведении σ^2 относительно n раскрывает генезис проблемы множества решений, которая реально может возникать в обратных задачах динамики близких спутников. Действительно, на практике орбитальные параметры спутника, как правило, определяются из условия достижения минимума некоторой функции вида (5.32), которая выражает степень близости наблюдаемых и моделируемых положений объекта в пространстве \mathbf{p} . По крайней мере, один из параметров обязательно связан с частотой n , но это обстоятельство как раз и становится причиной неоднозначного определения спутниковой орбиты.

Рассмотрим подробнее проблему множества решений на примере круговой задачи, где легко удается получить довольно ясные

и в то же время полезные для практики результаты, позволяющие, кроме того, оценить всю важность исследуемой проблемы.

В комплексной плоскости круговую орбиту можно представить как

$$x = p^C = ae^{iE}, \quad E = n(t - t_0) + E_0, \quad n = \sqrt{\mu/a^3}, \quad (5.33)$$

где x — положение точки; a — радиус орбиты (большая полуось); E — быстрая фаза (аномалия); n — частота обращения (среднее движение); t — время; E_0 — фаза на начальный момент времени t_0 ; μ — гравитационный параметр.

Рассмотрим семейство орбит вида (5.33) и оценим величину разности положений x и \bar{x} на орбитах (a, E_0) и (\bar{a}, \bar{E}_0) соответственно. Согласно (5.33), разность $\Delta x = x(t, a, E_0) - \bar{x}(t, \bar{a}, \bar{E}_0)$ нетрудно преобразовать к виду

$$\Delta x = \bar{a}e^{iE}[1 + \alpha - e^{-i(\lambda\nu + \beta)}], \quad (5.34)$$

где $\alpha = (a - \bar{a})/\bar{a}$; $\beta = E_0 - \bar{E}_0$; $\nu = (n - \bar{n})/\bar{n}$; $\lambda = \bar{n}(t - t_0)$. Полагая, что $\alpha \ll 1$, с точностью до первого порядка малости будем иметь оценку $2\nu = -3\alpha$. Тогда в соответствии с (5.34) квадрат величины Δx будет

$$|\Delta x|^2 = \bar{a}^2[\alpha^2 + 2(1 + \alpha)(1 - \cos \varphi)], \quad \varphi = \frac{3}{2}\lambda\alpha - \beta. \quad (5.35)$$

Отсюда видно, что поведение $|\Delta x|^2$ при малых α главным образом будет определяться функцией

$$f = \alpha^2 + 2(1 + \alpha)(1 - \cos \varphi). \quad (5.36)$$

На рис. 5.4 показано поведение функции f в зависимости от α и β при $\lambda = 10$. Как видно, рельеф поверхности, задаваемый f , имеет овражную структуру. При этом следует заметить, что увеличение λ приводит к повышению степени овражности f .

Теперь положим, что решение \bar{x} представляет наблюдаемые положения небесного объекта (p^O) на комплексной плоскости в некоторые моменты t_i , и поставим перед собой задачу отыскать

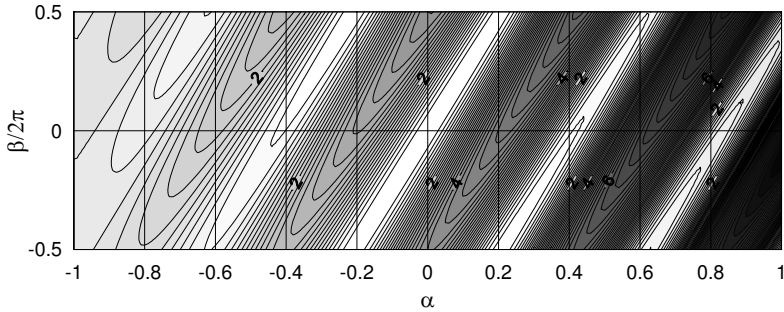


Рис. 5.4. Поведение функции $f = \alpha^2 + 2(1 + \alpha)(1 - \cos \varphi)$ при $\lambda = 10$

такие значения параметров α и β , которые бы доставляли минимум целевой функции

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N |\Delta x|_i^2, \quad (5.37)$$

где N — число наблюдений. Будем рассматривать такие наблюдения, для которых временная удаленность от начальной эпохи такова, что $|\lambda|$ — достаточно большие числа. Тогда поведение σ^2 в зависимости от α и β будет определяться главным образом тригонометрическими составляющими. Поэтому фактически нас должна интересовать функция вида

$$f_{\cos} = 1 - \frac{1}{N} \sum_{i=1}^N \cos \varphi_i, \quad \varphi_i = \frac{3}{2} \lambda_i \alpha - \beta. \quad (5.38)$$

Следует заметить, что при достаточно больших λ_i , отличающихся друг от друга на существенно меньшие величины, функция будет вести себя почти как одна тригонометрическая составляющая:

$$f_{\cos} \approx 1 - \cos \varphi,$$

где в качестве λ выбирается одно из значений λ_i . Следовательно, все наблюдения с такими λ_i можно объединить в группу и рассматривать как одно наблюдение. Если таких групп несколько, то

поведение функции можно представить как

$$f_{\cos} \approx F \equiv 1 - \sum_{j=1}^M k_j \cos(\omega_j \alpha - \beta), \quad k_j = \frac{N_j}{N}, \quad \omega_j = \frac{3}{2} \lambda_j, \quad (5.39)$$

где M — число групп; k_j — вес j -группы; определяемый числом наблюдений в группе N_j ; λ_j — одно из значений λ j -группы. При этом будем полагать, что $\lambda_1 \ll \lambda_2 \ll \dots \ll \lambda_M$.

Используя известные тригонометрические тождества, F можно переписать как

$$F = 1 - A \cos(\beta + \psi), \quad \text{где} \quad A = \sqrt{c^2 + s^2}, \quad \psi = \operatorname{arctg}(s/c), \quad (5.40)$$

$$c = \sum_{j=1}^M k_j \cos(\omega_j \alpha), \quad s = \sum_{j=1}^M k_j \sin(\omega_j \alpha).$$

Отсюда следует, что F 2π -периодична по β и при фиксированном α имеет единственный минимум на полуинтервале $(\beta_0 - \pi, \beta_0 + \pi]$ для любых $\beta_0 \in (-\infty, +\infty)$. Поэтому область исследования F по β можно ограничить до любого полуинтервала длиной в 2π .

Введем характеристику

$$\Phi(\alpha) = \min_{\beta \in (-\pi, +\pi]} F(\alpha, \beta), \quad (5.41)$$

которая примечательна тем, что не зависит от выбора начальной эпохи. Действительно, изменение начальной эпохи $t_0 \rightarrow t_0 + \Delta t$, согласно (5.39), приводит к преобразованию сдвига вдоль β :

$$F(\alpha, \beta)|_{t_0 + \Delta t} = F(\alpha, \beta + \Delta\beta)|_{t_0}, \quad \text{где} \quad \Delta\beta \propto \alpha \Delta t.$$

Тогда ввиду указанных выше свойств F

$$\min_{\beta \in (-\pi, +\pi]} F(\alpha, \beta)|_{t_0 + \Delta t} = \min_{\beta \in (-\pi, +\pi]} F(\alpha, \beta + \Delta\beta)|_{t_0} = \Phi(\alpha).$$

Заметим, что характеристика $\Phi(\alpha)$ является однозначной функцией α , поэтому минимумы $\Phi(\alpha)$ по α будут однозначно соответствовать минимумам $F(\alpha, \beta)$ по α и $\beta \in (-\pi, +\pi]$. Из (5.40) легко

получить формулу для Φ :

$$\Phi(\alpha) = 1 - A(\alpha). \quad (5.42)$$

Нетрудно также видеть, что коэффициенты c и s можно представить как

$$c = \frac{1}{N} \sum_{i=1}^N \cos(\omega_i \alpha), \quad s = \frac{1}{N} \sum_{i=1}^N \sin(\omega_i \alpha).$$

Интересно отметить, что функция Φ относительно переменной α , непосредственно связанной с частотой обращения спутника, может рассматриваться как аналог периодограмм, используемых в спектральном анализе для исследования периодичностей в измерительных данных (Витязев, 2001).

Необходимо также обратить особое внимание на то, что рельеф поверхности, задаваемой функцией $F(\alpha, \beta)$, в окрестностях минимумов имеет сильно овражную структуру. Это главным образом обусловлено большими числами ω_j . Действительно, с точностью до малых второго порядка поведение $F(\alpha, \beta)$ в окрестности тривиального минимума $\alpha = \beta = 0$ может быть представлено функцией

$$Q(\alpha, \beta) = \frac{1}{2}(\kappa_{11}\alpha^2 + 2\kappa_{12}\alpha\beta + \kappa_{22}\beta^2), \quad (5.43)$$

где

$$\kappa_{11} = \sum_{j=1}^M k_j \omega_j^2, \quad \kappa_{12} = - \sum_{j=1}^M k_j \omega_j, \quad \kappa_{22} = 1. \quad (5.44)$$

Уравнение $Q(\alpha, \beta) = \text{const}$ описывает некоторый эллипс в плоскости (α, β) с центром в тривиальном минимуме. Из аналитической геометрии известно, что квадрат отношения C большой полуоси эллипса к малой представим в виде

$$C^2 = \frac{\kappa_{11} + \kappa + \kappa_{22}}{\kappa_{11} - \kappa + \kappa_{22}}, \quad \text{где} \quad \kappa = \sqrt{4\kappa_{12}^2 + (\kappa_{11} - \kappa_{22})^2}.$$

Отсюда, учитывая (5.44), при достаточно больших ω_j будем иметь

$$C^2 \sim \kappa_{11}. \quad (5.45)$$

Впрочем, степень овражности можно слегка уменьшить путем выбора начальной эпохи t_0 . Согласно приближенной оценке (5.45), отношение C^2 будет близко к минимальному при

$$t_0 = \sum_{j=1}^M k_j t_j,$$

где t_j — один из временных моментов наблюдений j -группы. Фактически, это означает, что для уменьшения овражности $F(\alpha, \beta)$ в качестве начальной эпохи следует выбирать среднее арифметическое всех моментов наблюдений. Из известных автору источников об этом впервые упоминается в (Черницов, 1975). Кстати, такой выбор начальной эпохи приводит квадратическую форму (5.43) к каноническому виду, поскольку κ_{12} обращается в нуль.

Для исследования функции F удобнее рассматривать ее в виде

$$F = 1 - \sum_{j=1}^M k_j \cos((l_j - l_0)\zeta - \beta), \quad (5.46)$$

$$\zeta = \frac{3}{2}\alpha(\lambda_M - \lambda_1), \quad l_j = \frac{\lambda_j - \lambda_1}{\lambda_M - \lambda_1} \quad (j = 0, \dots, M), \quad \lambda_0 = 0.$$

Здесь величины l_0 и $l_j \in [0, 1]$ ($j = 1, \dots, M$) представляют временное распределение соответственно начальной эпохи и групп наблюдений относительно первой группы.

Рассмотрим случай двух групп наблюдений ($M = 2$) одинакового веса ($k = 1/2$). Этот случай соответствует распределению имеющихся на данный момент наземных наблюдений близкого спутника Юпитера Адрастеи (J15). На рис. 5.5 представлено поведение функции $F(\zeta, \beta)$ (5.46) при различных $l_0 \in [0, 1/2]$.

Из рисунка видно, что изменение параметра l_0 , т.е. изменение начальной эпохи, приводит к сдвигу рельефа поверхности

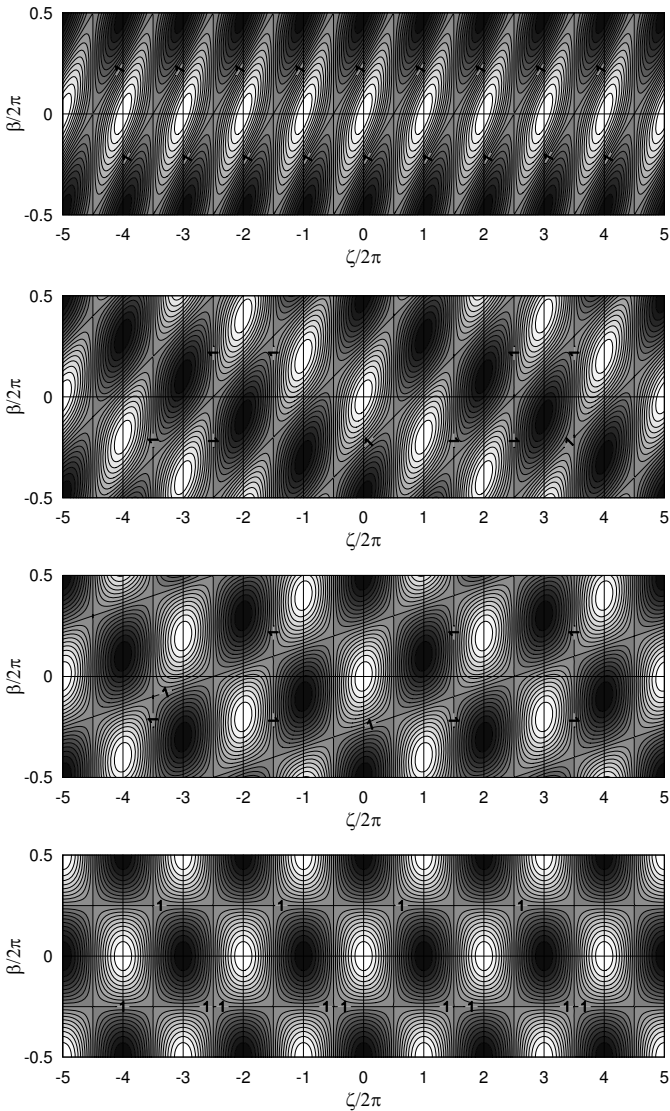


Рис. 5.5. Поведение $F(\zeta, \beta)$ для двух групп наблюдений одинакового веса и различных l_0 : 0.0, 0.2, 0.4 и 0.5 (сверху вниз соответственно) (спутник Адрастея)

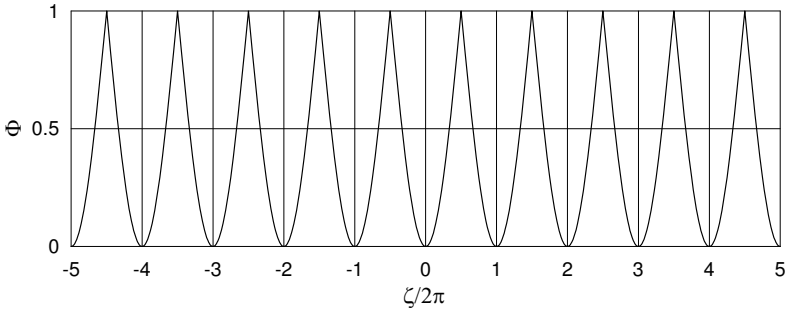


Рис. 5.6. Поведение $\Phi(\zeta)$ для двух групп наблюдений одинакового веса (спутник Адрастея)

$F(\zeta, \beta)$ вдоль β . При этом меняется и степень овражности рельефа: так, овраги при $l_0 = 0$ (начальная эпоха вблизи моментов первой группы наблюдений) заметно более вытянутые, нежели при $l_0 = 1/2$ (начальная эпоха находится в центре рассматриваемого временного интервала). Эти результаты подтверждают наши выводы, сделанные выше, о поведении функции F .

В то же время характеристика Φ остается неизменной для любых l_0 (рис. 5.6). В общем случае, при различных $k_1 = k$ и $k_2 = 1 - k$, где $k \in (0, 1)$, характеристика Φ , согласно (5.40) и (5.42), представима как

$$\Phi = 1 - \sqrt{1 - 2k(1 - k)(1 - \cos \zeta)}.$$

Из наличия тригонометрической составляющей в характеристике Φ следует, что для любых двух групп наблюдений существует множество равнозначных минимумов F , которые соответствуют значениям $\zeta = 2\pi m$ или

$$\alpha = \frac{4\pi m}{3(\lambda_M - \lambda_1)} = \frac{2m}{3R}, \quad (5.47)$$

где m — целые числа, а R — число оборотов объекта за $\lambda_M - \lambda_1$.

Случай трех групп наблюдений (даже равновесных) гораздо более сложен для исследования. Поэтому рассмотрим только

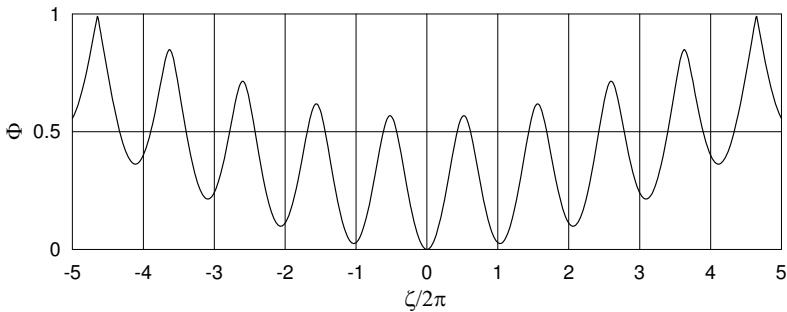


Рис. 5.7. Поведение $\Phi(\zeta)$ для трех групп наблюдений (спутник Метиды)

частный случай, соответствующий распределению наземных наблюдений спутника Метиды (J16): $k_1 = 0.28$, $k_2 = 0.16$, $k_3 = 0.56$, $l_1 = 0$, $l_2 = 0.92$, $l_3 = 1$.

Характеристика Φ (рис. 5.7) так же, как и в случае двух групп наблюдений, имеет множество минимумов. Правда, отличительной особенностью характеристики для трех групп является то, что эти минимумы не равнозначны и на рассматриваемом интервале изменения ζ явно выделяется абсолютный минимум, который соответствует истинному решению.

Впрочем, если распределение групп наблюдений таково, что величины l_j удовлетворяют резонансному соотношению типа

$$\sum_{j=2}^M m_j l_j = 0, \quad (5.48)$$

где $m_j \neq 0$ — целые числа, то может существовать множество абсолютных минимумов F . Причем плотность распределения таких минимумов вдоль ζ будет тем выше, чем ниже будет порядок резонанса или чем меньше будут величины m_j . Пример резонансного распределения групп наблюдений для $M = 3$ представлен на рис. 5.8, где использованы те же веса, что и для наблюдений Метиды. Здесь имеет место резонанс $l_2/l_3 = 2/3$, который приводит к

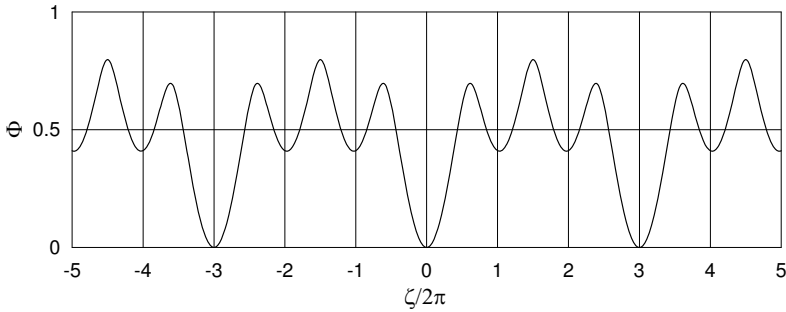


Рис. 5.8. Поведение $\Phi(\zeta)$ для трех групп наблюдений с резонансным распределением: $l_1 = 0$, $l_2 = 2/3$, $l_3 = 1$

множеству равномерно распределенных вдоль ζ абсолютных минимумов с интервалом 6π . Очевидно, вследствие повышения порядка резонанса (5.48) увеличение числа разрозненных групп M будет уменьшать плотность абсолютных минимумов.

Интересен также случай многочисленных равномерно распределенных и равновесных групп. Нетрудно показать, что при достаточно большом M

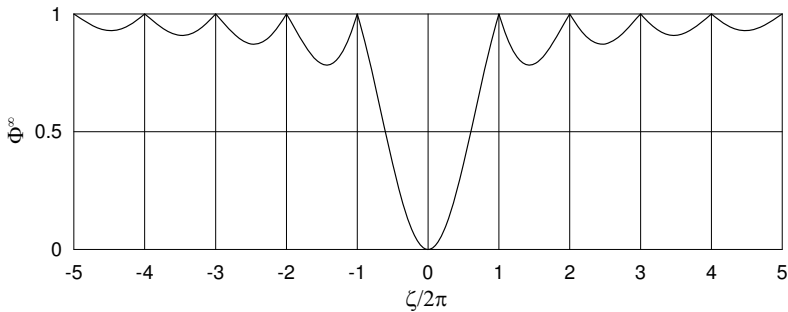
$$F \approx F^\infty \equiv F|_{M \rightarrow \infty} = 1 - \frac{1}{\zeta} [\sin((1 - l_0)\zeta - \beta) - \sin(-l_0\zeta - \beta)].$$

Тогда

$$\Phi \approx \Phi^\infty \equiv \Phi|_{M \rightarrow \infty} = 1 - \left| \frac{2}{\zeta} \sin \frac{\zeta}{2} \right|.$$

Поведение Φ^∞ представлено на рис. 5.9, откуда видно, что даже при плотном распределении групп наблюдений функция F также будет иметь множество минимумов, среди которых явно выделяется единственный абсолютный минимум.

Из рис. 5.9 можно извлечь несколько иную, также весьма полезную для практики информацию о поведении целевой функции. Допустим, для минимизации целевой функции σ^2 используется какой-либо численный (итерационный) метод типа Гаусса–Ньютона. Тогда график на рисунке фактически дает предельные

Рис. 5.9. Поведение $\Phi^\infty(\zeta)$

оценки максимально возможных значений ζ , при которых итерационный метод еще может иметь сходимость к тривиальному минимуму. Если начальное приближение орбитальных параметров таково, что $|\zeta| > 2\pi$ или, иначе говоря, ошибка в большой полуоси настолько большая, что

$$|\alpha| > \frac{2}{3} \frac{1}{R},$$

то, скорее всего, метод не будет сходящимся даже при плотном временном распределении моментов наблюдений. Очевидно, интервал сходимости (по α) к тривиальному минимуму будет только укорачиваться при прореживании наблюдательных данных (см., например, рис. 5.6–5.8). Мы видим в этом возможную причину неудачи П.В. Садбери (Sudbury, 1969) в попытке получить единую систему элементов Амальтеи по двум группам наблюдений, разведенных почти 30-летним временным интервалом. По-видимому, первоначальные оценки параметров были настолько грубыми, что они просто не попали в область сходимости метода, минимизирующего целевую функцию.

Несмотря на то, что при $|\lambda_j| \gg 0$ целевая функция σ^2 (5.37) достаточно хорошо представляется ее тригонометрической составляющей f_{\cos} (5.38), первая в отличие от последней всегда будет иметь только один абсолютный минимум в $\alpha = \beta = 0$, соответ-

ствующий истинному решению \bar{x} , где целевая функция обращается в нуль. Это связано с наличием в σ^2 квадратичного члена α^2 . В других минимумах целевая функция принимает самые разнообразные (ненулевые) значения. Поэтому в нашем случае за критерий истинности искомого решения x (по наблюдениям \bar{x}) можно принять условие наименьшего значения σ^2 на множестве всех минимумов.

Тем не менее заметим, что для околорезонансных распределений l_j могут иметь место такие минимумы, при которых значения функции σ^2 будут почти нулевыми. Причем плотность их появления тем выше, чем ниже порядок их соизмеримости. Назовем эти минимумы псевдотривиальными.

Кроме того, следует иметь в виду, что если зафиксировать a и в качестве оцениваемых параметров рассматривать n (независимо от a) и E_0 , то функция σ^2 будет подобна f_{\cos} со всеми ее особенностями, но уже относительно ν (5.34) и β .

До сих пор мы полагали, что наблюдения, представленные решением \bar{x} , являются точными, однако при наличии (случайных) ошибок $\Delta\bar{x}$ функция σ^2 , вообще говоря, не будет принимать нулевые значения и, кроме того, наиболее правдоподобное решение в окрестности $\alpha = \beta = 0$ может не доставлять абсолютный минимум: вследствие ошибок статус абсолютного может перейти к одному из псевдотривиальных минимумов. Следовательно, в этом случае условие наименьшего значения σ^2 нельзя принимать в качестве критерия максимального правдоподобия.

Весьма интересен случай наличия в аномалии E вековых ошибок $\lambda^* = n^*(t - t_0)$, которые проявляются на практике в результате немоделируемых эффектов, в частности, вследствие игнорирования либо упрощения некоторых моделей сил. Кстати, для Амальтеи оценки n^* приведены на рис. 2.1. Нетрудно показать, что влияние вековых ошибок λ^* приводит к сдвигу функции f_{\cos} (5.38) вдоль α , тогда как соответствующие значения минимумов функции сохраняются. При этом величина φ преобразуется к виду

$$\varphi = \lambda \left(\frac{3}{2}\alpha - \frac{n^*}{\bar{n}} \right) - \beta.$$

Отсюда, например, следует, что даже если упрощенная (грубая) и усложненная (более точная) модели дают одни и те же значения минимумов целевой функции σ^2 , это еще не может являться основанием в пользу выбора первой, так как определяемые при минимизации σ^2 значения α будут принципиально разными.

Наконец, заметим, что множество минимумов σ^2 имеет место только для $M > 1$. В случае одной группы достаточно близких наблюдений

$$\min_{\beta \in (-\pi, +\pi]} \sigma^2 = \alpha^2.$$

Поэтому целевая функция для $M = 1$ должна иметь единственный минимум в $\alpha = \beta = 0$.

Итак, целевая функция σ^2 (5.37) (для $M > 1$) имеет множество минимумов. Поэтому при использовании численных (итерационных) методов для минимизации σ^2 по a и E_0 (5.33) получаемое решение будет непосредственно зависеть от выбираемого начального приближения. При этом из всех минимумов не всегда узнаваем тот, что соответствует истинному решению \bar{x} . В особенности это может иметь место, когда $|\lambda_j| \gg 0$.

В действительности используемые на практике модели \mathbf{p}^C гораздо сложнее, нежели рассмотренные нами выше. Тем не менее проблема неоднозначного определения спутниковых орбит в общем случае также имеет место, и, кроме того, вследствие ошибок наблюдений она куда серьезнее.

На рис. 5.10 в плоскости большая полуось–эксцентриситет (a, e) представлено распределение минимумов целевой функции S для Адрастеи, близкого спутника Юпитера, у которого на данный момент имеются всего две группы наземных наблюдений (Nicholson, Matthews, 1991; Kulyk et al., 2002). За промежуток времени между этими группами (12 лет) спутник совершает порядка 14500 оборотов. Для представления наблюдений спутника использовалась высокоточная численная модель орбитального движения (Авдюшев, Баныщикова, 2008).

Как видно из рис. 5.10, целевая функция дает довольно малые среднеквадратические ошибки в окрестности абсолютного мини-

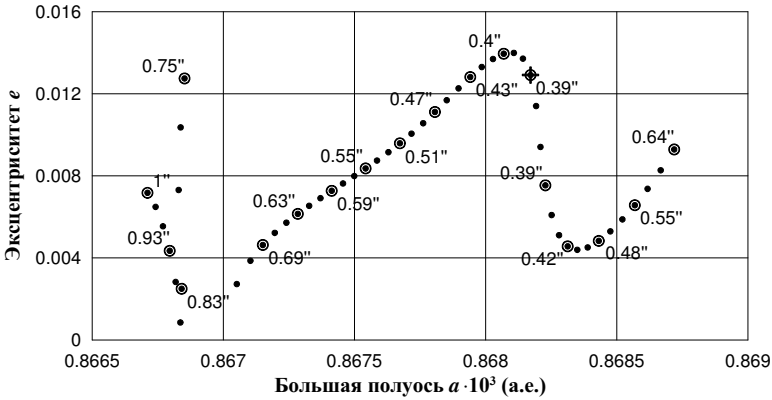


Рис. 5.10. Распределение орбитальных параметров a и e , доставляющих минимумы целевой функции S (Адрастея). Для некоторых решений, обозначенных кружочками, даны среднеквадратические ошибки. Решение в абсолютном минимуме обозначено крестиком

му: $\sigma = \sqrt{S/N}$ ($N = 180$). В то же время нельзя утверждать, что именно в абсолютном минимуме находятся наилучшие оценки орбитальных параметров, поскольку среднеквадратические ошибки в соседних минимумах отличаются на величины, гораздо меньшие ошибок наблюдений.

Опасность наличия у целевой функции множества минимумов состоит, очевидно, в том, что при формальной минимизации можно получить такие оценки, которые при достаточно хороших среднеквадратических ошибках будут давать неудовлетворительный прогноз движения спутника на длительном интервале времени.

На рис. 5.11 приведено соответствие между последними наблюдениями 2000 г. спутника Метиды (Kulyk et al., 2002) и его моделируемыми положениями, полученными при улучшении орбиты по двум предыдущим группам 1988 г. и 1999 г. (Nicholson, Matthews, 1991; Kulyk et al., 2002) для случаев двух соседних решений, соответствующих $\zeta/2\pi = \pm 1$. Рисунок показывает, что мо-

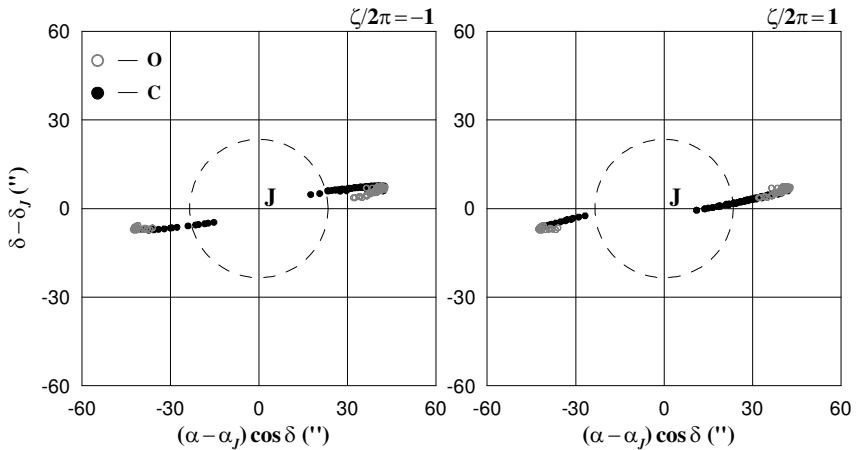


Рис. 5.11. Распределение наблюдаемых (О) и вычисленных (С) положений Метиды на небесной сфере для третьей группы наблюдений (В18 Terskol) (пунктирной линией показана граница видимого с Земли диска Юпитера)

дель, основанная на этих решениях, дает очень плохой прогноз на 2000 г., и если бы она использовалась для планирования наблюдений спутника в одной из его элонгаций, вполне возможно, что в ожидаемое время он бы не оказался в ожидаемом месте. Вместе с тем решение, соответствующее $\zeta = 0$, обеспечивает хорошее согласие с наблюдениями.

5.7.2. Наблюдения на короткой орбитальной дуге

Проблема неоднозначного определения орбит может также возникать в иных случаях, а именно когда наблюдения сгруппированы на малом временном интервале и покрывают короткую орбитальную дугу. Очевидно, что такой состав наблюдений имеют почти все новые объекты. Впрочем, возникновение этой проблемы возможно лишь при определенных условиях наблюдения, что хотя и сужает класс проблемных объектов, их тем не менее остается

еще достаточно много и значительная часть из них — это новые далекие спутники больших планет.

Для общего понимания проблемы можно обратиться, как это может показаться не удивительным, к театру теней. Как невозможно по проекции на экран определить, в какую сторону вращается фигурка-персонаж, так же по аналогии невозможно по наблюдениям в угловых координатах на небесной сфере определить, в какую сторону на круговой орбите движется объект: к наблюдателю или от него³⁹ (см., например, наблюдения Метиды на рис. 5.11). Формально трудность состоит просто в отсутствии информации о расстоянии между наблюдателем и объектом. В итоге мы имеем две альтернативные орбиты и только одна из них, неизвестно какая, соответствует реальному движению.

Спутники планет, как правило, наблюдаются в угловых координатах, и проблема неоднозначного определения орбит для новых далеких спутников фактически связана с известной проблемой, возникающей в предварительном определении орбит по трем угловым положениям, когда при определенных условиях наблюдения существуют два приемлемых решения уравнения Лагранжа восьмой степени относительно расстояния, которые дают две предварительные кеплеровские орбиты (Charlier, 1910) (см. также Приложение). Условия двойных решений выполняются в том случае, если наблюдаемый объект находится в области, указанной серым цветом на рис. 5.12, где в качестве центрального тела выступает Юпитер. Как видно из рисунка, эта область целиком содержит сферу влияния Юпитера. Поэтому для любого спутника планеты при предварительном определении его орбиты всегда будут существовать два решения, одинаково хорошо представляющие наблюдения. Если неизвестна какая-либо дополнительная информация о геоцентрическом расстоянии до наблюдаемого спутника, то оказывается совершенно невозможным выбрать ту орбиту, которая будет соответствовать реальной.

³⁹Если, конечно же, плоскость орбиты не перпендикулярна направлению от наблюдателя на центральное тело.

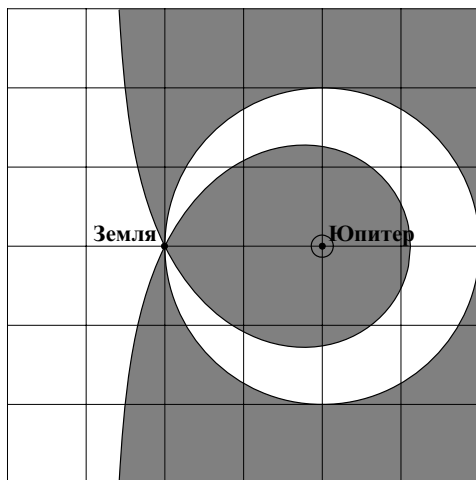


Рис. 5.12. Область двойных решений (около Юпитера показана сфера влияния планеты). Следует заметить, что эта схема получена на основе формул из работы (Charlier, 1910), где предполагается, что центральным телом является Солнце, а не Юпитер, а наблюдения выполняются за астероидом с Земли, движение которой подчинено только притяжению светила. Тем не менее структура области двойных решений для Юпитера как центрального тела должна быть подобной

Впрочем, проблема двужначности обычно разрешается на этапе уточнения орбиты с привлечением дополнительных наблюдений, которые, как правило, имеются в наличии. При решении обратной задачи с использованием избыточной выборки наблюдений разумным критерием выбора нужного решения, соответствующего реальной орбите, является малость среднеквадратической ошибки: из двух решений необходимо выбирать то, для которого среднеквадратическая ошибка принципиально меньше. Однако, если в обоих случаях среднеквадратические ошибки сравнимы, выбор решения без привлечения какой-либо еще дополнительной информации об орбите оказывается под вопросом. Такая ситуация вполне возможна для объектов, наблюдавшихся на очень ко-

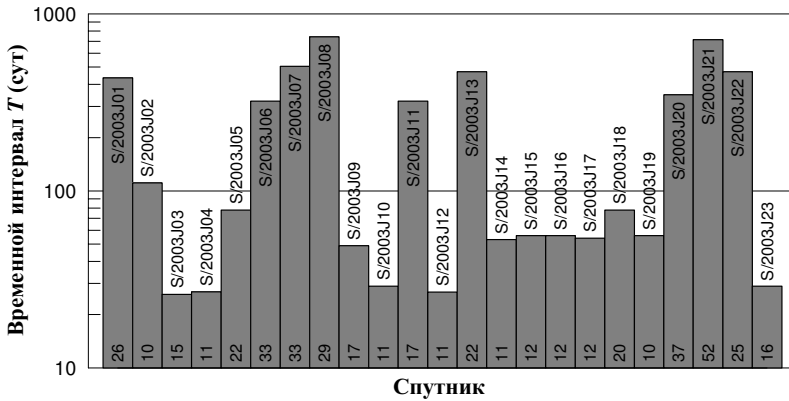


Рис. 5.13. Временные интервалы наблюдения новых далеких спутников Юпитера, открытых в 2003 г. В нижней части графика указано количество наблюдений для каждого из спутников

роткой дуге, что как раз характерно для новых неименованных спутников Юпитера семейства S/2003. Рассмотрим их на предмет двойных решений.

Уже предварительно зная периоды спутниковых наблюдений (рис. 5.13), можно с уверенностью утверждать, что для объектов, которые наблюдались менее 100 сут, должны существовать альтернативные решения, одинаково приемлемые для представления наблюдений.

Проблема двойных решений для спутников семейства S/2003 детально исследована в работе (Авдюшев, Баньщикова, 2010). Приведем ее основные результаты.

Определение спутниковых орбит выполнялось по имеющимся астрометрическим наблюдениям (Arlot, Emelyanov, 2009) на основе высокоточного численного моделирования спутникового движения (Авдюшев, Баньщикова, 2008) в рамках нелинейной задачи наименьших квадратов. В качестве начальных приближений для каждого спутника бралась пара решений, полученных из предварительного определения орбиты методом Лапласа (Charlier, 1910).

Таблица 5.2. Альтернативные оценки орбитальных элементов

S/2003	a (а.е.)	e	σ (")	a^* (а.е.)	e^*	σ^* (")
J02	$1.81 \cdot 10^{-1}$	0.34	0.14	$8.04 \cdot 10^{-2}$	0.74	0.20
J03	$1.29 \cdot 10^{-1}$	0.23	0.22	-1.89	1.05	0.30
J04	$1.23 \cdot 10^{-1}$	0.43	0.20	$1.53 \cdot 10^{-1}$	0.27	0.17
J05	$1.63 \cdot 10^{-1}$	0.21	0.25	$-1.31 \cdot 10^{-2}$	12.29	1.10
J09	$1.58 \cdot 10^{-1}$	0.21	0.32	$-1.17 \cdot 10^{-2}$	8.96	4.22
J12	$1.17 \cdot 10^{-1}$	0.49	0.20	$1.92 \cdot 10^{-1}$	0.40	0.21
J14	$1.68 \cdot 10^{-1}$	0.28	0.19	$1.33 \cdot 10^{-1}$	0.08	5.45
J15	$1.52 \cdot 10^{-1}$	0.13	0.24	$-2.35 \cdot 10^{-2}$	4.73	3.78
J16	$1.38 \cdot 10^{-1}$	0.24	0.16	$-3.44 \cdot 10^{-2}$	4.35	0.92
J17	$1.53 \cdot 10^{-1}$	0.19	0.15	$-3.97 \cdot 10^{-1}$	1.26	2.21
J19	$1.59 \cdot 10^{-1}$	0.30	0.20	$-3.64 \cdot 10^{-2}$	1.71	1.28
J23	$1.56 \cdot 10^{-1}$	0.36	0.35	$8.08 \cdot 10^{-2}$	0.42	0.85

В результате только для 12 спутников из 23 были получены пары орбит, элементы которых приведены в табл. 5.2, в то время как для каждого из оставшихся 11 спутников только одно из двух начальных приближений обеспечивало сходимость итерационного процесса. В таблице a — большая полуось; e — эксцентриситет; σ — среднеквадратическая ошибка; звездочкой обозначены альтернативные орбитальные элементы, ранее не публиковавшиеся в литературе. Интересно, что среди 24 орбит в 12 полученных парах 7 альтернативных орбит оказались гиперболическими, т.е. уже 7 объектов могут являться астероидами, временно захваченными Юпитером. Впрочем, из альтернативных решений все же можно принять лишь те, среднеквадратические ошибки которых достаточно малы. Таким образом, в действительности заслуживают внимание только 4 объекта: S/2003 J02, S/2003 J03, S/2003 J04, S/2003 J12, один из которых (S/2003 J03) имеет гиперболическую орбиту.

Легче найти ошибку, чем истину.
И. Гете

Глава 6. ОЦЕНИВАНИЕ ПАРАМЕТРИЧЕСКОЙ ТОЧНОСТИ

Наблюдения любого небесного тела, как бы тщательно они не выполнялись, обременены ошибками различного происхождения, неустранимыми в силу их случайности⁴⁰. При определении орбиты они неизбежно передаются параметрам принятой орбитальной модели. Таким образом, ошибки наблюдений задают некий барьер (модельной) точности, который не может быть преодолен практически, какой бы высокоточный вычислительный инструментарий не использовался. Встает вопрос: как же тогда оценить неопределенности в орбитальных параметрах, вызванные ошибками наблюдений?

С появлением в последнее время быстродействующих и многопроцессорных компьютеров для исследования точности орбит, определяемых из наблюдений, все чаще прибегают к стохастическому моделированию возможных значений орбитальных параметров методами Монте-Карло (Chernitsov et al., 1998; Virtanen et al., 2001; Bordovitsyna et al., 2001; Williams et al., 2005; Muinonen et al., 2006; Avdyushev, Vanschikova, 2007; Desmars et al., 2009; Avdyushev, 2009; Emel'yanov, 2010; Avdyushev, 2011), что широко используется при планировании наблюдений и идентификации небесных тел, а также в задачах астероидной опасности.

Множество возможных значений, плотность которого отвечает вероятностной плотности нахождения истинных значений орбитальных параметров, обычно строится на основе оценок линейризированной задачи наименьших квадратов (НК), несмотря

⁴⁰Например, щедрый источник ошибок для оптических наблюдений — это дрожание атмосферы, злейший враг наземного наблюдателя.

на то, что связь между представлениями наблюдений и параметрами, вообще говоря, нелинейна. Таким образом, использование линейных оценок НК-задачи для стохастического моделирования возможных значений является обоснованным только в том случае, если эти значения покрывают достаточно малую область в параметрическом пространстве, где указанная связь может хорошо представляться линейной аппроксимацией. В ином случае линейные оценки будут недостоверно описывать вероятностный разброс. Следовательно, тогда нужно прибегать к каким-либо нелинейным методам стохастического моделирования.

В данной главе дается обзор линейных и нелинейных методов, основанных на концепциях классического регрессионного анализа. Их эффективность исследуется применительно к оцениванию точности орбиты нового спутника Юпитера S/2003 J04, параметры движения которого плохо определены вследствие скудного состава имеющихся на данный момент наблюдений. Вводится показатель нелинейности НК-задачи, полезный для принятия решения в выборе между линейными и нелинейными методами.

6.1. Задача наименьших квадратов, ее геометрия и доверительное оценивание

Пусть требуется решить переопределенную систему уравнений

$$\mathbf{p}^O = \mathbf{p}^C(\mathbf{q}), \quad (6.1)$$

где \mathbf{q} — K -мерный вектор определяемых параметров; \mathbf{p}^O — N -мерный вектор наблюдений (измеряемых величин), содержащий ошибки $\delta\mathbf{p}^O$; \mathbf{p}^C — заданная N -мерная вектор-функция параметров \mathbf{q} (представление наблюдений принятой моделью). Причем предполагается, что ошибки наблюдений случайны и независимо распределены по нормальному закону с нулевой средней и дисперсией σ^2 , т.е. $\delta\mathbf{p}^O \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ (\mathbf{I} — единичная матрица). Решение переопределенной системы $\hat{\mathbf{q}}$ (наилучшая оценка) находится из принципа наименьших квадратов: $S(\hat{\mathbf{q}}) \rightarrow \min$, где $S(\mathbf{q})$ — целе-

вая функция, характеризующая близость наблюдаемых и моделируемых величин:

$$S(\mathbf{q}) = \|\mathbf{p}^O - \mathbf{p}^C(\mathbf{q})\|^2. \quad (6.2)$$

В линейном случае

$$\mathbf{p}^C(\mathbf{q}) = \mathbf{A}\mathbf{q}, \quad (6.3)$$

где \mathbf{A} — прямоугольная постоянная матрица размера $N \times K$ ранга K . Тогда оценка $\hat{\mathbf{q}}$, доставляющая минимум S (6.2), будет

$$\hat{\mathbf{q}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{p}^O. \quad (6.4)$$

Поскольку связь между $\hat{\mathbf{q}}$ и \mathbf{p}^O линейная, $\hat{\mathbf{q}} \sim \mathcal{N}(\bar{\mathbf{q}}, \mathbf{C})$, где $\bar{\mathbf{q}}$ — вектор истинных значений параметров, а \mathbf{C} — ковариационная матрица параметрической оценки $\hat{\mathbf{q}}$:

$$\mathbf{C} = \sigma^2 (\mathbf{A}^T \mathbf{A})^{-1}. \quad (6.5)$$

В нелинейном случае для нахождения оценки $\hat{\mathbf{q}}$ обычно прибегают к итерационному методу Гаусса–Ньютона (5.23):

$$\hat{\mathbf{q}} = \lim_{k \rightarrow \infty} \mathbf{q}_k,$$

где

$$\mathbf{q}_k = \mathbf{q}_{k-1} - \frac{1}{2} \mathbf{Q}_{\mathbf{q}_{k-1}}^{-1} \left(\frac{\partial S}{\partial \mathbf{q}} \right)_{\mathbf{q}_{k-1}}^T, \quad \mathbf{Q} = \left(\frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right)^T \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}}, \quad (6.6)$$

однако о распределении оценки $\hat{\mathbf{q}}$ как случайной величины здесь уже ничего определенного сказать нельзя.

Согласно геометрической интерпретации линейной НК-задачи (6.3) (рис. 6.1), вектор измеряемых величин задает точку \mathbf{p}^O в некотором N -мерном пространстве наблюдений \mathbf{p} , тогда как образом $\mathbf{A}\mathbf{q}$ параметрического пространства \mathbf{q} в пространстве наблюдений \mathbf{p} является K -мерное подпространство \mathbf{p}^C , называемое подпространством оценок (Bates, Watts, 1980; Draper, Smith, 1981).

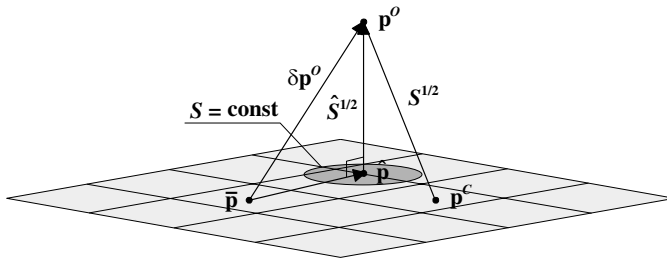


Рис. 6.1. Геометрическая интерпретация линейной НК-задачи при $N = 3$ и $K = 2$. Здесь \mathbf{p}^O — точка наблюдений; $\hat{\mathbf{p}}$ и $\bar{\mathbf{p}}$ на плоскости оценок (сетка, заполненная серым цветом) — точки, задаваемые линейной моделью $\mathbf{p}^C(\mathbf{q}) = \mathbf{A}\mathbf{q}$ с истинными значениями параметров $\bar{\mathbf{q}}$ и их НК-оценками $\hat{\mathbf{q}}$ соответственно; \mathbf{p}^C — произвольная точка плоскости оценок, и $\delta\mathbf{p}^O$ — вектор ошибок наблюдений. Доверительная область на плоскости оценок представлена темно-серым кругом

Очевидно, так как вектор $\delta\mathbf{p}^O$ представляет только измерительные ошибки, точные измерения $\bar{\mathbf{p}}$, соответствующие истинным значениям параметров $\bar{\mathbf{q}}$, должны принадлежать этому подпространству, иначе говоря, $\bar{\mathbf{p}} = \mathbf{p}^C(\bar{\mathbf{q}}) = \mathbf{A}\bar{\mathbf{q}}$. Наконец, важно заметить, что в линейном случае подпространство оценок плоское, т.е. его можно натянуть на некоторую прямоугольную систему координат в пространстве наблюдений.

Таким образом, после введения геометрических объектов \mathbf{p}^O и \mathbf{p}^C целевая функция S (6.2) становится не чем иным, как квадратом расстояния в пространстве наблюдений \mathbf{p} от точки \mathbf{p}^O до произвольной точки подпространства \mathbf{p}^C , соответствующей некоторому прообразу параметрического пространства \mathbf{q} , и только точка $\hat{\mathbf{p}} = \mathbf{p}^C(\hat{\mathbf{q}})$ обеспечивает минимум целевой функции $\hat{S} = S(\hat{\mathbf{q}})$.

Между тем, поскольку \hat{S} фактически является квадратом расстояния от точки \mathbf{p}^O до подпространства оценок, точку $\hat{\mathbf{p}}$ можно рассматривать как ортогональную проекцию \mathbf{p}^O на \mathbf{p}^C . Введем случайную J -мерную стандартную нормально распределен-

ную величину ν_J . Тогда ввиду того, что вектор $\mathbf{p}^O - \bar{\mathbf{p}} = \delta\mathbf{p}^O = \sigma\nu_N$, вектор $\hat{\mathbf{p}} - \bar{\mathbf{p}} = \sigma\nu_K$, а квадрат его величины, поделенный на дисперсию ошибок, будет распределен по закону χ_K^2 со степенью свободы K :

$$\frac{\|\hat{\mathbf{p}} - \bar{\mathbf{p}}\|^2}{\sigma^2} = \chi_K^2. \quad (6.7)$$

Если же рассматривать точку $\bar{\mathbf{p}}$ относительно случайно распределенной точки $\hat{\mathbf{p}}$, то ее случайное распределение будет таким же, как $\hat{\mathbf{p}} - \bar{\mathbf{p}}^{41}$, поэтому доверительную область можно определить неравенством

$$\frac{\|\hat{\mathbf{p}} - \mathbf{p}^C\|^2}{\sigma^2} \leq \chi_{K,\alpha}^2, \quad (6.8)$$

где $\chi_{K,\alpha}^2$ — α -квантиль функции χ_K^2 . Вероятностный смысл неравенства (6.8) состоит в том, что оно задает в K -мерном подпространстве оценок шарообразную область с центром $\hat{\mathbf{p}}$, которая с вероятностью α содержит точку $\bar{\mathbf{p}}$.

Если дисперсия ошибок σ^2 неизвестна, то формула (6.8) неприемима. Чтобы определить доверительную область в этом случае, воспользуемся несмещенной оценкой дисперсии

$$\hat{\sigma}^2 = \frac{S(\hat{\mathbf{q}})}{N - K},$$

которая как статистика имеет связь с распределением χ_{N-K}^2 :

$$\hat{\sigma}^2 = \sigma^2 \frac{\chi_{N-K}^2}{N - K}, \quad (6.9)$$

поскольку $S(\hat{\mathbf{q}}) = \|\mathbf{p}^O - \hat{\mathbf{p}}\|^2$, а вектор $\mathbf{p}^O - \hat{\mathbf{p}} = \sigma\nu_{N-K}$ распределен по нормальному закону в $(N - K)$ -мерном подпространстве, ортогональном к подпространству оценок.

Умножим (6.7) слева и справа на $\sigma^2/(K\hat{\sigma}^2)$. Тогда будем иметь

$$\frac{\|\hat{\mathbf{p}} - \bar{\mathbf{p}}\|^2}{K\hat{\sigma}^2} = \frac{\chi_K^2}{K} \frac{N - K}{\chi_{N-K}^2} = \mathcal{F}_{K,N-K}, \quad (6.10)$$

⁴¹Точнее, $\bar{\mathbf{p}} - \hat{\mathbf{p}}$, но это в данном случае не принципиально в силу симметричности нормального распределения относительно среднего.

где $\mathcal{F}_{K,N-K}$ — распределение Фишера со степенями свободы K и $N-K$. Отсюда доверительная область определяется неравенством

$$\frac{\|\hat{\mathbf{p}} - \mathbf{p}^C\|^2}{K\hat{\sigma}^2} \leq \mathcal{F}_{K,N-K,\alpha}, \quad (6.11)$$

где $\mathcal{F}_{K,N-K,\alpha}$ — α -квантиль функции вероятности Фишера.

Так же, как и (6.8), неравенство (6.11) задает в K -мерном подпространстве оценок шарообразную область с центром $\hat{\mathbf{p}}$, которая с вероятностью α содержит точку $\bar{\mathbf{p}}$, однако при одних и тех же α радиус данной области больше и тем значительнее, чем меньше число измерений N .

Выполнив в (6.11) замену $\hat{\mathbf{p}} = \mathbf{A}\hat{\mathbf{q}}$ и $\mathbf{p}^C = \mathbf{A}\mathbf{q}$, после несложных выкладок получим классическое определение доверительной области в параметрическом пространстве:

$$(\hat{\mathbf{q}} - \mathbf{q})^T \mathbf{A}^T \mathbf{A} (\hat{\mathbf{q}} - \mathbf{q}) \leq K\hat{\sigma}^2 \mathcal{F}_{K,N-K,\alpha} = (\kappa\hat{\sigma})^2$$

или

$$(\hat{\mathbf{q}} - \mathbf{q})^T \hat{\mathbf{C}}^{-1} (\hat{\mathbf{q}} - \mathbf{q}) \leq \kappa^2, \quad (6.12)$$

где $\kappa = \sqrt{K\mathcal{F}_{K,N-K,\alpha}}$; $\hat{\mathbf{C}} = \hat{\sigma}^2(\mathbf{A}^T \mathbf{A})^{-1}$ — оценка ковариационной матрицы \mathbf{C} (6.5). Согласно (6.12), в пространстве \mathbf{q} доверительная область — это эллипсоидальное множество с центром $\hat{\mathbf{q}}$, содержащее с вероятностью α истинные значения параметров $\bar{\mathbf{q}}$. Как видно, форма и ориентация эллипсоида (6.12) в параметрическом пространстве задается только оценкой $\hat{\mathbf{C}}$: доверительная область ориентирована вдоль собственных векторов матрицы $\hat{\mathbf{C}}$, а ее собственные числа задают отношения полуосей эллипсоида. Между тем следует заметить, что поскольку оценка дисперсии $\hat{\sigma}^2$ в ковариационной матрице является случайной, размер и объем доверительной области (6.12) также будут случайными.

В общем нелинейном случае определения (6.8), (6.11) и (6.12) уже не справедливы и доверительные области могут быть существенно нешарообразными в оценочном подпространстве и неэллиптическими в параметрическом пространстве, хотя бы потому,

что подпространство оценок уже не является плоским в пространстве наблюдений. Впрочем, интересен частный нелинейный случай, когда подпространство оценок все же плоское, а нелинейность вызвана неудачным выбором параметров \mathbf{q} , т.е. на самом деле существует такой набор K параметров \mathbf{r} , при переходе к которым НК-задача становится линейной⁴². Для таких нелинейных задач, которые называют еще внутренне линейными (или внешне нелинейными) (Draper, Smith, 1981), доверительные области в подпространстве оценок будут определяться, как и в линейном случае, неравенствами (6.8) и (6.11), но тогда соответствующие области в параметрическом пространстве должны строиться путем их обратного (нелинейного) отображения посредством модели $\mathbf{p}^C(\mathbf{q})$.

6.2. Моделирование доверительной области

При компьютерном моделировании в линейном случае доверительную область (6.12) в параметрическом пространстве можно представить как конечное множество равномерно и случайно заполняющих ее точек \mathbf{q} , для которых справедливо неравенство (6.12). Однако, чтобы судить о форме и размерах доверительной области, достаточно иметь представление лишь о ее границе. Для моделирования границы можно воспользоваться простой схемой

$$\mathbf{q} = \hat{\mathbf{q}} + \hat{\mathbf{C}}^{1/2} \boldsymbol{\rho}, \quad (6.13)$$

где $\hat{\mathbf{C}}^{1/2}$ — такая матрица размера $K \times K$ (например, матрица Холецкого), перемножение которой на транспонированную дает ковариационную матрицу параметров $\hat{\mathbf{C}}$:

$$\hat{\mathbf{C}}^{1/2} (\hat{\mathbf{C}}^{1/2})^T = \hat{\mathbf{C}}; \quad (6.14)$$

а $\boldsymbol{\rho}$ — K -мерный случайный вектор, равномерно описывающий границу многомерной сферы радиуса κ . Действительно, согласно

⁴²В качестве параметров \mathbf{r} могут выступать координаты прямоугольной системы, на которую натянуто подпространство \mathbf{p}^C .

(6.13) и (6.14), случайный вектор \mathbf{q} будет представлять границу эллипсоидальной области (6.12), ведь $\boldsymbol{\rho} = (\hat{\mathbf{C}}^{1/2})^{-1}(\mathbf{q} - \hat{\mathbf{q}})$, тогда

$$\kappa^2 = \boldsymbol{\rho}^T \boldsymbol{\rho} = (\hat{\mathbf{q}} - \mathbf{q})^T \hat{\mathbf{C}}^{-1}(\hat{\mathbf{q}} - \mathbf{q}).$$

Подобного типа алгоритм, основанный на проецировании нормального распределения в параметрическом пространстве на поверхность доверительного эллипсоида, предложен в работе (Сюсина и др., 2007).

В нелинейном, но плоском случае границу доверительной области следует моделировать в три этапа. Сначала в K -мерном подпространстве оценок моделируется множество нормально распределенных точек с дисперсией $\hat{\sigma}^2$, например, по формуле

$$\mathbf{p}^C = \hat{\mathbf{p}} + \mathbf{\Pi} \delta \mathbf{p}, \quad \mathbf{\Pi} = \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \mathbf{Q}^{-1} \left(\frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right)^T, \quad (6.15)$$

где $\delta \mathbf{p} = \hat{\sigma} \boldsymbol{\nu}_N$, а идемпотентная матрица $\mathbf{\Pi}$ выступает в качестве ортогонального проектора на подпространство оценок. Затем точки нормального разброса \mathbf{p}^C радиально от центра $\hat{\mathbf{p}}$ проецируются на сферу радиуса $\kappa \hat{\sigma}$. Таким способом получаем границу доверительной области (6.11) в подпространстве оценок. Кстати, подобным образом можно получить и равномерное распределение $\boldsymbol{\rho}$ на многомерной сфере в схеме (6.13), а именно $\boldsymbol{\rho} = \kappa \boldsymbol{\nu}_K / \|\boldsymbol{\nu}_K\|$. Наконец, остается выполнить обратное отображение \mathbf{p}^C в пространство параметров \mathbf{q} с использованием модели $\mathbf{p}^C(\mathbf{q})$, что осуществимо путем многократного численного решения нелинейной НК-задачи методом Гаусса–Ньютона (6.6) для каждой из полученных выборок измерений на сферической границе доверительной области \mathbf{p}^C , где за неимением лучшего в качестве начального приближения выбирается оценка $\hat{\mathbf{q}}$.

Схему (6.15) можно применять также для приближенного построения доверительной области в общем нелинейном случае, но тогда получаемые точки \mathbf{p}^C первого и второго этапов будут принадлежать уже не подпространству оценок (рис. 6.2), а некоторому касательному к нему в точке $\hat{\mathbf{p}}$ плоскому подпространству, и

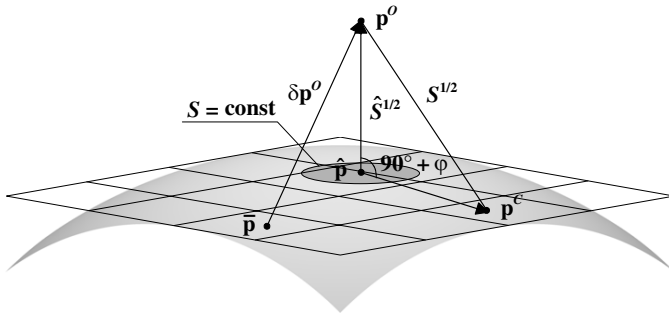


Рис. 6.2. То же самое, что и на рис. 6.1, но для нелинейной НК-задачи. Поверхность оценок и касательная плоскость к ней представлены серым цветом и сеткой соответственно

в таком случае, конечно же, чем меньше кривизна подпространства оценок, тем достовернее моделируемая область будет представлять достоверную.

Наконец, хотелось бы особо отметить, что нужно быть осторожным при отображении достоверной области на пространство размерности L , меньшей K , поскольку отображенная область в новом пространстве, вообще говоря, уже не будет достоверной для заданной вероятности α . Если все же предполагается линейное отображение, то, по крайней мере, для внутренне линейных задач область первоначально следует моделировать с меньшим коэффициентом, а именно $\kappa = \sqrt{L\mathcal{F}_{L,N-L,\alpha}}$. Это подобно ситуации, когда нормальный разброс трехмерной случайной величины ν_3 ($K = 3$) проецируется (отображается) на плоскость. В результате получается разброс уже двумерной случайной величины ν_2 ($L = 2$). В соответствии с вероятностной плотностью радиусы круга для ν_2 и сферы для ν_3 , центрированных в начале координат и содержащих одинаковую вероятностную массу (аналог вероятности α), должны быть разными: первый меньше последнего. Следовательно, круг не является проекцией сферы. Эти же особенности имеют место и при проецировании достоверных областей.

6.3. Моделирование возможных значений параметров

Задача доверительного оценивания состоит в определении области, которая с заданной вероятностью содержит точку, представляющую истинные значения параметров. Решение этой задачи неоднозначно, поскольку построение любой доверительной области основано на некоей вероятностной плотности и, таким образом, каждой вероятности α будет соответствовать в действительности бесконечное множество доверительных областей, содержащих одинаковую вероятностную массу. Что касается определений (6.8), (6.11) и (6.12), то каждое из них задает единственную область фактически лишь вследствие дополнительного налагаемого на нее условия симметричности относительно оценки. В связи с этим неоднозначность общего определения вероятностной области порой может поставить под сомнение практическую полезность доверительного оценивания.

Впрочем, на практике очень часто возникает обратная проблема: необходимо, наоборот, уже в выделенном, интересующем нас объеме параметрического (или другого) пространства оценить вероятность нахождения точки, отвечающей истинным значениям параметров. Эта проблема решается стохастическим моделированием так называемых возможных значений параметров, которые дискретно представляют вероятностную плотность распределения истинных значений $\bar{\mathbf{q}}$ относительно оценочных $\hat{\mathbf{q}}$. Искомая вероятность оценивается как отношение числа попавших в объем точек, определяемых возможными значениями, к общему числу. Если объем-мишень задана в каком-либо другом (непараметрическом) пространстве, то, очевидно, следует выполнить соответствующее отображение выборочного множества точек на это пространство.

6.3.1. Линейные методы

Возможные значения параметров моделируются на основе тех же вероятностных распределений, что и доверительные области. В ли-

нейном случае при известной дисперсии ошибок σ^2 точка истинных измерений \mathbf{p} относительно точечных оценок $\hat{\mathbf{p}}$ распределяется по K -мерному нормальному закону с дисперсией σ^2 : $\mathbf{p} - \hat{\mathbf{p}} = \sigma \boldsymbol{\nu}_K$. Следовательно, множество возможных значений в подпространстве оценок должно моделироваться в соответствии с этим законом, для чего можно воспользоваться схемой (6.15) применительно к линейному случаю

$$\mathbf{p}^C = \hat{\mathbf{p}} + \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \delta \mathbf{p} = \hat{\mathbf{p}} + \mathbf{P} \delta \mathbf{p}, \quad (6.16)$$

что, в свою очередь, эквивалентно многократному решению НК-задачи:

$$\mathbf{p}^C : \|\hat{\mathbf{p}} + \delta \mathbf{p} - \mathbf{p}^C\|^2 \rightarrow \min \quad \text{или} \quad \mathbf{q} : \|\hat{\mathbf{p}} + \delta \mathbf{p} - \mathbf{p}^C(\mathbf{q})\|^2 \rightarrow \min. \quad (6.17)$$

Тогда схема стохастического моделирования для \mathbf{q} , аналогичная (6.16), будет

$$\mathbf{q} = \hat{\mathbf{q}} + (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \delta \mathbf{p}. \quad (6.18)$$

Нетрудно показать, что $\mathbf{q} \sim \mathcal{N}(\hat{\mathbf{q}}, \mathbf{C})$.

Однако имеется более простая схема для моделирования нормального распределения \mathbf{q} , а именно с использованием матрично-множителя $\mathbf{C}^{1/2}$ разложения Холецкого. Она имеет вид (6.13):

$$\mathbf{q} = \hat{\mathbf{q}} + \mathbf{C}^{1/2} \boldsymbol{\nu}_K. \quad (6.19)$$

Если при неизвестной дисперсии ошибок используется ее оценка $\hat{\sigma}^2$, то в соответствии с распределением (6.10) возможные \mathbf{p}^C должны быть распределены по закону Хотеллинга (Anderson, 1958) так, чтобы случайная величина $\|\hat{\mathbf{p}} - \mathbf{p}^C\|^2 / (K \hat{\sigma}^2)$ на множестве \mathbf{p}^C была распределена по закону Фишера со степенями свободы K и $N - K$. Для этого достаточно в $\delta \mathbf{p}$ в схемах (6.17) и (6.18) ввести модулирующий множитель, случайную величину $r = \sqrt{(N - K) / \chi_{N-K}^2}$: $\delta \mathbf{p} = r \hat{\sigma} \boldsymbol{\nu}_N$, вследствие чего упрощенный аналог (6.19) переписется как (Avdyushev, 2009)

$$\mathbf{q} = \hat{\mathbf{q}} + r \hat{\mathbf{C}}^{1/2} \boldsymbol{\nu}_K. \quad (6.20)$$

Согласно (6.18),

$$\begin{aligned}\|\hat{\mathbf{p}} - \mathbf{p}^C\|^2 &= (\hat{\mathbf{q}} - \mathbf{q})^T \mathbf{A}^T \mathbf{A} (\hat{\mathbf{q}} - \mathbf{q}) = \\ &= \delta \mathbf{p}^T \mathbf{\Pi}^T \mathbf{\Pi} \delta \mathbf{p} = r^2 \hat{\sigma}^2 \chi_K^2 = K \hat{\sigma}^2 \mathcal{F}_{K, N-K}.\end{aligned}\quad (6.21)$$

Следовательно, как мы видим, после введения множителя r схема (6.18) моделирует случайные величины \mathbf{q} , отвечающие распределению (6.10).

Очевидно, что $r \rightarrow 1$ и $\hat{\sigma}^2 \rightarrow \sigma^2$ при $N \rightarrow \infty$, поэтому при достаточно большом количестве наблюдений вместо усложненной схемы (6.20) вполне может быть использована схема (6.19) с $\hat{\sigma}^2$ вместо σ^2 .

6.3.2. Нелинейные методы

В нелинейных задачах, если ошибки $\delta \mathbf{p}^O$ достаточно малы и влекут малые параметрические ошибки в \mathbf{q} , НК-оценки могут рассматриваться в контексте линейной задачи⁴³, и поэтому для моделирования возможных значений параметров целесообразно прибегать либо к схеме (6.20), либо при больших N к (6.18) и (6.19), где, впрочем, предполагается, что

$$\mathbf{C} = \hat{\sigma}^2 \mathbf{Q}^{-1} \quad \text{и} \quad \mathbf{A} = \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}}.$$

При достаточно большом разбросе параметров \mathbf{q} применение явных схем оказывается необоснованным, однако в нашем распоряжении имеются также неявные аналоги (6.17), которые вполне могут быть использованы для моделирования вероятностного разброса возможных значений параметров во многих нелинейных задачах. Так как использование схем фактически предполагает случайные вариации величин $\hat{\mathbf{p}}$, будем называть их схемами возмущенных оценок.

⁴³Ниже водится так называемый показатель нелинейности, который позволяет судить о том, насколько оправдано применения линейных оценок.

В частности, совершенно очевидно, что применение неявных схем вполне обоснованно во всех внешне нелинейных задачах, поскольку они имеют геометрию линейных задач ввиду того, что их подпространства оценок являются плоскими. Интересно заметить, что тогда в схемах (6.17) замена оценки $\hat{\mathbf{p}}$ на \mathbf{p}^O абсолютно не влияет на результат моделирования, т.е. случайные вариации $\delta\mathbf{p}$ можно вносить не в представления наблюдений моделью, а в сами наблюдения. Действительно, ведь минимизация целевой функции S по \mathbf{p}^C дает разброс решений, являющийся ортогональной проекцией разброса искусственных наблюдений $\hat{\mathbf{p}} + \delta\mathbf{p}$ на подпространство оценок, но в то же время результат проецирования инвариантен относительно любого переноса наблюдений вдоль вектора $\mathbf{p}^O - \hat{\mathbf{p}}$, ортогонального к плоскому подпространству оценок. Таким образом, множество возможных значений параметров можно моделировать еще до того, как известно решение НК-задачи $\hat{\mathbf{q}}$.

Если подпространство оценок не плоское, то замена $\hat{\mathbf{p}}$ на \mathbf{p}^O может привести к существенно иным результатам, в особенности при достаточно больших $\hat{\sigma}$ (Avdyushev, 2009). Разумеется, выбор должен быть за $\hat{\mathbf{p}}$. Например, в случае, когда кривизна подпространства постоянна (как у многомерной сферы), а дисперсия ошибок σ^2 известна, именно вариация оценки $\hat{\mathbf{p}}$ в (6.17) будет давать случайный (шарообразный) разброс \mathbf{p}^C относительно $\hat{\mathbf{p}}$, точно соответствующий распределению $\hat{\mathbf{p}}$ относительно $\bar{\mathbf{p}}$. Замена же $\hat{\mathbf{p}}$ на \mathbf{p}^O приведет либо к меньшему, либо к большему разбросу в зависимости от расположения \mathbf{p}^O относительно центра кривизны подпространства оценок.

Между тем, даже если кривизна подпространства оценок постоянная, но в то же время дисперсия ошибок σ^2 неизвестна и необходимо прибегать к схемам (6.17), возникает проблема с выбором случайной величины r . Модулирующий множитель r при возмущениях $\delta\mathbf{p}$ по сути необходим для учета того факта, что величина $\hat{\sigma}^2$ как статистическая оценка дисперсии ошибок наблюдений σ^2 , получаемая из выборки наблюдательных данных со случайными ошибками, сама является случайной. Если ошибки

наблюдений распределены по нормальному закону, то в линейном случае она распределена в соответствии с формулой (6.9). В нелинейном (неплоском) случае распределение $\hat{\sigma}^2$, вообще говоря, неизвестно.

Чтобы как-то компенсировать отсутствие информации о распределении $\hat{\sigma}^2$, предлагается моделировать возможные значения параметров в два этапа (Avdyushev, 2009). На первом этапе используется схема (6.17) с некоторой сгенерированной выборкой случайных возмущений $\delta \mathbf{p} = \hat{\sigma} \boldsymbol{\nu}_N$, где в результате решения обратной задачи получаем оценку $\hat{\sigma}_*^2$ дисперсии $\hat{\sigma}^2$. Затем на втором этапе выполняется процесс моделирования по той же схеме, но уже с $\delta \mathbf{p} = r \hat{\sigma} \boldsymbol{\nu}_N$ и множителем $r = \hat{\sigma} / \hat{\sigma}_*$. В таком подходе предполагается, что случайная величина $\hat{\sigma}_*^2 / \hat{\sigma}^2$ распределена так же, как и $\hat{\sigma}^2 / \sigma^2$. Очевидно, что в плоском случае двухэтапный подход будет давать те же результаты, что и схема (6.17) с множителем r , вычисляемым формульно. Разумеется, этот подход не решает проблему полностью, ведь успех предприятия зависит главным образом от того, насколько близкой окажется оценка $\hat{\sigma}^2$ к дисперсии σ^2 . Поэтому его следует рассматривать только как некую полумеру.

Очевидный и главный недостаток неявных схем состоит в том, что минимизация в (6.17) выполняется итерационно, как правило, методом Гаусса–Ньютона (6.6), для чего порой могут потребоваться значительные затраты компьютерного времени, в особенности при моделировании большого разброса возможных значений параметров. Объемные вычисления вызваны медленной сходимостью итерационного процесса. Это связано главным образом с тем, что начальное приближение \mathbf{q}_0 , в качестве которого выбирается оценка $\hat{\mathbf{q}}$, становится очень грубым для периферийных, отдаленных от него решений. Более того, метод Гаусса–Ньютона при таком начальном приближении вообще может не сходиться. В принципе, последняя проблема практически разрешается путем простого демпфирования поправок метода (Ortega, Reinboldt, 2000), но, с другой стороны, такой подход существенно усугубляет проблему низкого быстродействия.

6.4. Взвешенная задача наименьших квадратов

До сих пор мы полагали, что все наблюдения \mathbf{p}^O равноточные. Если же известна ковариационная матрица ошибок $\mathbf{D} = \sigma^2 \mathbf{W}^{-1}$ с неединичной \mathbf{W} , то $\delta \mathbf{p}^O = \mathbf{D}^{1/2} \boldsymbol{\nu}_N$, и поэтому разброс ошибок при всевозможных измерениях будет уже не шарообразный, а эллипсоидальный, как, собственно, и разброс оценок $\hat{\mathbf{p}}$, по крайней мере, в плоском случае. Тогда, естественно, в схемах (6.17) и (6.18) следует вносить вариацию $\delta \mathbf{p} = \mathbf{D}^{1/2} \boldsymbol{\nu}_N$.

Впрочем, когда ковариационная матрица ошибок (или хотя бы матрица \mathbf{W}) известна, имеет смысл выполнять стохастическое моделирование в рамках взвешенной НК-задачи (5.3) с весовой матрицей \mathbf{W} (Черницов и др., 2007):

$$S(\mathbf{q}) = \|(\mathbf{W}^{1/2})^T(\mathbf{p}^O - \mathbf{p}^C(\mathbf{q}))\|^2 \rightarrow \min. \quad (6.22)$$

Норма в (6.22) на самом деле соответствует евклидовой норме пространства $(\mathbf{W}^{1/2})^T \mathbf{p}$. Разброс ошибок наблюдений в этом пространстве оказывается распределенным по нормальному закону с дисперсией σ^2 , т.е. $(\mathbf{W}^{1/2})^T \delta \mathbf{p}^O = \sigma \boldsymbol{\nu}_N$, что нетрудно увидеть, принимая во внимание соотношение $(\mathbf{W}^{1/2})^T = \sigma(\mathbf{D}^{1/2})^{-1}$, тогда как $\delta \mathbf{p}^O = \mathbf{D}^{1/2} \boldsymbol{\nu}_N$. Поэтому моделирование возможных значений параметров сводится к уже полученным схемам, но применительно к преобразованному пространству наблюдений $(\mathbf{W}^{1/2})^T \mathbf{p}$. В частности, схему возмущенных оценок (6.17) можно записать в виде

$$\mathbf{q} : \|(\mathbf{W}^{1/2})^T(\hat{\mathbf{p}} + \delta \mathbf{p} - \mathbf{p}^C(\mathbf{q}))\|^2 \rightarrow \min, \quad (6.23)$$

хотя здесь уже $\delta \mathbf{p} = r \hat{\sigma} [(\mathbf{W}^{1/2})^T]^{-1} \boldsymbol{\nu}_N$, а $\hat{\sigma}^2 = S(\hat{\mathbf{q}})/(N - K)$, как и выше, но с целевой функцией (6.22).

Понятно, что оценка $\hat{\mathbf{p}}$ взвешенной задачи (6.22) будет уже, вообще говоря, косоугольной, а не ортогональной проекцией точки наблюдений \mathbf{p}^O на подпространство оценок \mathbf{p}^C , хотя ортогональное проецирование в то же время имеет место в преобразованном пространстве $(\mathbf{W}^{1/2})^T \mathbf{p}$. Это существенно может уменьшить разброс оценок $\hat{\mathbf{p}}$, что, конечно же, приводит к уточнению статисти-

ки $\hat{\mathbf{q}}$. Поэтому при использовании весов следует также ожидать уменьшения разброса возможных значений параметров.

Повышение точности взвешенных оценок достаточно просто объяснить, если обратиться к геометрической интерпретации НК-задачи. Рассмотрим эллипсоид наблюдений \mathbf{p}^O , задаваемый ковариационной матрицей \mathbf{D} . Его можно представить в виде

$$(\mathbf{p}^O - \bar{\mathbf{p}})^T \mathbf{W} (\mathbf{p}^O - \bar{\mathbf{p}}) \leq \omega = \text{const}. \quad (6.24)$$

В пространстве $(\mathbf{W}^{1/2})^T \mathbf{p}$ он преобразуется в N -мерный шар:

$$[(\mathbf{W}^{1/2})^T \mathbf{p}^O - (\mathbf{W}^{1/2})^T \bar{\mathbf{p}}]^T [(\mathbf{W}^{1/2})^T \mathbf{p}^O - (\mathbf{W}^{1/2})^T \bar{\mathbf{p}}] \leq \omega. \quad (6.25)$$

Совершенно очевидно, что в плоском случае все оценки $(\mathbf{W}^{1/2})^T \hat{\mathbf{p}}$ как ортогональные проекции наблюдений $(\mathbf{W}^{1/2})^T \mathbf{p}^O$ будут заполнять K -мерный шар подпространства оценок $(\mathbf{W}^{1/2})^T \mathbf{p}^C$, который, с другой стороны, можно рассматривать как пересечение шара наблюдений (6.25) с подпространством. Следовательно, в исходном пространстве наблюдений \mathbf{p} оценки $\hat{\mathbf{p}}$ взвешенной НК-задачи (6.22) должны находиться внутри пересечения эллипсоида наблюдений (6.24) с подпространством оценок \mathbf{p}^C . Оценки же НК-задачи (6.1) без взвешивания как ортогональные проекции наблюдений \mathbf{p}^O на \mathbf{p}^C будут, вообще говоря, покрывать более обширную область подпространства оценок, включающую в себя и область оценок взвешенной задачи, в чем, собственно, и состоит низкая эффективность НК-оценок при игнорировании весов. Впрочем, следует заметить, что с теоретической точки зрения использование весов не всегда повышает точность НК-оценок. Например, когда эллипсоид (6.24) ориентирован так, что какие-либо K из его осей лежат в подпространстве оценок, то его пересечение с подпространством и ортогональная проекция будут совпадать, поэтому введение весов не даст уменьшение разброса оценок.

Таким образом, в качестве показателя эффективности взвешивания можно рассматривать отношение объемов этих эллипсоидов либо в подпространстве оценок, либо в параметрическом пространстве, где эллипсоиды описываются соответствующими ковариационными матрицами.

Если задача (6.22) линейна, т.е. $\mathbf{p}^C(\mathbf{q}) = \mathbf{A}\mathbf{q}$, то решение представимо в виде

$$\hat{\mathbf{q}} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{p}^O. \quad (6.26)$$

В соответствии с известным распределением ошибок наблюдений $\delta \mathbf{p}^O = \mathbf{D}^{1/2} \boldsymbol{\nu}_N$ будем иметь ковариационную матрицу параметрических оценок

$$\mathbf{C}_W = \sigma^2 (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1}. \quad (6.27)$$

В то же время, если при том же разбросе ошибок $\delta \mathbf{p}^O$ оценки находятся в рамках невзвешенной НК-задачи (6.1), то разброс оценок, согласно (6.4), должен описываться ковариационной матрицей

$$\mathbf{C}_I = \sigma^2 (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}^{-1} \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1}. \quad (6.28)$$

Тогда отношение объемов эллипсоидов ошибок w как показатель эффективности взвешивания вычисляется через определители ковариационных матриц (6.27) и (6.28) по формуле (Bard, 1974; Draper, Smith, 1981)

$$w = \sqrt{\frac{\det \mathbf{C}_W}{\det \mathbf{C}_I}} \leq 1.$$

Чем меньше w , тем выше эффективность взвешивания. Разумеется, эту формулу можно использовать и для нелинейных задач, где, однако, требуется положить $\mathbf{A} = \partial \mathbf{p}^C / \partial \mathbf{q}$.

6.5. Бутстрэп-метод

Альтернативным методом стохастического моделирования возможных значений параметров является бутстрэп-метод, предложенный Б. Эфроном в конце семидесятых годов прошлого века (Efron, 1979). Насколько нам известно, метод не получил широкое применение на практике в небесной механике и со времени его появления удостоивался внимания лишь немногих авторов (Шапорев, 1996; Desmars et al., 2009).

Бутстрэп-метод также основан на многократном решении обратной задачи при использовании искусственных выборок наблюдений объема N (того же, что и \mathbf{p}^O), но составленных из реальных наблюдательных данных путем так называемого выбора с возвращением. Таким образом, в выборках одни наблюдения могут встречаться несколько раз, а другие — ни разу. Следовательно, бутстрэп-метод может быть реализован посредством решения взвешенной НК-задачи, где в качестве весовой выступает случайная диагональная матрица \mathbf{W} с целочисленными неотрицательными диагональными элементами, каждый из которых указывает на то, сколько раз в выборке используется соответствующее реальное измерение. Причем полагается, что появление каждого измерения в выборке равновероятно. Тогда формально схема бутстрэп-метода представима в виде (6.22), где взвешенная НК-задача решается многократно при использовании различных случайных весовых матриц \mathbf{W} .

Из N наблюдений можно получить N^N различных выборок, однако соответствующих им уникальных орбит будет гораздо меньше. Пользуясь элементарными формулами комбинаторики и учитывая, что реальных наблюдений, составляющих выборку, должно быть не меньше K , получаем число таких орбит

$$C_N^{2N-1} - \sum_{i=1}^{K-1} C_i^N C_{i-1}^{N-1}. \quad (6.29)$$

Например, при $N = 22$ и $K = 6$ в распоряжении имеются более 10^{12} орбит, хотя при $N = 8$ — всего 645, что весьма мало для детального исследования распределения вероятностной плотности.

Примечательной особенностью бутстрэп-метода является то, что для него не требуется знание о действительном распределении ошибок наблюдений. Несмотря на это, эффективность метода еще слабо изучена. Доказано только, что для оценки математического ожидания бутстрэп-методом справедливы аналоги закона больших чисел и центральной предельной теоремы (Efron, 1979). В то же время что касается малых выборок (интересующий нас случай), то ничего определенного сказать нельзя.

6.6. Показатели нелинейности

Если нелинейность НК-задачи слабая, то для стохастического моделирования возможных значений параметров вполне может быть использована очень простая явная схема (6.20), иначе нет. Встает вопрос: возможно ли вообще оценить, насколько слаба нелинейность, с тем чтобы использовать линейные схемы вместо сложных нелинейных? Для разрешения этого вопроса введем показатель нелинейности.

Поскольку $\|\hat{\mathbf{p}} - \mathbf{p}^C\|^2 = S(\mathbf{q}) - S(\hat{\mathbf{q}})$, то для точек \mathbf{q} границы доверительной области (6.11) имеем

$$\frac{S(\mathbf{q}) - S(\hat{\mathbf{q}})}{\hat{\sigma}^2} = \kappa^2. \quad (6.30)$$

При вариации параметра κ на величину $\delta\kappa$ размеры эллипсоидальной области (6.30) изменятся и целевая функция на ее границе примет новое значение $\tilde{S} = \bar{S} + \delta S$, где \bar{S} — значение целевой функции на границе доверительной области при κ . Тогда, согласно (6.30), с точностью до малых первого порядка вариации будут удовлетворять соотношению

$$\frac{1}{2} \frac{\delta S}{\bar{S} - \hat{S}} = \frac{\delta\kappa}{\kappa}.$$

Таким образом, если мы будем знать вариацию целевой функции δS , то по величине $\delta\kappa/\kappa$ сможем определить, насколько изменилась доверительная область в размерах.

Используя полученные оценки, введем некий коэффициент \varkappa , позволяющий в нелинейном случае приближенно оценить степень отклонения урванной поверхности, задаваемой уравнением (6.30), от эллипсоидальной, определяемой ковариационной матрицей параметров:

$$\varkappa(\mathbf{q}) = \frac{1}{2} \frac{\tilde{S} - \bar{S}}{\bar{S} - \hat{S}}, \quad \text{где} \quad \bar{S} = \hat{S} \left(1 + \frac{\kappa^2}{N - K} \right).$$

Здесь в качестве \tilde{S} выступают значения целевой функции нелинейной задачи на поверхности доверительного эллипсоида линеаризированной задачи при заданном κ . Обычно представляет интерес доверительная область, покрывающая истинные параметры с вероятностью, близкой к единице. Например, при $\alpha = 0.999$ и $K = 6$, когда N достаточно большое, $\kappa \approx 4.7$.

Естественно, коэффициент

$$\varkappa_{\max} = \max_{\mathbf{q}} |\varkappa(\mathbf{q})| \quad (6.31)$$

можно принять в качестве показателя нелинейности задачи. Очевидно, в линейном случае он равен нулю, а в нелинейном $\varkappa_{\max} \rightarrow 0$ при $\alpha \rightarrow 0$. Насколько известно автору, показатели нелинейности типа (6.31) впервые были предложены в работе (Черницов и др., 2006). Чтобы практически оценить степень нелинейности, достаточно найти максимум величины \varkappa лишь на конечном множестве точек \mathbf{q} , в качестве которых, как правило, рассматривают вершины доверительного эллипсоида. Исходя из личного опыта, автор рекомендует пользоваться нелинейными схемами стохастического моделирования, если $\varkappa_{\max} > 0.1$.

Достоверность результатов стохастического моделирования непосредственно зависит от внутренней кривизны подпространства оценок, фактически информирующей о том, насколько НК-задача внутренне нелинейна. Конечно же, не имеет смысла оценивать ее глобально, ведь нас в действительности интересует, насколько подпространство оценок отличается от плоского лишь в точках моделируемого выборочного множества.

В качестве показателя внутренней нелинейности предлагается ввести величину φ , определяемую соотношением (Avdyushev, 2009)

$$\cos(90^\circ + \varphi) = \frac{(\mathbf{p}^O - \hat{\mathbf{p}}) \cdot (\mathbf{p}^C - \hat{\mathbf{p}})}{\|\mathbf{p}^O - \hat{\mathbf{p}}\| \|\mathbf{p}^C - \hat{\mathbf{p}}\|}. \quad (6.32)$$

Геометрический смысл величины φ состоит в том, что она представляет собой угол между вектором $\mathbf{p}^C - \hat{\mathbf{p}}$ и плоским подпространством, касающимся подпространства оценок в точке $\hat{\mathbf{p}}$. Оче-

видно, для плоского подпространства $\varphi = 0$ в силу ортогональности векторов $\mathbf{p}^O - \hat{\mathbf{p}}$ и $\mathbf{p}^C - \hat{\mathbf{p}}$. Заметим, что подобная мера внутренней нелинейности предлагалась ранее в работе (Beale, 1960), и она фактически выражается через величину φ как $\sin^2 \varphi$ (Bates, Watts, 1980).

6.7. Сравнительный анализ эффективности методов

Рассмотренные алгоритмы моделирования возможных значений параметров были исследованы в сильно нелинейной обратной задаче спутниковой динамики. По немногочисленным наблюдениям (угловым координатам) \mathbf{p}^O нового спутника Юпитера S/2003 J04 (Emel'yanov, Arlot, 2008) на основе высокоточного моделирования его движения $\mathbf{p}^C(\mathbf{q})$ (Avdyushev, Banschikova, 2007) были определены НК-оценки орбитальных параметров спутника, в качестве которых взяты шесть компонент вектора динамического состояния (положения и скорости) $\mathbf{q} = (\mathbf{x}_0, \dot{\mathbf{x}}_0)^T$ в эпоху t_0 . При этом среднеквадратическая ошибка составила $\hat{\sigma} = 0.20''$.

Выбор объекта S/2003 J04 среди других спутников Юпитера был обусловлен прежде всего тем, что на его примере очень хорошо проявляется эффективность нелинейных схем стохастического моделирования возможных значений параметров. Астрометрические наблюдения спутника распределены на очень короткой орбитальной дуге: в течение нескольких десятков суток 11 моментов наблюдений ($N = 22$)⁴⁴, поэтому орбитальные параметры спутника определяются крайне ненадежно, тогда как нелинейность обратной задачи чрезвычайно сильная: $\varkappa_{\max} > 100$ (см. рис. 6.3). Вследствие такого специфического распределения наблюдений вероятностный разброс возможных значений параметров для S/2003 J04 настолько большой, что спутник может быть уверенно причислен к классу потерянных объектов (Avdyushev, Banschikova, 2007). Кстати, из рис. 6.3 видно, что есть ряд других

⁴⁴Каждому моменту соответствуют две угловые координаты: прямое восхождение и склонение.

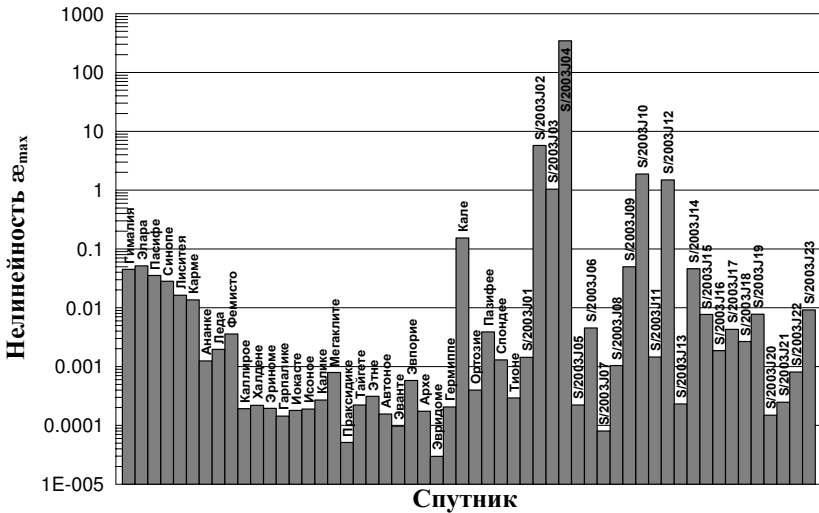


Рис. 6.3. Показатель нелинейности \varkappa_{\max} (6.31) для внешних спутников Юпитера при $\alpha = 0.999$

новых спутников, для которых сильная нелинейность обратной задачи ($\varkappa_{\max} > 0.1$) также имеет место.

НК-оценки орбитальных параметров S/2003 J04, определенные по реальным наблюдениям, были приняты за истинные $\bar{\mathbf{q}}$ и на их основе моделировались точные наблюдения спутника $\bar{\mathbf{p}} = \mathbf{p}^C(\bar{\mathbf{q}})$ на реальные моменты наблюдений. Затем в $\bar{\mathbf{p}}$ вносилась выборка нормально распределенных ошибок $\delta\mathbf{p}^O$ с дисперсией $\sigma^2 = \hat{\sigma}^2$ и по наблюдениям $\bar{\mathbf{p}} + \delta\mathbf{p}^O$ была получена НК-оценка $\hat{\mathbf{q}}$, для которой $\hat{\sigma} = 0.21''$.

Рассмотрим сначала два алгоритма: (6.19) с заменой σ^2 на $\hat{\sigma}^2$ и (6.17) с возмущениями $\delta\mathbf{p} = \hat{\sigma}\boldsymbol{\nu}_N$. С использованием этих алгоритмов построены множества возможных значений параметров, представленные 10000 решениями. Чтобы усилить эффект нелинейности, начальная эпоха t_0 была намеренно вынесена из временного интервала наблюдений, причем ее удаленность во времени от

моментов наблюдений составила порядка одного орбитального периода спутника.

При использовании неявной схемы (6.17) в качестве начального приближения \mathbf{q}_0 в методе Гаусса–Ньютона (6.6) выбиралась оценка $\hat{\mathbf{q}}$, но ввиду обширности области, покрываемой НК-решениями, такое приближение не всегда обеспечивало сходимость итерационного процесса. В этом случае применялся демпфированный метод Гаусса–Ньютона (Ortega, Reinboldt, 2000), а именно в схеме (6.6) уменьшалась поправка в 100 раз. Несмотря на то, что итерационный процесс существенно замедлялся, его сходимость достигалась для любых выборок наблюдений.

Результаты моделирования возможных значений орбитальных параметров спутника в йовицентрической координатной системе представлены на рис. 6.4. Как видно, облака возможных положений спутника чрезвычайно протяженные. Облако, моделируемое по линейной схеме (6.19), эллипсоидальное, однако оно не покрывает истинное положение спутника, в отличие от бананообразного облака, моделируемого по неявной схеме (6.17).

Необходимо отметить, что вероятностные облака перемешиваются только вблизи оценки $\hat{\mathbf{q}}$, поэтому можно полагать, что использование алгоритма (6.19) будет обоснованным только в случае, если разброс возможных спутниковых положений будет существенно меньше, в несколько десятков раз и более.

Как уже говорилось, моделирование возможных значений параметров на основе многократного решения НК-задачи обосновывается лишь фактом существования системы параметров \mathbf{r} , относительно которых НК-задача становится линейной. Таковую систему параметров можно получить, если K -мерное подпространство оценок, задаваемое моделью $\mathbf{p}^C(\mathbf{q})$ в N -мерном пространстве измеряемых величин \mathbf{p} , является плоским (Beale, 1960; Bates, Watts, 1980), т.е. любая точка этого подпространства может быть задана в некоторой прямоугольной системе координат, состоящей из K базисных векторов.

В нелинейной НК-задаче подпространство оценок \mathbf{p}^C , вообще говоря, неплоское и системы параметров \mathbf{r} , в которой задача

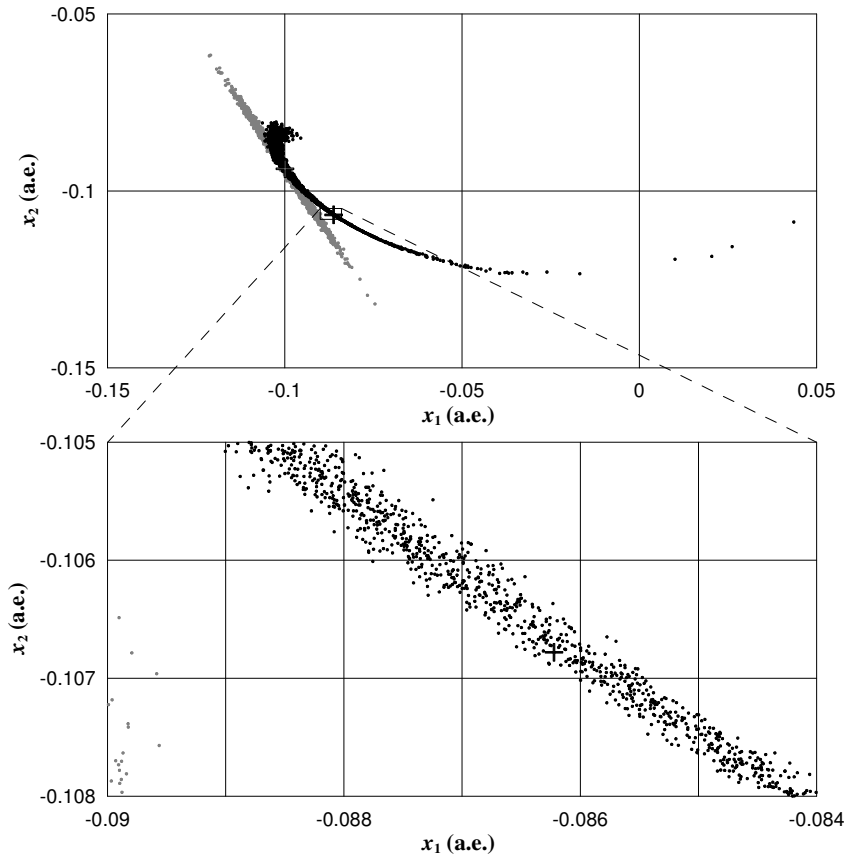


Рис. 6.4. Возможные положения спутника S/2003 J04 на координатной плоскости, построенные по алгоритмам (6.19) (серые точки) и (6.17) (черные точки), в окрестности истинного решения в разных масштабах. Проекция истинного решения $\bar{\mathbf{q}}$ обозначена черным крестиком, проекция оценки $\hat{\mathbf{q}}$ — серым

становится линейной, может не существовать. Тем не менее, если подпространство несущественно отличается от плоского (в интересующей нас области изменения параметрических оценок), то использование метода возмущенных оценок для моделирования возможных значений параметров будет вполне оправданным.

На рис. 6.5 показано распределение угловых отклонений φ (6.32) подпространства \mathbf{p}^O в точках возможных значений \mathbf{q} от касательного в НК-оценке $\hat{\mathbf{q}}$ плоского подпространства. Как видно из рисунка, значения φ достаточно малы и большая их часть не превосходит 0.1° . Следовательно, сильная нелинейность задачи вызвана главным образом неудачным выбором оцениваемых параметров (Bates, Watts, 1980; Bates, Watts, 1988). Так или иначе, можно полагать, что построенное множество НК-решений должно достаточно хорошо представлять распределение возможных значений орбитальных параметров. Кроме того, при использовании неявной схемы (6.17) вариации $\delta\mathbf{p}$ можно вносить не в оценочные представления наблюдений $\hat{\mathbf{p}}$, а в сами наблюдения \mathbf{p}^O .

Модифицированная схема метода возмущенных оценок с $\delta\mathbf{p} = r\hat{\sigma}\nu_N$ исследовалась в той же задаче, но при использовании лишь некоторых наблюдений, отнесенных на 4 момента времени ($N = 8$), также покрывающих весь период наблюдения спутника. Именно при таких условиях обратной задачи (со скудной наблюдательной информацией) можно ожидать проявление высокой эффективности схемы.

Как и выше, по выборке искусственно созданных наблюдений $\bar{\mathbf{p}} + \delta\mathbf{p}^O$ объема $N = 8$ была получена новая оценка $\hat{\mathbf{q}}$, для которой среднеквадратическая ошибка составила $\hat{\sigma} = 0.13''$. Затем моделировались возможные положения спутника по схемам с возмущениями $\delta\mathbf{p} = \hat{\sigma}\nu_N$ и $\delta\mathbf{p} = r\hat{\sigma}\nu_N$. Результаты приведены на рис. 6.6.

Как видно из рисунка, множества возможных положений, построенные по неявным схемам с разными возмущениями $\delta\mathbf{p}$, существенно отличаются друг от друга и описывают принципиально разные распределения вероятностной плотности в пространстве координат. При этом вероятностная плотность по схеме с

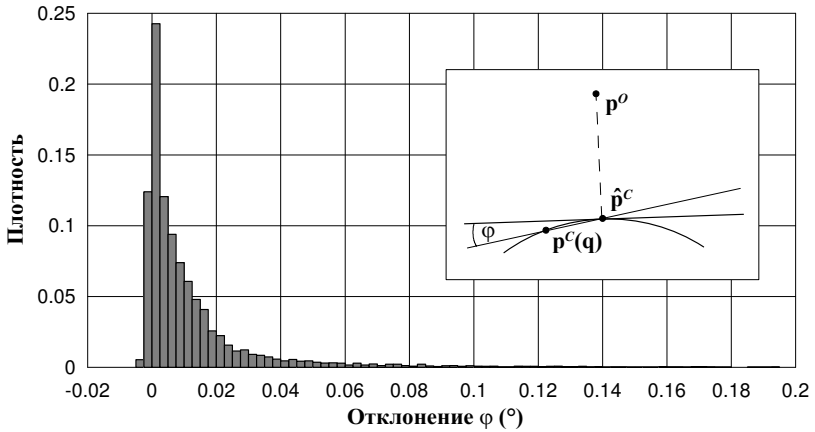


Рис. 6.5. Распределение отклонений φ (6.32) подпространства \mathbf{p}^C в точках множества возможных значений параметров от касательного в НК-оценке $\hat{\mathbf{q}}$ плоского подпространства

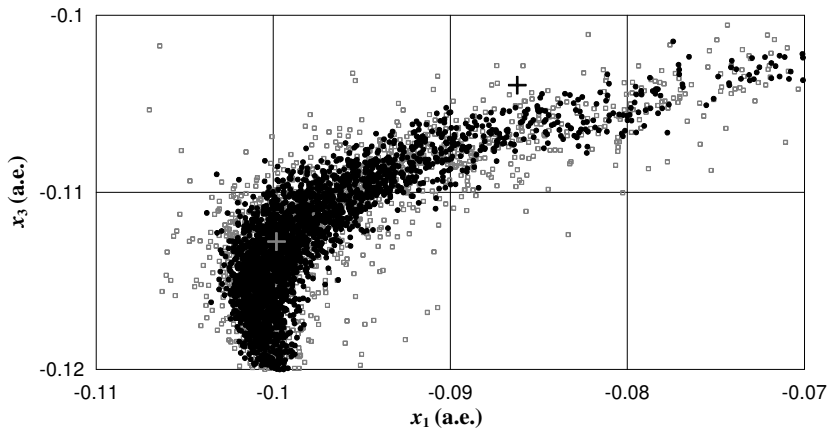


Рис. 6.6. Возможные положения спутника S/2003 J04 на координатной плоскости, построенные по схеме (6.17) с возмущениями $\delta \mathbf{p} = \hat{\sigma} \nu_N$ (точки) и $\delta \mathbf{p} = r \hat{\sigma} \nu_N$ (серые квадратики)

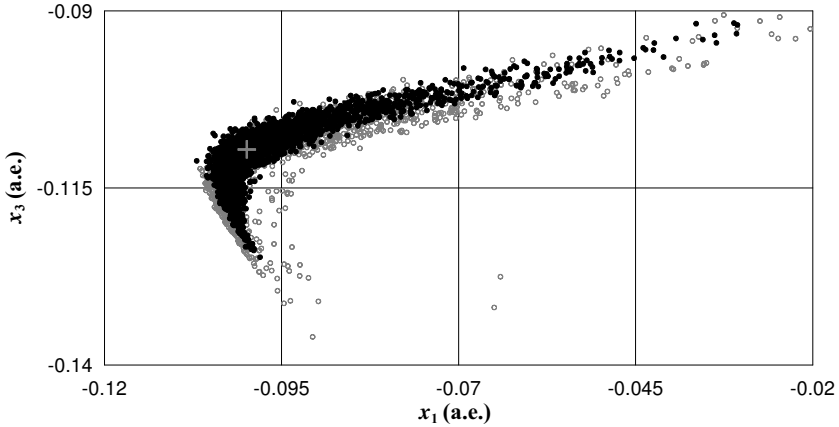


Рис. 6.7. Возможные положения спутника S/2003 J04 на координатной плоскости, построенные методом возмущенных оценок (точки) и бутстрэп-методом (серые кружочки)

$\delta \mathbf{p} = \hat{\sigma} \boldsymbol{\nu}_N$ в точке истинного положения спутника оказывается значительно меньше, нежели по схеме с $\delta \mathbf{p} = r \hat{\sigma} \boldsymbol{\nu}_N$. Более того, согласно распределению возможных положений, полученных по первой схеме, появление спутника в окрестности истинного положения оказывается почти невероятным событием. Очевидно, это в данном случае является веским аргументом в пользу модифицированной схемы с модулирующим множителем r как более эффективной с точки зрения стохастического моделирования возможных значений параметров.

Ввиду того, что при $N = 8$ и $K = 6$ бутстрэп-метод дает недостаточное число возможных орбит (6.29), он тестировался в задаче с исходным количеством наблюдений $N = 22$. Предполагалось, что бутстрэп-распределение возможных решений будет хорошо согласовываться с распределением, полученным по схеме метода возмущенных оценок (6.17), поскольку заранее было известно, что ошибки в наблюдениях представляют собой реализацию нормально распределенных случайных величин, но между тем имен-

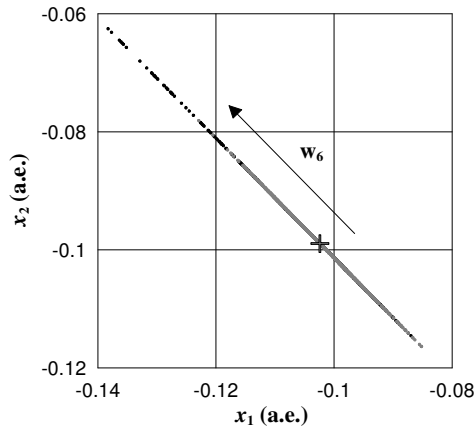


Рис. 6.8. Результаты стохастического моделирования для S/2003 J04 на момент времени внутри интервала наблюдений (линейные оценки представлены серым цветом, нелинейные — черным)

но на основе этого предположения моделируются возможные значения параметров по схеме (6.17). Вопреки ожиданиям, согласие между распределениями (рис. 6.7), в особенности на периферии множеств возможных положений спутника, оказалось не очень хорошим. Кроме того, имеют место непонятные выбросы некоторых положений в бутстрэп-множестве. По этой причине мы бы не рекомендовали использовать бутстрэп-метод для исследования неопределенностей в орбитальных параметрах новых спутников с бедной хронологией наблюдений.

Как уже говорилось, чтобы усилить эффект нелинейности, начальная эпоха была намеренно вынесена из временного интервала наблюдений на один оборот вперед. Если же моделировать возможные значения на эпоху внутри интервала наблюдений, то результаты линейного и нелинейного моделирования становятся визуально очень похожими (рис. 6.8). Характерной особенностью в распределении результатов стохастического моделирования является то, что возможные спутниковые положения располагают-

ся главным образом вдоль направления от наблюдателя на спутник (Avdyushev, Vanschikova, 2007), что как раз вызвано отсутствием измерительной информации о дальности объекта, поскольку спутник наблюдается на небесной сфере в топоцентрических угловых координатах, причем на очень короткой орбитальной дуге. В действительности, результаты моделирования близки лишь с точки зрения представления наблюдений, но уже существенные различия могут проявиться при динамическом прогнозе возможных спутниковых положений, полученных на основе возможных начальных значений орбитальных параметров.

Для более детального анализа конфигурации возможных положений в параметрическом пространстве ориентируем их вдоль собственных векторов \mathbf{w}_i ковариационной матрицы и рассмотрим в различных проекциях и в разных масштабах. На рис. 6.9 приведены пять проекций на плоскостях \mathbf{w}_i – \mathbf{w}_6 ($i = 1, \dots, 5$), где собственный вектор \mathbf{w}_6 соответствует наибольшему собственному числу (т.е. он как раз указывает направление, вдоль которого вытягивается доверительная область), и, как видно, конфигурации возможных значений разительно отличаются друг от друга.

В случае новых спутников искажение областей возможных параметрических значений нелинейностью связано главным образом с их явно большими размерами. Несмотря на то, что вероятностные области для близких спутников существенно меньше, они тем не менее оказываются еще достаточно большими для их описания в контексте линейной НК-задачи. Данные табл. 6.1 наглядно свидетельствуют, насколько существенно влияние нелинейности с точки зрения оценивания точности орбитальных параметров: так, показатель \varkappa_{\max} оказывается больше 10^3 .

Интересно заметить, что если для новых (далеких) спутников группы S/2003 на доверительном эллипсоиде линеаризированной задачи имеют место лишь две аномальные вершины, ориентированные вдоль направления \mathbf{w}_6 (см. рис. 6.8 и 6.9), то для близких спутников Юпитера, наоборот, только в двух вершинах, соответствующих вектору \mathbf{w}_1 , значения \varkappa достаточно малы (см. \varkappa_{\min} в табл. 6.1), тогда как в других \varkappa значительно больше единицы.

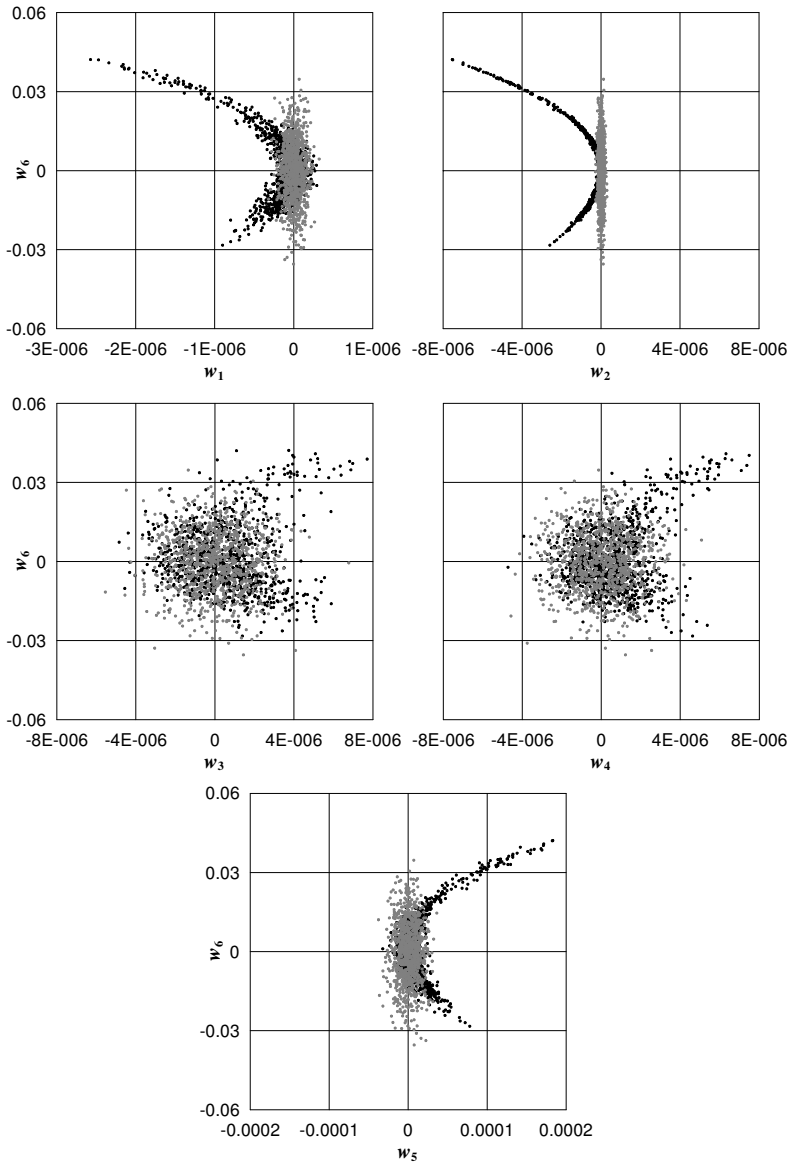


Рис. 6.9. Возможные значения параметров для S/2003 J04 относительно НК-оценок

Таблица 6.1. Минимальный и максимальный показатели нелинейности \varkappa для близких спутников Юпитера

Спутник	\varkappa_{\min}	\varkappa_{\max}
Амальтея	$4.7 \cdot 10^{-3}$	$4.0 \cdot 10^5$
Теба	$2.8 \cdot 10^{-3}$	$2.7 \cdot 10^3$
Адрастея	$8.6 \cdot 10^{-2}$	$2.5 \cdot 10^4$
Метида	$1.9 \cdot 10^{-2}$	$1.5 \cdot 10^5$

В действительности собственные числа ковариационных матриц для близких спутников выявляют блинообразную форму доверительной области (Авдюшев, Баныщикова, 2008): одно из собственных чисел существенно меньше других. Причем область в параметрическом пространстве \mathbf{q} оказывается сплюснутой вдоль частотной координаты типа среднего движения n , и она выстилается вдоль поверхности $n(\mathbf{q}) = n(\hat{\mathbf{q}}) = \hat{n}$. Однако поскольку гиперповерхность $n(\mathbf{q})$, вообще говоря, не плоская, то эллипсоидальная вероятностная область, построенная по ковариационной матрице, при достаточно большом разбросе параметрических ошибок будет существенно отличаться от реальной доверительной области, что и объясняет разительно большие значения \varkappa .

В такой ситуации для исследования распределения параметрических ошибок необходимо прибегать к нелинейному оцениванию. Единственное и эффективное, на наш взгляд, подспорье, которое у нас имеется к настоящему времени для разрешения данной проблемы, — это предложенный выше метод возмущенных оценок. Однако, как отмечалось, его значительным недостатком является многократное численное решение нелинейной НК-задачи, что делает практически невозможным его реализацию применительно к оцениванию точности орбитальных параметров близких спутников, поскольку выполнение каждой итерации метода Гаусса–Ньютона вследствие долгосрочного численного интегрирования даже на современных компьютерах занимает несколько часов. В связи с этими трудностями не будем строить области возмож-

ных значений параметров для близких спутников с использованием их высокоточных динамических моделей, но, по крайней мере, поставим перед собой цель высветить проблему нелинейности.

Заметим, что построение облака возможных параметрических значений на основе НК-разброса оказывается выполнимым при использовании аналитических моделей движения, в частности, на основе задачи двух тел. Несмотря на то, что результаты определения орбит из семейства кеплеровских не имеют практическую ценность, на их примере, так или иначе, можно увидеть проявление влияния нелинейности на оценки параметрической точности.

Используя модель кеплеровского движения (см. Приложение), как и в предыдущей главе, воспроизведем обратную задачу для близкого спутника Юпитера Адрастеи и по схемам (6.19) и (6.17) (линейного и нелинейного оценивания) смоделируем множества возможных значений параметров. Причем среднеквадратическую ошибку наблюдений примем $0.2''$. На рис. 6.10 показаны проекции этих множеств, ориентированные относительно собственных векторов \mathbf{w}_i соответствующей ковариационной матрицы (иначе говоря, относительно главных осей доверительного эллипсоида линеаризированной задачи). Здесь приводятся только проекции вдоль собственного вектора \mathbf{w}_1 с наименьшим собственным числом, и, как видно из рисунка, вероятностные облака совершенно не согласуются друг с другом.

Принимая во внимание масштаб оси абсцисс, может показаться, что отличие между вероятностными облаками в действительности незначительно. Однако напомним, что именно вдоль направления \mathbf{w}_1 целевая функция наиболее чувствительна к изменению параметров, точнее, к изменению среднего движения или большой полуоси (см., например, рис. 5.1).

Вместе с тем интересно заметить, что распределения величин $\nu(\mathbf{q}) = n(\mathbf{q})/\hat{n} - 1$, получаемых по возможным решениям \mathbf{q} , при линейном и нелинейном оценивании имеют соответственно среднеквадратичные отклонения $1.1 \cdot 10^{-5}$ и $2.6 \cdot 10^{-8}$. Последнее свидетельствует о том, что вероятностная область нелинейного оценивания распространяется преимущественно вдоль поверхности

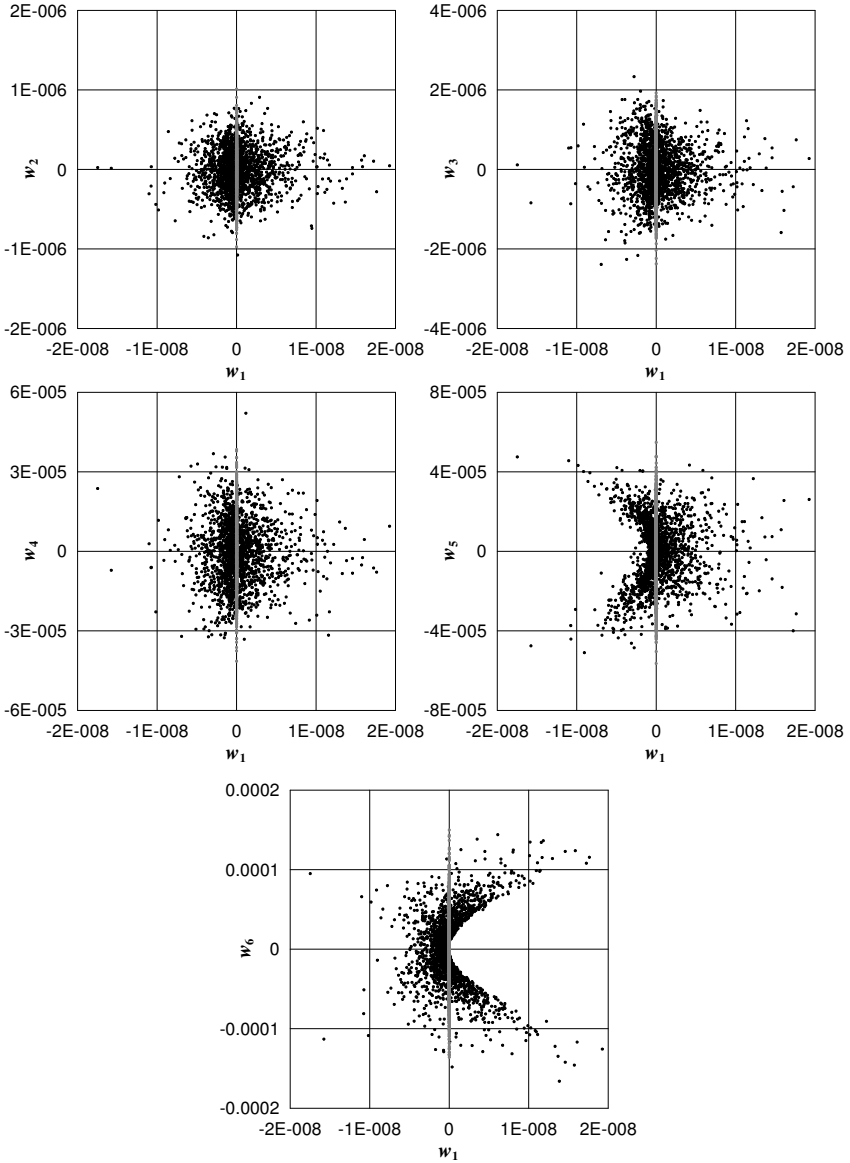


Рис. 6.10. Возможные значения параметров для Адрастеи относительно НК-оценок

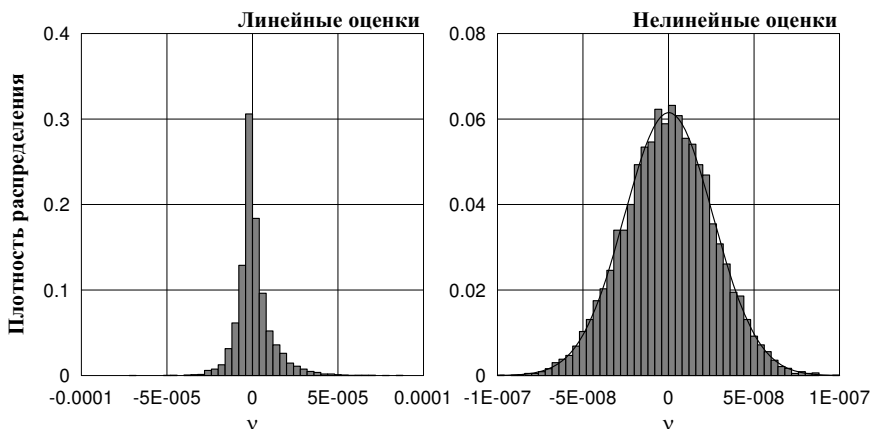


Рис. 6.11. Распределение величины ν . Колоколообразной линией в нелинейном случае представлена аппроксимирующая функция нормального распределения

$n(\mathbf{q})$. Кроме того, в этом случае, в отличие от линейного, значения ν , как и ожидалось, распределяются по нормальному закону (рис. 6.11), причем дисперсия аппроксимирующей функции нормального распределения согласуется с дисперсией распределения ν с точностью до четвертого знака.

Таким образом, поскольку нелинейность слабо проявляется только в одном направлении, вдоль изменения частотной переменной типа среднего движения, то именно для этой переменной можно получить оценку точности из ковариационной матрицы. Примечательно, что как раз эта оценка имеет наибольшее прикладное значение с точки зрения планирования наблюдений близкого спутника и его идентификации, так как частотная переменная непосредственно определяет движение объекта по орбите, в то время как конфигурация самой орбиты определяется из наблюдений с приемлемой точностью.

В табл. 6.2 для внутренних спутников Юпитера в качестве примера приведены оценки точности σ_n частотной переменной,

Таблица 6.2. Оценки точности частотной переменной

Спутник	σ_n (сут ⁻¹)	σ_ν	$\Delta\tilde{t}$ (г.)
Амальтея	$1.0 \cdot 10^{-7}$	$8.2 \cdot 10^{-9}$	440
Теба	$2.9 \cdot 10^{-7}$	$3.2 \cdot 10^{-8}$	125
Адрастея	$1.3 \cdot 10^{-6}$	$6.3 \cdot 10^{-8}$	48
Метида	$5.1 \cdot 10^{-7}$	$2.4 \cdot 10^{-8}$	124

как если бы она была средним движением $n = n(\mathbf{q})$. Оценки вычислены по вариациям параметров \mathbf{q} , полученным из соответствующей ковариационной матрицы вдоль интересующего нас направления. В таблице также представлены относительная ошибка частотной переменной σ_ν и интервал времени $\Delta\tilde{t}$, за который ошибки в прогнозируемых координатах спутника, вызванные ошибкой σ_n , будут отклонять геоцентрическое угловое положение объекта на $1''$. В частности из таблицы видно, что при требуемой точности прогноза движения порядка $1''$ дуги модели спутниковых орбит с полученными из наблюдений параметрами вполне могут быть использованы для планирования будущих наблюдений спутников на 50 лет вперед и более.

Впрочем, следует иметь в виду, что результаты в табл. 6.2 относятся к вполне определенной системе параметрических оценок, тогда как в случае близких спутников, как мы знаем, таких систем, доставляющих малые минимальные значения целевой функции, может быть много. Поэтому в соответствии с особенностями обратных задач динамики близких спутников необходимо рассматривать всю совокупность несвязанных между собой вероятностных областей, отнесенных к своей системе орбитальных параметров. Хотя к построению такой совокупности, очевидно, имеет смысл прибегать, когда в действительности не известно, какая из систем параметров наилучшим образом представляет реальную спутниковую орбиту. В таком случае для идентификации и планирования наблюдений спутников, пожалуй, будет более полезна совокупность самих НК-оценок, а не вероятностных областей,

ввиду чрезвычайно малых размеров последних в сравнении с расстояниями между ними в параметрическом пространстве.

Таким образом, в главе представлен обзор основанных на концепциях классического регрессионного анализа линейных и нелинейных методов стохастического моделирования возможных значений параметров для исследования точности орбит, определяемых из наблюдений. На примере спутника Юпитера S/2003 J04, параметры движения которого плохо определены вследствие скудного состава имеющихся на данный момент наблюдений, демонстрируется высокая эффективность нелинейных методов. Поскольку сильная нелинейность обратной задачи проявляется, как правило, в случае скудного состава наблюдательных данных, то на основании теоретических и численных результатов для эффективного стохастического моделирования возможных значений параметров можно рекомендовать неявную схему (6.17) с возмущениями $\delta \mathbf{p} = r \hat{\sigma} \boldsymbol{\nu}_N$ либо ее взвешенный аналог (6.23), если веса наблюдений известны.

Экспериментально также показано, что нелинейность используемых моделей для близких (внутренних) спутников Юпитера настолько значительна, что ковариационные матрицы, строго говоря, оказываются совершенно бесполезными для вероятностного описания распределений параметрических ошибок и в данном случае следует прибегать к нелинейному оцениванию.

— Извините, я опоздал.

— Что случилось?

— Да ничего, я просто не хотел приходить.

*Шелдон Купер из ситкома
«Теория большого взрыва»*

Глава 7. ЛИНЕАРИЗАЦИЯ ОРБИТАЛЬНОЙ МОДЕЛИ

Эта небольшая глава не является обязательной и ее можно было бы опустить, поскольку рассматриваемые в ней вопросы частные. Однако из соображения того, что они могут вызвать у некоторых читателей определенный практический интерес (Авдюшев, 2012), глава все же была оставлена в монографии.

Итак, если орбита исследуемого небесного тела с орбитальными параметрами $\hat{\mathbf{q}}$ известна:

$$\mathbf{p}^C = \mathbf{p}^C(t, \hat{\mathbf{q}}),$$

то орбиту любого близкого попутчика с параметрами \mathbf{q} приближенно (с точностью до малых первого порядка) можно представить в виде

$$\mathbf{p} = \mathbf{p}^C(t, \hat{\mathbf{q}}) + \left. \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right|_{t, \hat{\mathbf{q}}} (\mathbf{q} - \hat{\mathbf{q}}), \quad (7.1)$$

где $\partial \mathbf{p}^C / \partial \mathbf{q}$ — по сути те же изохронные производные (5.15), которые участвуют в решении обратной задачи. Таким образом, механический смысл выражения (7.1) состоит в том, что оно описывает локальную динамику около номинальной орбиты $\hat{\mathbf{q}}$, и чем ближе к номинальной орбите, тем лучше. Поэтому очевидное назначение (7.1) — это отображение близких (к номинальному) возможных положений небесного тела на заданные моменты времени. Далее рассмотрим некоторые ситуации, когда целесообразно прибегать к линейным отображениям.

7.1. Временной перенос доверительной области

Линейные отображения могут служить полезным подспорьем для исследования эволюции малых доверительных областей, а также компактных разбросов возможных значений параметров.

Предположим, доверительная область, характеризующая параметрическую неопределенность в \mathbf{q} , настолько мала, что может рассматриваться как гиперэллипсоид (6.12), описываемый ковариационной матрицей \mathbf{C}_q . Тогда, согласно линейному преобразованию (7.1), неопределенность в величинах \mathbf{p} , порождаемая неопределенностью в параметрах \mathbf{q} , на заданные моменты времени t будет приближенно описываться доверительным эллипсоидом с центром $\hat{\mathbf{p}} = \mathbf{p}^C(t, \hat{\mathbf{q}})$ и ковариационной матрицей

$$\mathbf{C}_p = \left. \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right|_{t, \hat{\mathbf{q}}} \mathbf{C}_q \left. \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right|_{t, \hat{\mathbf{q}}}^T. \quad (7.2)$$

Производные $\partial \mathbf{p}^C / \partial \mathbf{q}$ в (7.2) вычисляются через $\partial \mathbf{x} / \partial \mathbf{q}$ (5.15), которые, в свою очередь, определяются из уравнений в вариациях (5.16). Таким образом, для исследования эволюции доверительного эллипсоида, как и при решении обратной задачи, необходимо численно интегрировать расширенную систему уравнений

$$\frac{d^2 \mathbf{x}}{dt^2} = \mathcal{P}, \quad \frac{d^2}{dt^2} \frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \frac{\partial \mathcal{P}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{q}} + \frac{\partial \mathcal{P}}{\partial \dot{\mathbf{x}}} \frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{q}}.$$

Со временем значения изохронных производных, как правило, возрастают, и это свидетельствует об увеличении доверительной области, что сопряжено с усилением эффекта нелинейности, вызванного нелинейностью орбитальной модели. Поэтому нужно ясно осознавать, что какой бы малой ни была изначально доверительная область, линейные отображения (7.1) пригодны лишь на умеренных интервалах времени, когда нелинейность слаба. На длительных же интервалах получаемая ковариационная матрица \mathbf{C}_p (7.2) будет грубо описывать доверительную область, хотя в некоторых случаях доставляемая ею информация бывает вполне

достаточной для качественного представления неопределенностей в орбите. Между тем, чтобы оценить степень влияния нелинейности, можно, как и при оценивании параметрической точности, воспользоваться уже известной нам формулой (6.31).

Разумеется, ковариационная матрица \mathbf{C}_p применима не только для описания доверительной области, но и для стохастического моделирования возможных положений в пространстве \mathbf{p} с использованием линейной схемы (6.19). Однако с вычислительной точки зрения, если требуются оценки на многочисленные моменты времени, проще осуществлять временной перенос уже смоделированных возможных значений параметров \mathbf{q} линейными отображениями (7.1), нежели моделировать разброс величин \mathbf{p} , постоянно перевычисляя матрицу Холецкого $\mathbf{C}_p^{1/2}$ для схемы (6.19) на каждый момент времени.

7.2. Вычисление ляпуновских показателей

Поскольку изохронные производные хорошо описывают локальную динамику около номинальной орбиты, они могут быть полезны для определения скорости разбегания возможных бесконечно близких положений, что бывает необходимо при вычислении ляпуновских показателей с целью выявления хаоса в орбитальном движении.

Показатель Ляпунова χ определяется как (Murray, Dermott, 1999)

$$\chi(t) = \frac{\ln \|\Delta \mathbf{p}(t)\| / \|\Delta \mathbf{p}(t_0)\|}{t - t_0}, \quad (7.3)$$

где $\Delta \mathbf{p} = \mathbf{p} - \hat{\mathbf{p}}$ — отклонение смежного возможного положения от номинального, которое вызвано вариацией параметров модели $\hat{\mathbf{q}}$. Если $\chi \rightarrow \text{const}$ при $t \rightarrow \infty$, орбита — хаотична, и в этом случае показатель Ляпунова (7.3) будет представлять собой скорость экспоненциального разбегания возможных положений.

Чтобы точнее определить скорость разбегания, вариации параметров должны быть достаточно малыми. Между тем они должны быть достаточно большими, чтобы вычислительные ошибки

не влияли существенно на оценку скорости. Эта дилемма разрешается, если прибегнуть к линейным отображениям.

Поскольку величина $\Delta \mathbf{p}$ должна быть малой, то для ее оценки можно воспользоваться формулой (7.1) и тогда

$$\frac{\|\Delta \mathbf{p}(t)\|}{\|\Delta \mathbf{p}(t_0)\|} = \frac{\|(\partial \mathbf{p}^C / \partial \mathbf{q})_{t, \hat{\mathbf{q}}}(\mathbf{q} - \hat{\mathbf{q}})\|}{\|(\partial \mathbf{p}^C / \partial \mathbf{q})_{t_0, \hat{\mathbf{q}}}(\mathbf{q} - \hat{\mathbf{q}})\|}. \quad (7.4)$$

В свою очередь, для вычисления этого отношения требование малости отклонений уже не обязательно.

Обычно при исследовании орбитального движения на предмет хаотичности $\mathbf{p} = (\mathbf{x}, \dot{\mathbf{x}})^T$ и $\mathbf{q} = (\mathbf{x}_0, \dot{\mathbf{x}}_0)^T$ — векторы динамического состояния на текущий момент времени t и начальный t_0 соответственно. Тогда матрица изохронных производных становится не чем иным, как матрицей перехода (Montenbruck, Gill, 2000)

$$\frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} = \Phi(t, t_0) = \begin{pmatrix} \partial \mathbf{x} / \partial \mathbf{x}_0 & \partial \mathbf{x} / \partial \dot{\mathbf{x}}_0 \\ \partial \dot{\mathbf{x}} / \partial \mathbf{x}_0 & \partial \dot{\mathbf{x}} / \partial \dot{\mathbf{x}}_0 \end{pmatrix}.$$

Очевидно, что величина отношения (7.4) зависит от выбора вариации $\mathbf{q} - \hat{\mathbf{q}}$, однако для определения ляпуновского показателя имеет смысл использовать максимальную величину, которая фактически является нормой матрицы перехода (Lekien, Ross, 2010):

$$\max_{\mathbf{q} - \hat{\mathbf{q}}} \frac{\|\Delta \mathbf{p}(t)\|}{\|\Delta \mathbf{p}(t_0)\|} = \|\Phi(t, t_0)\|.$$

Вычисление нормы симметричной матрицы на практике не составляет особого труда, но матрица перехода Φ не симметрична. Чтобы найти ее норму, можно прибегнуть к симметризации $\Phi^T \Phi$. Корень из максимального собственного числа λ_{\max} симметричной матрицы $\Phi^T \Phi$ как раз и будет нормой матрицы перехода Φ . Тогда показатель Ляпунова можно переопределить как (Ross, Tallapragada, 2012)

$$\chi(t) = \frac{\ln \sqrt{\lambda_{\max}}}{t - t_0}.$$

Для простоты вычислений вполне допустимо в качестве матрицы Φ рассматривать матрицу производных $\partial \mathbf{x} / \partial \mathbf{x}_0$ в предположении, что, возникнув в фазовом пространстве, хаос непременно даст о себе знать и в физическом.

7.3. Оценивание вероятности попадания объекта в малый объем

Имея множество возможных значений параметров, можно непосредственно вычислить вероятность нахождения истинных значений в любом ограниченном объеме параметрического пространства как отношение числа возможных решений, попавших в объем, к общему числу. Очевидно, чем больше используемых решений, тем точнее будет оценка вероятности. Когда разброс возможных решений описывается в рамках линейной обратной задачи, никаких серьезных затруднений не возникает: используя лишь ковариационную матрицу для формирования разброса, можно достаточно быстро получить нужное число решений для оценивания искомой вероятности с заданной точностью.

Проблема возникает при сильной нелинейности, что предполагает многократное численное решение нелинейной обратной задачи. Более того, ситуация усугубляется, если на основе построенного разброса значений орбитальных параметров, отнесенного на начальный момент времени, требуется определить вероятность попадания небесного тела в ограниченный объем какого-либо другого, непараметрического пространства на некоторый удаленный от начального момент времени. Это, например, часто необходимо в задачах астероидной опасности для оценки вероятности столкновения астероида с Землей, либо может быть полезным при планировании наблюдений объекта в будущем для оценки вероятности появления объекта на наблюдаемом участке неба. Отображение разброса возможных решений в заданное пространство и на заданный момент времени осуществляется на основе численного моделирования орбитального движения, что с необходимостью предполагает пошаговое численное интегрирование, требу-

ющее значительных затрат процессорного времени. К тому же на практике рассматриваемые объемы, как правило, очень малы, что, разумеется, сопряжено с малыми вероятностями попадания в них объектов. Следовательно, вычисление таких вероятностей должно быть чрезвычайно трудоемким, поскольку оно фактически выполнимо лишь при наличии весьма большого количества возможных решений.

Для разрешения указанных трудностей можно использовать быстрые отображения типа (7.1). Вследствие того, что эти отображения приближенные, их оценки вероятности могут не совпадать с оценками, получаемыми при использовании исходных орбитальных моделей, хотя различия в них могут быть практически незначительными. Тем не менее, если все же, скажем, в задачах астероидной опасности требуется точная оценка, то быстрые отображения, по крайней мере, можно использовать для предварительного отбора потенциально опасных возможных решений, увеличив в несколько раз область-мишень. Конечно, в этом случае численного интегрирования орбит не избежать, однако их будет уже существенно меньше, нежели всех возможных орбит.

Предположим, используя модель движения $\mathbf{p}^C(t, \mathbf{q})$, необходимо оценить вероятность попадания объекта в малый объем $V_{\mathbf{p}}$ в L -мерном пространстве \mathbf{p} на момент времени t_1 , когда на момент t_0 известен вероятностный разброс K -мерного параметрического вектора \mathbf{q} , который хорошо описывается ковариационной матрицей. При этом возможные движения внутри малого объема $V_{\mathbf{p}}$ происходят почти прямолинейно и равномерно.

Для определения плотности возможных положений объекта в области $V_{\mathbf{p}}$ на момент t_1 на основе начального вероятностного разброса \mathbf{q} воспользуемся линейным отображением, сконструированным относительно точки \mathbf{q}^* (целевая орбита), являющейся прообразом некоторой точки $\mathbf{p}^* \in V_{\mathbf{p}}$. Вообще говоря, точка \mathbf{q}^* не является единственной, поскольку прообразом точки \mathbf{p}^* в параметрическом K -мерном пространстве является некое $(K - L)$ -мерное подпространство целевых орбит. Резонным было бы выбрать на нем ту точку \mathbf{q}^* , в которой плотность множества возможных зна-

чений параметров максимальна. Такая точка должна доставлять минимум квадратичной формы

$$\mathcal{Q}(\mathbf{q}) = \frac{1}{2}(\mathbf{q} - \hat{\mathbf{q}})^T \mathbf{Q}(\mathbf{q} - \hat{\mathbf{q}})$$

на подпространстве преобразов. Таким образом, отыскание прообраза \mathbf{q}^* фактически сводится к задаче условной минимизации функции \mathcal{Q} по \mathbf{q} , где в качестве ограничения выступает векторное равенство $\mathbf{p}^C(t_1, \mathbf{q}) = \mathbf{p}^*$. В свою очередь, решение этой задачи находится из минимизации функции Лагранжа (Бертсекас, 1987)

$$\mathcal{L}(\mathbf{q}) = \mathcal{Q}(\mathbf{q}) + \boldsymbol{\lambda} \cdot (\mathbf{p}^C(t_1, \mathbf{q}) - \mathbf{p}^*) \quad (7.5)$$

по параметрическому вектору \mathbf{q} и L -мерному вектору множителей Лагранжа $\boldsymbol{\lambda}$, для чего можно воспользоваться методом Гаусса–Ньютона с начальным приближением $\hat{\mathbf{q}}$.

Наконец, пользуясь линейным отображением $\mathbf{q}(t_0) \rightarrow \mathbf{p}(t_1)$:

$$\mathbf{p} = \mathbf{p}^* + \left. \frac{\partial \mathbf{p}^C}{\partial \mathbf{q}} \right|_{t_1, \mathbf{q}^*} (\mathbf{q} - \mathbf{q}^*), \quad (7.6)$$

для каждого возможного \mathbf{q} вычисляем \mathbf{p} и регистрируем попадание отображенного положения в объем $V_{\mathbf{p}}$, т.е. выполнение условия $\mathbf{p} \in V_{\mathbf{p}}$. Отношение числа возможных решений, для которых это условие выполняется, к их общему числу будет давать статистическую оценку вероятности попадания объекта в малый объем $V_{\mathbf{p}}$ на момент t_1 .

7.4. Оценивание вероятности столкновения астероида с планетой

Рассмотрим теперь особый случай, когда $\mathbf{p} = \mathbf{x}$ и любое возможное движение объекта в объеме $V_{\mathbf{p}}$ существенно отличается от прямолинейного и равномерного вследствие наличия в нем гравитирующей материальной точки (сингулярности). Это типично имеет место при оценивании вероятности столкновения астероида

с планетой. Очевидно, что в данном случае использование линейных преобразований типа (7.6) уже будет необоснованным, каким бы малым не был объем $V_{\mathbf{p}}$, содержащий сингулярность.

Явление тесного сближения астероида с планетой, как правило, кратковременно. Несмотря на это, оно приводит к серьезной трансформации астероидной орбиты. Между тем движение астероида во время сближения происходит главным образом под влиянием массивной планеты, поэтому оно достаточно хорошо может быть описано простыми (нелинейными) формулами задачи двух тел. Эта особенность указывает на целесообразность применения составного отображения для разрешения проблемы сильной нелинейности, вызванной сингулярностью.

Составной подход состоит в том, что до тесного сближения облако виртуальных астероидов отображается линейно по формулам (7.6), тогда как во время сближения — нелинейно по формулам задачи двух тел (Авдюшев, 2012). Сфера влияния планеты может рассматриваться как область сильной нелинейности. Тогда t_1 в (7.6) будет моментом времени прохождения астероида на целевой орбите с параметрами \mathbf{q}^* через сферу влияния.

Для функции Лагранжа, очевидно, следует принять ограничение

$$|\mathbf{p}^C(t_1, \mathbf{q}) - \mathbf{p}^*|^2 = b_S^2,$$

где \mathbf{p}^* — положение планеты в физическом пространстве, а b_S — радиус сферы влияния. Следовательно, имеем целевую функцию

$$\mathcal{L}(\mathbf{q}) = \mathcal{Q}(\mathbf{q}) + \lambda(|\mathbf{p}^C(t_1, \mathbf{q}) - \mathbf{p}^*|^2 - b_S^2). \quad (7.7)$$

Проблема, однако, состоит в том, что итерационные методы, минимизирующие функцию (7.7), вследствие сильной нелинейности чрезвычайно плохо сходятся, и практически редко удается найти значения параметров \mathbf{q}^* целевой орбиты. Впрочем, они могут быть определены иным способом (Авдюшев, 2012).

Возьмем сначала в качестве \mathbf{q}^* вектор оценок $\hat{\mathbf{q}}$ и выполним интегрирование орбиты до момента времени t_1 , когда либо будет достигнуто минимальное расстояние до планеты, либо номинальный астероид пересечет сферу влияния. На этот момент

отображаем начальное облако виртуальных астероидов по формулам (7.6). Находим среди отображенных астероидов тот, который максимально близок к центру планеты. Выбираем теперь его параметры в качестве номинальных, и повторяем описанные действия. И так далее до тех пор, пока до и после очередной итерации не получим те же самые номинальные значения параметров. Собственно, на этом этапе уже нужны не столько они, сколько линейно отображенное относительно них облако виртуальных астероидов, которое впоследствии отображается нелинейно по формулам задачи двух тел. Как показывает практика (Авдюшев, Галушина, 2014), поиск номинальных значений параметров на множестве возможных обычно требует всего несколько итераций, поэтому составное отображение может быть весьма востребованным для оперативного определения вероятности столкновения астероида с планетой, в особенности, если эта вероятность достаточно мала, а столкновение ожидается в неблизком будущем.

Приложение

ЗАДАЧА ДВУХ ТЕЛ. ЭЛЛИПТИЧЕСКОЕ РЕШЕНИЕ

Будем рассматривать движение одного тела относительно другого под действием их взаимного притяжения в некоторой невращающейся прямоугольной системе координат. Тогда уравнение задачи двух тел (кеплеровского движения) можно представить в виде

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F} = -\mu \frac{\mathbf{x}}{|\mathbf{x}|^3} = \left(\frac{\partial}{\partial \mathbf{x}} \frac{\mu}{|\mathbf{x}|} \right)^T. \quad (\text{П.1})$$

Умножим уравнение справа и слева сначала векторно на \mathbf{x} , а затем скалярно на $\dot{\mathbf{x}}$. После несложных преобразований получаем

$$\begin{aligned} \mathbf{x} \times \frac{d^2\mathbf{x}}{dt^2} &= \frac{d}{dt}(\mathbf{x} \times \dot{\mathbf{x}}) = | = \mathbf{x} \times \mathbf{F} = \mathbf{0}; \\ \dot{\mathbf{x}} \cdot \frac{d^2\mathbf{x}}{dt^2} &= \frac{d}{dt} \left(\frac{\dot{\mathbf{x}}^2}{2} \right) = | = \dot{\mathbf{x}} \cdot \left(\frac{\partial}{\partial \mathbf{x}} \frac{\mu}{|\mathbf{x}|} \right)^T = \frac{d}{dt} \left(\frac{\mu}{|\mathbf{x}|} \right). \end{aligned}$$

Здесь слэш разделяет выражения преобразованных левой и правой частей. Интегрируя полученные уравнения, получаем интеграл площадей или момента количества движения

$$\mathbf{x} \times \dot{\mathbf{x}} = \mathbf{c} = \mathbf{const} \quad (\text{П.2})$$

и интеграл живых сил или энергии

$$\frac{\dot{\mathbf{x}}^2}{2} - \frac{\mu}{|\mathbf{x}|} = h = \mathbf{const}. \quad (\text{П.3})$$

Здесь \mathbf{c} — вектор момента количества движения; h — энергия (за единицу массы). Из (П.2), в частности, следует, что орбитальное движение происходит в некоторой неподвижной плоскости, перпендикулярной вектору \mathbf{c} (рис. 1).

Умножим теперь левую и правую части уравнения (П.1) векторно на \mathbf{c}/μ , учитывая (П.2):

$$\begin{aligned} \frac{d^2\mathbf{x}}{dt^2} \times \frac{\mathbf{c}}{\mu} &= \frac{d}{dt} \left(\frac{\dot{\mathbf{x}} \times \mathbf{c}}{\mu} \right) = \left| = -\frac{\mathbf{x} \times (\mathbf{x} \times \dot{\mathbf{x}})}{|\mathbf{x}|^3} = \right. \\ &= \left. -\frac{\mathbf{x}(\mathbf{x} \cdot \dot{\mathbf{x}}) - \dot{\mathbf{x}}(\mathbf{x} \cdot \mathbf{x})}{|\mathbf{x}|^3} = \frac{d}{dt} \left(\frac{\mathbf{x}}{|\mathbf{x}|} \right). \right. \end{aligned}$$

Интегрируя это уравнение, получим еще один так называемый интеграл Лапласа

$$\frac{\dot{\mathbf{x}} \times \mathbf{c}}{\mu} - \frac{\mathbf{x}}{|\mathbf{x}|} = \mathbf{e} = \mathbf{const}, \quad (\text{П.4})$$

где \mathbf{e} — вектор Лапласа (рис. 2), хотя часто в качестве него рассматривают вектор $\mathbf{g} = \mu\mathbf{e}$.

Из всех семи интегралов (П.2)–(П.4) независимыми являются только пять, поскольку, как нетрудно проверить, они связаны двумя соотношениями

$$\mathbf{c} \cdot \mathbf{e} = 0, \quad 2hc^2 = \mu^2(e^2 - 1). \quad (\text{П.5})$$

Из первого соотношения, в частности, следует, что вектор Лапласа лежит в орбитальной плоскости.

Умножим (П.4) скалярно на \mathbf{x}

$$\frac{c^2}{\mu} - |\mathbf{x}| = |\mathbf{x}|e \cos v, \quad \text{отсюда} \quad |\mathbf{x}| = \frac{p}{1 + e \cos v}, \quad (\text{П.6})$$

где v — угол между векторами \mathbf{x} и \mathbf{e} (рис. 2); $p = c^2/\mu$. Полученное соотношение (П.6) представляет собой уравнение конического сечения (эллипса, параболы или гиперболы) в полярных координатах $|\mathbf{x}|$ (радиус-вектор) и v (истинная аномалия) с фокальным параметром p и эксцентриситетом e . Центральное тело находится в фокусе орбиты. Минимальное расстояние до центрального тела достигается при $v = 0$, т.е. когда вектор положения и Лапласа

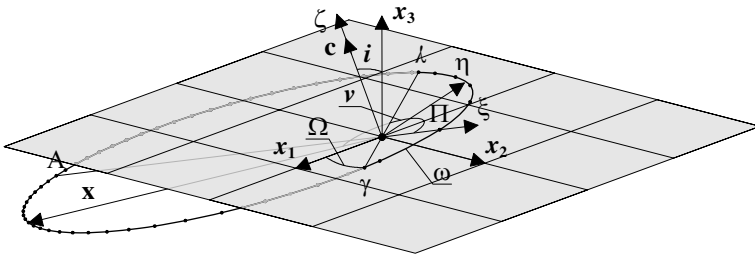


Рис. 1. Орбита в пространстве. Здесь $e = 0.7$; $i = 20^\circ$; $\Omega = 30^\circ$; $\omega = 60^\circ$; А — апоцентр; П — перицентр; γ — восходящий узел. На орбите точками представлены положения объекта через временные интервалы $\Delta t = T/50$. Пунктирной линией показана проекция орбиты на основную плоскость

коллинеарны. Точка минимального расстояния на орбите называется перицентром. Таким образом, если начало вектора Лапласа поместить в фокус орбиты, то он будет направлен на ее перицентр, при этом его величина равна значению эксцентриситета (рис. 2).

Введем орбитальную систему координат с осями

$$\mathbf{S} = \frac{\mathbf{e}}{|\mathbf{e}|}, \quad \mathbf{T} = \mathbf{W} \times \mathbf{S}, \quad \mathbf{W} = \frac{\mathbf{c}}{|\mathbf{c}|}.$$

Тогда соответствующие координаты в этой системе ξ , η и ζ (см. рис. 1 и 2) будут связаны с полярными как

$$\xi = |\mathbf{x}| \cos v, \quad \eta = |\mathbf{x}| \sin v, \quad \zeta = 0.$$

Интеграл момента количества движения (П.2) в орбитальных координатах можно представить в виде

$$\xi \dot{\eta} - \dot{\xi} \eta = c, \quad \text{отсюда} \quad |\mathbf{x}|^2 \dot{v} = c. \quad (\text{П.7})$$

Далее ограничимся рассмотрением эллиптического решения (П.6), т.е. когда $0 \leq e < 1$. Наиболее отдаленная от фокуса точка

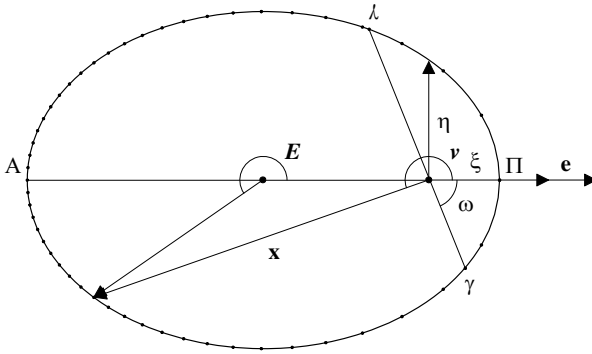


Рис. 2. Та же орбита, что и на рис. 1, но на ее плоскости. Угол E (эксцентрисическая аномалия) измеряется на плоскости круговой орбиты, ортогональной проекцией которой на основную плоскость орбитальной системы координат является эллиптическая орбита

на эллиптической орбите ($v = \pi$) называется апоцентром. В то же время полурастояние a между перицентром и апоцентром называется большой полуосью, а линия, проходящая через них, — линией апсид. Суммируя полурастояния от фокуса (П.6) при $v = 0$ и $v = \pi$, находим

$$p = a(1 - e^2).$$

Тогда, используя второе соотношение (П.5), получим

$$h = -\frac{\mu}{2a}.$$

Эллиптическую орбиту можно представить как ортогональную проекцию на орбитальную плоскость круговой орбиты радиуса a , проходящей через перицентр и апоцентр, т.е. линия апсид является пересечением плоскостей орбит. Угол ε между этими плоскостями выражается через эксцентриситет как $e = \sin \varepsilon$. В плоскости круговой орбиты введем угол E (рис. 2), называемый эксцентрисической аномалией, между линией апсид и направлением

из центра орбиты на точку, которая является прообразом положения на эллиптической орбите. Тогда из геометрических соображений нетрудно представить орбитальные координаты ξ и η как функции эксцентрической аномалии E :

$$\begin{aligned}\xi &= |\mathbf{x}| \cos v = a \cos E - ae, \\ \eta &= |\mathbf{x}| \sin v = a \cos \varepsilon \sin E = a\sqrt{1-e^2} \sin E,\end{aligned}\quad (\text{П.8})$$

отсюда

$$\dot{\xi} = -a\dot{E} \sin E, \quad \dot{\eta} = a\dot{E} \sqrt{1-e^2} \cos E. \quad (\text{П.9})$$

Из (П.8) следует, что

$$|\mathbf{x}| = \sqrt{\xi^2 + \eta^2} = a(1 - e \cos E). \quad (\text{П.10})$$

Дифференцируя (П.10) и пользуясь соотношениями (П.6)–(П.8), получим дифференциальное уравнение для E , а также формулу для \dot{E} в (П.9)

$$\frac{dE}{dt} = \frac{n}{1 - e \cos E} = \frac{na}{|\mathbf{x}|}, \quad (\text{П.11})$$

где $n = \sqrt{\mu/a^3}$ — так называемое среднее движение. Учитывая, что $E = v = 0$ в момент прохождения через перигеум $t = \tau$, находим решение дифференциального уравнения (П.11)

$$E - e \sin E = M = n(t - \tau) = n(t - t_0) + M_0, \quad (\text{П.12})$$

где M — средняя аномалия; M_0 — значение M на момент t_0 . Соотношение (П.12) называется уравнением Кеплера. Оно связывает время t и аномалии M , E и v , которые однозначно задают положение объекта на орбите. Поскольку орбитальное движение 2π -периодично по аномалиям, то орбитальный период \mathcal{T} по времени, согласно (П.12), будет

$$\mathcal{T} = \frac{2\pi}{n}.$$

Мы определили движение в орбитальных координатах. Чтобы перейти к исходным координатам, нужно выполнить преобразование вращения $\mathbf{R} = \mathbf{R}_3(-\Omega)\mathbf{R}_1(-i)\mathbf{R}_3(-\omega)$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{R} \begin{pmatrix} \xi \\ \eta \\ 0 \end{pmatrix}.$$

Здесь Ω , i и ω — эйлеровы углы (см. рис. 1), задающие ориентацию орбиты в пространстве \mathbf{x} : прецессионный, нутационный и собственный углы, которые в небесной механике традиционно называют долгой восходящего узла⁴⁵, наклонением орбиты и аргументом перицентра соответственно, а \mathbf{R}_1 и \mathbf{R}_3 — матрицы вращения относительно осей абсцисс и аппликат:

$$\mathbf{R}_1(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & -\sin \varphi & \cos \varphi \end{pmatrix}, \quad \mathbf{R}_3(\varphi) = \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В итоге имеем

$$\begin{aligned} x_1 &= |\mathbf{x}|(\cos \Omega \cos u - \sin \Omega \sin u \cos i), \\ x_2 &= |\mathbf{x}|(\sin \Omega \cos u + \cos \Omega \sin u \cos i), \\ x_3 &= |\mathbf{x}| \sin u \sin i, \end{aligned}$$

где $u = v + \omega$ — аргумент широты.

В орбитальной системе координат $(0, 0, c)^T$ и $(e, 0, 0)^T$ — векторы момента количества движения и Лапласа. Применяя к ним преобразование вращения \mathbf{R} , получим следующие полезные соотношения:

$$\begin{aligned} c_1 &= c \sin \Omega \sin i, & e_1 &= e(\cos \Omega \cos \omega - \sin \Omega \sin \omega \cos i), \\ c_2 &= -c \cos \Omega \sin i, & e_2 &= e(\sin \Omega \cos \omega + \cos \Omega \sin \omega \cos i), \\ c_3 &= c \cos i; & e_3 &= e \sin \omega \sin i. \end{aligned}$$

⁴⁵Узлы — это точки пересечения орбиты с основной плоскостью исходной системы координат (см. рис. 1). Если при прохождении узла аппликата исследуемого тела возрастает, узел восходящий, иначе — нисходящий. Линия, проходящая через узлы, называется линией узлов.

Они фактически позволяют из динамических переменных \mathbf{x} и $\dot{\mathbf{x}}$ посредством \mathbf{c} и \mathbf{e} получить эйлеровы углы Ω , i и ω .

Впрочем, имеются другие, более удобные для практики формулы, представляющие решение задачи двух тел. Вследствие того, что кеплеровское движение плоское, вектор положения \mathbf{x} можно представить как линейную комбинацию любых двух неколлинеарных векторов, лежащих в орбитальной плоскости. Возьмем в качестве таковых \mathbf{x}_0 и $\dot{\mathbf{x}}_0$, положение и скорость небесного тела в начальный момент времени t_0 . Тогда вектор положения на эллиптической орбите в любой другой момент времени $t = t_0 + \Delta t$ будет (Бэттин, 1966)

$$\mathbf{x} = f\mathbf{x}_0 + g\dot{\mathbf{x}}_0. \quad (\text{П.13})$$

Чтобы определить коэффициенты f и g , векторно умножим (П.13) на $\dot{\mathbf{x}}_0$ и \mathbf{x}_0 :

$$\mathbf{x} \times \dot{\mathbf{x}}_0 = f\mathbf{c}, \quad \mathbf{x}_0 \times \mathbf{x} = g\mathbf{c}.$$

Отсюда, воспользовавшись орбитальными координатами, коэффициенты можно выразить как

$$f = \frac{\xi\dot{\eta}_0 - \dot{\xi}_0\eta}{c}, \quad g = \frac{\xi_0\eta - \xi\eta_0}{c}.$$

Далее, применяя формулы, полученные выше, будем иметь

$$f = 1 - \frac{a}{|\mathbf{x}_0|}(1 - \cos \Delta E), \quad g = \Delta t - \frac{\Delta E - \sin \Delta E}{n},$$

$$a = -\frac{\mu}{2h}, \quad h = \frac{\dot{\mathbf{x}}_0^2}{2} - \frac{\mu}{|\mathbf{x}_0|}.$$

Разность текущей и начальной аномалий $\Delta E = E - E_0$ определяется обобщенным уравнением Кеплера

$$n\Delta t = \Delta E + \frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2}(1 - \cos \Delta E) - \left(1 - \frac{|\mathbf{x}_0|}{a}\right) \sin \Delta E,$$

которое является следствием уравнения Кеплера (П.12). Здесь использованы соотношения

$$e \cos E_0 = 1 - \frac{|\mathbf{x}_0|}{a}, \quad e \sin E_0 = \frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2},$$

получаемые из (П.10) и (П.11).

На практике порой бывают необходимы производные от вектора положения на некоторый произвольный момент времени по начальным параметрам \mathbf{x}_0 и $\dot{\mathbf{x}}_0$. В соответствии с формулами (П.13) производные от \mathbf{x} по $\mathbf{q} = (\mathbf{x}_0, \dot{\mathbf{x}}_0)^T$ можно представить в виде

$$\frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \mathbf{x}_0 \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{x}}_0 \frac{\partial g}{\partial \mathbf{q}} + f \frac{\partial \mathbf{x}_0}{\partial \mathbf{q}} + g \frac{\partial \dot{\mathbf{x}}_0}{\partial \mathbf{q}},$$

где

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{q}} &= -(1 - \cos \Delta E) \frac{\partial}{\partial \mathbf{q}} \left(\frac{a}{|\mathbf{x}_0|} \right) - \frac{a}{|\mathbf{x}_0|} \sin \Delta E \frac{\partial \Delta E}{\partial \mathbf{q}}, \\ \frac{\partial g}{\partial \mathbf{q}} &= \frac{\Delta E - \sin \Delta E}{n^2} \frac{\partial n}{\partial \mathbf{q}} - \frac{1 - \cos \Delta E}{n} \frac{\partial \Delta E}{\partial \mathbf{q}}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \Delta E}{\partial \mathbf{q}} &= \left[\frac{\partial n}{\partial \mathbf{q}} \Delta t - \frac{\partial}{\partial \mathbf{q}} \left(\frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2} \right) (1 - \cos \Delta E) - \right. \\ &\left. - \frac{\partial}{\partial \mathbf{q}} \left(\frac{|\mathbf{x}_0|}{a} \right) \sin \Delta E \right] \left[1 + \frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2} \sin \Delta E - \left(1 - \frac{|\mathbf{x}_0|}{a} \right) \cos \Delta E \right]^{-1}, \end{aligned}$$

$$\frac{\partial n}{\partial \mathbf{q}} = -\frac{3n}{2a} \frac{\partial a}{\partial \mathbf{q}}, \quad \frac{\partial a}{\partial \mathbf{q}} = -\frac{a}{h} \frac{\partial h}{\partial \mathbf{q}}, \quad \frac{\partial h}{\partial \mathbf{x}_0} = n^2 \frac{a^3}{|\mathbf{x}_0|^3} \mathbf{x}_0^T, \quad \frac{\partial h}{\partial \dot{\mathbf{x}}_0} = \dot{\mathbf{x}}_0^T,$$

$$\frac{\partial}{\partial \mathbf{x}_0} \left(\frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2} \right) = \frac{1}{na^2} \left(\dot{\mathbf{x}}_0^T - \frac{1}{2a} \frac{\partial a}{\partial \mathbf{x}_0} \right),$$

$$\frac{\partial}{\partial \dot{\mathbf{x}}_0} \left(\frac{\mathbf{x}_0 \cdot \dot{\mathbf{x}}_0}{na^2} \right) = \frac{1}{na^2} \left(\mathbf{x}_0^T - \frac{1}{2a} \frac{\partial a}{\partial \dot{\mathbf{x}}_0} \right),$$

$$\frac{\partial}{\partial \mathbf{x}_0} \left(\frac{a}{|\mathbf{x}_0|} \right) = \frac{1}{|\mathbf{x}_0|} \left(-\frac{a}{|\mathbf{x}_0|^2} \mathbf{x}_0^T + \frac{\partial a}{\partial \mathbf{x}_0} \right), \quad \frac{\partial}{\partial \dot{\mathbf{x}}_0} \left(\frac{a}{|\mathbf{x}_0|} \right) = \frac{1}{|\mathbf{x}_0|} \frac{\partial a}{\partial \dot{\mathbf{x}}_0},$$

$$\frac{\partial}{\partial \mathbf{x}_0} \left(\frac{|\mathbf{x}_0|}{a} \right) = \frac{1}{a} \left(\frac{\mathbf{x}_0^T}{|\mathbf{x}_0|} - \frac{|\mathbf{x}_0|}{a} \frac{\partial a}{\partial \mathbf{x}_0} \right), \quad \frac{\partial}{\partial \dot{\mathbf{x}}_0} \left(\frac{|\mathbf{x}_0|}{a} \right) = -\frac{|\mathbf{x}_0|}{a^2} \frac{\partial a}{\partial \dot{\mathbf{x}}_0}.$$

Наконец, $\partial \mathbf{x}_0 / \partial \mathbf{q} = (\mathbf{E}, \mathbf{0})$ и $\partial \dot{\mathbf{x}}_0 / \partial \mathbf{q} = (\mathbf{0}, \mathbf{E})$, где \mathbf{E} и $\mathbf{0}$ — единичная и нулевая матрицы размера 3×3 соответственно.

Далее рассмотрим задачу определения кеплеровской орбиты по трем наблюдениям α (прямое восхождение) и δ (склонение) на моменты времени t_{-1} , t_0 и t_1 . Используя метод Лапласа (Escobal, 1965), найдем параметры \mathbf{x}_0 и $\dot{\mathbf{x}}_0$ на средний момент t_0 .

Пусть нам известно движение наблюдателя $\mathbf{x}_O = \mathbf{x}_O(t)$ в пространстве \mathbf{x} . Тогда справедливо соотношение

$$\mathbf{x} = \rho \mathbf{l} + \mathbf{x}_O, \quad (\text{П.14})$$

где $\rho = |\mathbf{x} - \mathbf{x}_O|$ — расстояние от наблюдателя до объекта; \mathbf{l} — единичный вектор:

$$l_1 = \cos \alpha \cos \delta, \quad l_2 = \sin \alpha \cos \delta, \quad l_3 = \sin \delta.$$

Дифференцируя (П.14) дважды, будем иметь

$$\dot{\mathbf{x}} = \dot{\rho} \mathbf{l} + \rho \dot{\mathbf{l}} + \dot{\mathbf{x}}_O, \quad (\text{П.15})$$

$$\ddot{\mathbf{x}} = \ddot{\rho} \mathbf{l} + 2\dot{\rho} \dot{\mathbf{l}} + \rho \ddot{\mathbf{l}} + \ddot{\mathbf{x}}_O. \quad (\text{П.16})$$

Согласно (П.1) и (П.14), соотношение (П.16) можно переписать в виде

$$\mathbf{l} \ddot{\rho} + 2\dot{\mathbf{l}} \dot{\rho} + \left(\ddot{\mathbf{l}} + \frac{\mu}{|\mathbf{x}|^3} \mathbf{l} \right) \rho = -\ddot{\mathbf{x}}_O - \frac{\mu}{|\mathbf{x}|^3} \mathbf{x}_O. \quad (\text{П.17})$$

Если записать систему уравнений (П.17) для момента t_0 , то известными будут величины \mathbf{l}_0 , $(\mathbf{x}_O)_0$ и $(\ddot{\mathbf{x}}_O)_0$, а неизвестными — $\ddot{\rho}_0$, $\dot{\rho}_0$, ρ_0 , $\dot{\mathbf{l}}_0$, $\ddot{\mathbf{l}}_0$ и $|\mathbf{x}|_0$. Впрочем, величины $\dot{\mathbf{l}}_0$ и $\ddot{\mathbf{l}}_0$ можно вычислять приближенно, путем дифференцирования интерполяционного полинома Лагранжа, построенного на значениях \mathbf{l}_{-1} , \mathbf{l}_0 и \mathbf{l}_1 :

$$\dot{\mathbf{l}}_0 = \frac{\mathbf{l}_0 - \mathbf{l}_{-1}}{t_0 - t_{-1}} - \frac{\mathbf{l}_1 - \mathbf{l}_{-1}}{t_1 - t_{-1}} + \frac{\mathbf{l}_1 - \mathbf{l}_0}{t_1 - t_0},$$

$$\ddot{\mathbf{l}}_0 = \frac{2\mathbf{l}_{-1}}{(t_{-1} - t_0)(t_{-1} - t_1)} + \frac{2\mathbf{l}_0}{(t_0 - t_1)(t_0 - t_{-1})} + \frac{2\mathbf{l}_1}{(t_1 - t_{-1})(t_1 - t_0)}.$$

Решая методом Крамера уравнения (П.17) относительно $\ddot{\rho}_0$, $\dot{\rho}_0$, ρ_0 , в частности, получаем

$$\rho_0 = A + \frac{\mu}{|\mathbf{x}|_0^3} B, \quad \dot{\rho}_0 = C + \frac{\mu}{|\mathbf{x}|_0^3} D, \quad (\text{П.18})$$

где

$$A = -\frac{|\mathbf{l}\ddot{\mathbf{x}}_O|_0}{|\mathbf{l}\ddot{\mathbf{l}}|_0}, \quad B = -\frac{|\mathbf{l}\mathbf{x}_O|_0}{|\mathbf{l}\ddot{\mathbf{l}}|_0}, \quad C = -\frac{|\mathbf{l}\ddot{\mathbf{x}}_O\dot{\mathbf{l}}|_0}{2|\mathbf{l}\ddot{\mathbf{l}}|_0}, \quad D = -\frac{|\mathbf{l}\mathbf{x}_O\dot{\mathbf{l}}|_0}{2|\mathbf{l}\ddot{\mathbf{l}}|_0}.$$

Здесь запись типа $|\mathbf{abc}|$ означает определитель матрицы, составленной из векторов \mathbf{a} , \mathbf{b} и \mathbf{c} .

В соответствии с (П.14) имеем

$$|\mathbf{x}|_0^2 = \rho_0^2 + 2\rho_0\mathbf{l}_0 \cdot (\mathbf{x}_O)_0 + |\mathbf{x}_O|_0^2. \quad (\text{П.19})$$

Подставив первое соотношение (П.18) в (П.19), получим уравнение для радиус-вектора $|\mathbf{x}|_0$ (уравнение Лагранжа):

$$|\mathbf{x}|_0^8 - (2AG + A^2 + |\mathbf{x}_O|_0^2)|\mathbf{x}|_0^6 - 2\mu B(G + A)|\mathbf{x}|_0^3 - \mu^2 B^2 = 0, \quad (\text{П.20})$$

где $G = \mathbf{l}_0 \cdot (\mathbf{x}_O)_0$.

Численно решая это уравнение, находим $|\mathbf{x}|_0$. Далее, подставив соотношение (П.18) с вычисленным радиус-вектором в (П.14), получим начальный вектор положения \mathbf{x}_0 . В свою очередь, скорость $\dot{\mathbf{x}}_0$ находится из соотношения (П.15) после подстановки в него соотношений (П.18).

Следует отметить, что уравнение Лагранжа имеет четыре вещественных корня. Если наблюдатель движется под действием гравитации центрального тела, т.е. $\ddot{\mathbf{x}}_O = -\mu\mathbf{x}_O/|\mathbf{x}_O|^3$, то один корень будет соответствовать расстоянию до наблюдателя и естественно он не будет иметь отношение к наблюдаемому телу. Второй корень также может быть отброшен, поскольку он отрицательный. Два оставшихся положительных корня могут совпадать

(кратный корень). Однако, если они отличаются, то в действительности оказывается невозможным определить, какой из них соответствует реальной орбите. Согласно (Charlier, 1910), два решения будут иметь место при условии, если

$$|\mathbf{x}_O|_0^2 + 3AG > 0.$$

Область двойных решений в физическом пространстве показана на рис. 5.12.

ЛИТЕРАТУРА

- Авдюшев В.А.* Численное моделирование движения галилеевых спутников Юпитера // *Астрон. вест.* 1999. Т. 33. Вып. 3. С. 332–337.
- Авдюшев В.А.* Метод синхронного слежения // *Материалы все-росс. конф. «Фундаментальные и прикладные проблемы современной механики».* Томск, 6–8 июня 2000 г. Томск: Изд-во Том. ун-та, 2000. С. 117–118.
- Авдюшев В.А.* Определение оптимального стабилизирующего параметра в методе Баумгарта для моделирования почти круговых орбит // *Изв. вузов. Физика.* 2003а. № 12. Приложение. С. 5–12.
- Авдюшев В.А.* Метод синхронного слежения в ограниченной задаче трех тел // *Изв. вузов. Физика.* 2003б. № 12. Приложение. С. 13–15.
- Авдюшев В.А.* Новая система начальных параметров для численного моделирования движения галилеевых спутников Юпитера // *Астрон. вест.* 2004. Т. 38. Вып. 3. С. 273–276.
- Авдюшев В.А.* О численном интегрировании орбит с короткопериодическими возмущениями // *Изв. вузов. Физика.* 2006а. № 2. Приложение. С. 31–43.
- Авдюшев В.А.* Методы теории специальных возмущений. I. Принципы построения и обоснование к применению // *Изв. вузов. Физика.* 2006б. № 12. С. 73–80.
- Авдюшев В.А.* Методы теории специальных возмущений. II. Сравнительный анализ численной эффективности // *Изв. вузов. Физика.* 2007. № 1. С. 78–86.
- Авдюшев В.А.* Численное моделирование орбит. Томск: Изд-во НТЛ, 2010а. 284 с.
- Авдюшев В.А.* Интегратор Гаусса–Эверхарта // *Вычисл. технологии.* 2010б. Т. 15. № 4. С. 31–47.
- Авдюшев В.А.* Линейные отображения в исследовании локальной орбитальной динамики // *Изв. вузов. Физика.* 2012. № 10/2. С. 5–10.

- Авдюшев В.А., Баныщикова М.А.* Определение орбит близких спутников Юпитера // *Астрон. вест.* 2008. Т. 42. Вып. 4. С. 317–340.
- Авдюшев В.А., Баныщикова М.А.* Альтернативные орбиты новых спутников Юпитера // *Изв. вузов. Физика.* 2010. № 10. С. 27–30.
- Авдюшев В.А., Бордовицына Т.В.* Развитие теории преобразований Шперлингга–Боде // *Изв. вузов. Физика.* 2006. № 2. Приложение. С. 27–30.
- Авдюшев В.А., Галушина Т.Ю.* Быстрое численное оценивание вероятности столкновения астероида с планетой // *Астрон. вест.* 2014. Т. 48. Вып. 4. С. 309–317.
- Аксенов Е.П.* Теория движения искусственных спутников Земли. М.: Наука, 1977. 360 с.
- Антонов В.А., Тимошкова Е.И., Холшевников К.В.* Введение в теорию ньютоновского потенциала. М.: Наука, 1988. 270 с.
- Аттетков А.В., Галкин С.В., Зарубин В.С.* Методы оптимизации. М.: Изд-во МГТУ, 2001. 439 с.
- Батраков Ю.В., Макарова Е.Н.* Обобщенный метод Энке для изучения возмущенного движения // *Бюлл. ИТА АН СССР.* Л.: Наука, 1979. Т. 14. С. 397–401.
- Батраков Ю.В.* Промежуточные орбиты, аппроксимирующие начальный участок возмущенного движения // *Бюлл. ИТА АН СССР.* Л.: Наука, 1981. Т. 15. С. 1–5.
- Бертсекас Д.* Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987. 400 с.
- Бордовицына Т.В.* Современные численные методы в задачах небесной механики. М.: Наука, 1984. 136 с.
- Бордовицына Т.В., Авдюшев В.А., Титаренко В.П.* Численное моделирование общей задачи трех тел // *Исследование по баллистике и смежным вопросам механики.* Томск: Изд-во Том. ун-та, 1998а. Вып. 2. С. 164–168.
- Бордовицына Т.В., Авдюшев В.А.* Теория движения искусственных спутников Земли. Томск: Изд-во Том. ун-та, 2007. 220 с.

- Бордовицына Т.В., Быкова Л.Е.* Теория движения и эфемериды VI и VII спутников Юпитера на 1979–2000 годы. Томск: Изд-во Том. ун-та, 1978. 120 с.
- Бордовицына Т.В., Быкова Л.Е., Авдюшев В.А.* Проблемы применения регуляризирующих и стабилизирующих преобразований в задачах динамики спутников планет и астероидов // Астрон. и геод. Томск: Изд-во Том. ун-та, 1998b. Вып. 16. С. 33–57.
- Бордовицына Т.В., Галушина Т.Ю., Авдюшев В.А.* Стабилизирующие и регуляризирующие преобразования в задаче численного моделирования движения особых астероидов // Изв. вузов. Физика. 2003. № 12. Приложение. С. 23–34.
- Бордовицына Т.В., Сухоплюева Л.Е.* Исследование эффективности численных алгоритмов, использующих стабилизирующие преобразования // Бюлл. инст. теор. астрон. Л.: Изд-во ИТА АН СССР, 1980. Т. 10. Вып. 163. С. 591–596.
- Бордовицына Т.В., Шарковский Н.А.* Эффективные алгоритмы численного моделирования движения Фобоса, спутника Марса // Изв. вузов. Физика. 1994. № 10. С. 8–12.
- Брумберг В.А.* Релятивистская небесная механика. М.: Наука, 1972. 382 с.
- Бэттин Р.* Наведение в космосе. М.: Машиностроение, 1966. 448 с.
- Витязев В.В.* Анализ неравномерных временных рядов: учеб. пособие. СПб.: Изд-во СПбГУ, 2001. 68 с.
- Гельфанд И.М., Цетлин М.Л.* О некоторых способах управления сложными системами // УМН. 1962. Т. 17. Вып. 1. С. 3–25.
- Гилл Ф., Мюррей У., Райт М.* Практическая оптимизация. М.: Мир, 1985. 509 с.
- Демиденко Е.З.* Оптимизация и регрессия. М.: Наука, 1989. 296 с.
- Деккер К., Вервер Я.* Устойчивость методов Рунге–Кутты для жестких нелинейных дифференциальных уравнений. М.: Мир, 1988. 334 с.
- Дубошин Г.Н.* Теория притяжения. М.: ИФМЛ, 1961. 286 с.
- Дубошин Г.Н.* Небесная механика. Аналитические и качественные методы. М.: Наука, 1964. 560 с.

- Кислик М.Д.* Сферы влияния больших планет и Луны // Космич. исслед. 1964. Т. 2. Вып. 6. С. 853–858.
- Ланс Дж.Н.* Численные методы для быстродействующих вычислительных машин. М.: ИЛ, 1962. 208 с.
- Мудров В.И., Кушко В.Л.* Метод наименьших модулей. М.: Знание, 1971. 64 с.
- Нумеров Б.В.* Новый метод определения орбит и вычисления эфемерид с учетом возмущений // Тр. АО Петрогр. ун-та. 1923. Т. 4. С. 29–39.
- Орлов В.В., Рубинов А.В.* Задача N тел в звездной динамике. СПб.: Изд-во СПбГУ, 2008. 175 с.
- Полещиков С.М., Холопов А.А.* Теория L -матриц и регуляризация уравнений движения в небесной механике. Сыктывкар: Изд-во СЛИ, 1999. 255 с.
- Рой А.Е.* Движение по орбитам. М.: Мир, 1981. 544 с.
- Себекей В.* Теория орбит. М.: Наука, 1982. 656 с.
- Сорокин Н.А.* Дифференциальные уравнения движения ИСЗ в задаче двух неподвижных центров и их численное интегрирование // Научн. информации. М.: Изд-во ИА АН СССР, 1991. Т. 69. С. 114–123.
- Субботин М.Ф.* Введение в теоретическую астрономию. М.: Наука, 1968. 800 с.
- Сюсина О.М., Черницов А.М., Тамаров В.А.* Новые алгоритмы построения методом Монте-Карло начальных доверительных областей движения малых тел // Изв. вузов. Физика. 2009. № 10/2. С. 48–55.
- Херрик С.* Астродинамика. М.: Мир, 1977. Т. 2. 263 с.
- Херрик С.* Астродинамика. М.: Мир, 1978. Т. 3. 360 с.
- Холл Дж., Уатт Дж.* Современные численные методы решения обыкновенных дифференциальных уравнений. М.: Мир, 1979. 312 с.
- Холшевников К.В., Титов В.Б.* Задача двух тел. СПб.: СПбГУ, 2007. 180 с.

- Черницов А. М.* Анализ некоторых упрощенных схем определения оценок параметров движения небесных тел // Астрон. и геод. Томск: Изд-во Том. ун-та, 1975. Вып. 5. С. 6–19.
- Черницов А. М., Дубас О. М., Тамаров В. А.* Способы уменьшения нелинейности задачи наименьших квадратов при построении областей возможных движений астероидов // Изв. вузов. Физика. 2006. № 2. Приложение. С. 44–51.
- Черницов А. М., Тамаров В. А., Дубас О. М.* Влияние ошибок в задании весовых матриц на точность определения доверительных областей движения астероидов // Изв. вузов. Физика. 2007. № 12/2. С. 52–59.
- Шапорев С. Д.* Оценка корреляционной функции ошибок наблюдений статистическими методами бутстрепа и складного ножа // Препринт ИТА РАН. СПб.: ИТА РАН, 1996. № 62. 25 с.
- Шефер В. А.* Алгоритм численного исследования движения особых малых планет, основанный на двойной регуляризации уравнений движения // Астрон. и геод. Томск: Изд-во Том. ун-та, 1980. Вып. 8. С. 81–91.
- Шефер В. А.* Влияние временных преобразований на эффективность численного интегрирования регуляризованных уравнений движения // Анализ движения тел Солнечной системы и их наблюдения / Отв. ред. Л. Лауцениекс. Рига: ЛГУ им. П. Стучки, 1986. С. 103–125.
- Шефер В. А.* Линеаризация и регуляризация уравнений кеплеровского движения с помощью интегралов // Астрон. журн. 1991. Т. 68. С. 197–205.
- Шефер В. А.* Обобщенные методы Энке для исследования возмущенного движения // Астрон. и геод. Томск: Изд-во Том. ун-та, 1998. Вып. 16. С. 149–171.
- Шефер В. А.* Новый метод определения промежуточной орбиты по трем положениям малого тела на небесной сфере // Астрон. вест. 2010. Т. 44. Вып. 2. С. 167–181.
- Aarseth S. J.* Gravitational N-Body Simulations. Cambridge University Press, 2003. 430 p.

- Aarseth S.J., Zare K.* A Regularization of the Three-Body Problem // *Celest. Mech.* 1974. Vol. 10. P. 185–205.
- Anderson T.W.* An Introduction to Multivariate Statistical Analysis. John Wiley & Sons Inc., 1958. 374 p.
- Arlot J.E., Emelyanov N.V.* The NSDB Natural Satellites Astrometric Database // *Astron. Astrophys.* 2009. Vol. 503. P. 631–638.
- Ascher U.M., Chin H., Petzold L., Reich S.* Stabilization of Constrained Mechanical Systems with DAEs and Invariant Manifolds // *J. Mech. Struct. Machines.* 1995. Vol. 23. P. 135–138.
- Avdyushev V.* Numerical Stabilization of Orbital Motion // *Celest. Mech.* 2003. Vol. 87. I. 4. P. 383–409.
- Avdyushev V.A.* A New Method for the Statistical Simulation of the Virtual Values of Parameters in Inverse Orbital Dynamics Problems // *Sol. Syst. Res.* 2009. Vol. 43. No. 6. P. 543–551.
- Avdyushev V.A.* Nonlinear Methods of Statistic Simulation of Virtual Parameter Values for Investigating Uncertainties in Orbits Determined from Observations // *Celest. Mech.* 2011. Vol. 110. No. 4. P. 369–388.
- Avdyushev V.A., Banschikova M.A.* Regions of Possible Motions for New Jovian Satellites // *Sol. Syst. Res.* 2007. Vol. 41. No. 5. P. 413–419.
- Avdyushev V., Bordovitsyna T.* Algorithms of Numerical Simulation of the Motion of Satellites // *Proc. of US/European Celestial Mechanics Workshop. Poznan, 3–7 July 2000.* Dordrecht: Kluwer Academic Publishers, 2000. P. 223–224.
- Bard Y.* Nonlinear Parameter Estimation. New York: Academic Press, 1974.
- Bates D.M., Watts D.G.* Relative Curvature Measures of Nonlinearity // *J. R. Statist. Soc.* 1980. Vol. 42. No. 1. P. 1–25.
- Bates D.M., Watts D.G.* Nonlinear Regression Analysis and Its Applications. John Wiley & Sons Inc., 1988. 365 p.
- Baumgarte J.* Stabilization of Constraints and Integrals of Motion in Dynamical Systems // *Comp. Math. Appl. Mech. Eng.* 1972. Vol. 1. P. 1–16.

- Baumgarte J.* Numerical Stabilization of the Differential Equations of Keplerian Motion // *Celest. Mech.* 1972. Vol. 5. I. 4. P. 490–501.
- Baumgarte J.* Stabilized Kepler Motion Connected with Analytic Step Adaptation // *Celest. Mech.* 1976a. Vol. 13. P. 105–109.
- Baumgarte J.* Stabilization, Manipulation, and Analytic Step Adaptation // *Long-Time Predictions in Dynamics; Proc. of the Advanced Study Institute, Cortina d'Ampezzo, Italy, August 3–16, 1975* / ed. by V. Szebehely and B.D. Tapley. D. Reidel Publishing Co., 1976b. P. 153–163.
- Baumgarte J., Stiefel E.L.* Examples of Transformations Improving the Numerical Accuracy of the Integration of Differential Equations // *Lect. Notes in Math.* 1974. Vol. 362. P. 207–236.
- Beale E.M.L.* Confidence Regions in Non-Linear Estimation // *J. R. Statist. Soc.* 1960. Vol. 22. No. 1. P. 41–88.
- Bond G.P.* On Some Application of the Method of Mechanical Quadratures // *Memoirs of the American Academy of Arts and Sciences.* 1849. Vol. 4. No. 1. P. 189–208.
- Bordovitsyna T., Avdyushev V., Chernitsov A.* New Trends in Numerical Simulation of the Motion of Small Bodies of the Solar System // *Celest. Mech.* 2001. Vol. 80. I. 3. P. 227–247.
- Brumberg V.A., Brumberg E.V.* *Celestial Dynamics at High Eccentricities.* Amsterdam: Gordon & Breach, 1999. 220 p.
- Brumberg V.A.* *Analytical Techniques of Celestial Mechanics.* Springer, 1995. 236 p.
- Burdet C.A.* Theory of Kepler Motion: the General Perturbed Two Body Problem // *Z. Angew. Math. Phys.* 1968. I. 19. P. 345–368.
- Butcher J.C.* Implicit Runge–Kutta Processes // *Math. Comput.* 1964. Vol. 18. P. 50–64.
- Charlier C.V.L.* On Multiple Solutions in the Determination of Orbits from Three Observations // *Mon. Notic. Roy. Astron. Soc.* 1910. Vol. 71. P. 120–124.
- Chernitsov A.M., Baturin A.P., Tamarov V.A.* Analysis of Some Methods of Determining Probabilistic Evolutions of the Motion

- of Solar System Small Bodies // Sol. Syst. Res. 1998. Vol. 32. No. 5. P. 459–467.
- Chin H.* Stabilization Methods for Simulations of Constrained Multi-body Dynamics // PhD thesis. Institute of Applied Mathematics, University of British Columbia, Canada, 1995. 179 p.
- Conte S.P.* The Computation of Satellite Orbit Trajectories // Adv. in Comp. New York: Academic Press Inc., 1962. Vol. 3. P. 1–76.
- Dahlquist G.* Convergence and Stability in the Numerical Integration of Ordinary Differential Equations // Math. Scand. 1956. Vol. 4. P. 33–53.
- Desmars J., Arlot S., Arlot J.-E., Lainey V., Vienne A.* Estimating the Accuracy of Satellite Ephemerides Using the Bootstrap Method // Astron. Astrophys. 2009. Vol. 62. P. 321–330.
- Draper N.R., Smith H.* Applied Regression Analysis. John Wiley & Sons Ltd., 1981. 709 p.
- Droźyner A.* Recurrent Calculation of Gravitational Acceleration of a Satellite // Acta Astron. 1977. Vol. 27. P. 15–22.
- Efron B.* Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics. 1979. Vol. 7. P. 1–26.
- Emel'yanov N.* Precision of the Ephemerides of Outer Planetary Satellites // PSS. 2010. Vol. 58. I. 3. P. 411–420.
- Emel'yanov N.V., Arlot J.-E.* The Natural Satellites Ephemerides Facility MULTI-SAT // Astron. Astrophys. 2008. Vol. 487. P. 759–765.
- Encke J.F.* Über eine neue Methode der Berechnung der Planetenstörungen // Astr. Nach. 1852. Vol. 33. P. 377–398.
- Escobal P.* Methods of Orbit Determination. New York–London–Sydney: John Wiley and Sons, Inc., 1965. 463 p.
- Everhart E.* A New Method for Integrating Orbits // Bulletin of the American Astronomical Society. 1973. Vol. 5. P. 389.
- Everhart E.* Implicit Single Sequence Methods for Integrating Orbits // Celest. Mech. 1974. Vol. 10. P. 35–55.
- Everhart E.* An Efficient Integrator That Uses Gauss–Radau Spacings // Dynamics of Comets: Their Origin and Evolution, Proceedings of IAU Colloq. 83, held in Rome, Italy, June 11–15, 1984 /

- ed. by A. Carusi and G. B. Valsecchi. Dordrecht: Reidel, Astrophysics and Space Science Library, 1985. Vol. 115. P. 185–202.
- Ferrándiz J.M., Ferrer S., Sein-Echaluce M.L.* Generalized Elliptic Anomalies // *Celest. Mech.* 1987. Vol. 40. I. 3–4. P. 315–328.
- Fukushima T.* A Fast Procedure Solving Kepler's Equation for Elliptic Case // *Astron. J.* 1996. Vol. 112. I. 6. P. 2858–2861.
- Fukushima T.* Efficient Orbit Integration by Scaling for Kepler Energy Consistency // *Astron. J.* 2003. Vol. 126. P. 1097–1111.
- Gragg W.B.* Repeated Extrapolation to the Limit in the Numerical Solution of Ordinary Differential Equations: thesis ... Univ. California, 1964.
- Gragg W.B.* On Extrapolation Algorithms for Ordinary Initial Value Problems // *SIAM J. Numer. Anal.* 1965. Vol. 2. No. 3. P. 384–403.
- Guillou A., Soulé J.L.* La résolution numérique des problèmes différentiels aux conditions initiales par des méthodes de collocation // *Rev. Française Informat. Recherche Opérationnelle.* 1969. Vol. 3. Ser. R-3. P. 17–44.
- Hairer E., Lubich C., Wanner G.* Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations. Springer, 2002. 659 p.
- Hairer E., Norsett S.P., Wanner G.* Solving Ordinary Differential Equations. Nonstiff Problems. Springer, 1993. 539 p.
- Hartley H.O.* Exact Confidence Regions for the Parameters in Non-Linear Regression Laws // *Biometrika.* 1964. Vol. 51. No. 3/4. P. 347–353.
- Heggie D.C.* A Global Regularization of the Gravitational N-body Problem // *Celest. Mech.* 1974. Vol. 10. P. 217–242.
- Heun K.* Neue Methoden zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen // *Z. Math. Phys.* 1900. Vol. 45. P. 23–38.
- Herrick S.* Jupiter IX and Jupiter XII // *Publications of the Astronomical Society of the Pacific.* 1952. Vol. 64. No. 380. P. 237–241.

- Himmelblau D.M.* Applied Nonlinear Programming. Austin: the University of Texas, 1972.
- Jacobson R.A.* The Orbits of the Outer Jovian Satellites // *Astron. J.* 2000. Vol. 120. P. 2679–2686.
- Kahan W., Li R.-C.* Composition Constants for Raising the Orders of Unconventional Schemes for Ordinary Differential Equations // *Math. Comput.* 1997. Vol. 66. P. 108–1099.
- Kulyk I., Jockers K., Karpov N., Sergeev A.* Astrometric CCD Observations of the Inner Jovian Satellites in 1999–2000 // *Astron. Astrophys.* 2002. Vol. 383. P. 724–728.
- Kuntzmann J.* Neuere Entwicklungen der Methode von Runge und Kutta // *Z. Angew. Math. Mech.* 1961. Vol. 41. P. 28–31.
- Kustaanheimo P., Stiefel E.* Perturbation Theory of Kepler Motion Based on Spinor Regularization // *J. Reine Angew. Math.* 1965. Vol. 218. P. 204–219.
- Kutta W.* Beitrag zur näherungsweise Integration totaler Differentialgleichungen // *Z. Math. Phys.* 1901. Vol. 46. P. 435–453.
- Kyner W.T., Bennett M.M.* Modified Encke Special Perturbation Method // *Astron. J.* 1966. Vol. 71. P. 579–582.
- Lagrange J.L.* Sur le Problème de la détermination des orbites des comètes d'après trois observations. Troisième Mémoire, dans lequel on donne une solution directe et générale du problème. // *Nouveaux mémoires de l'Académie royale des sciences et belles-lettres.* 1783. P. 296–332.
- Lainey V., Duriez L., Vienne A.* New Accurate Ephemerides for the Galilean Satellites of Jupiter. I. Numerical Integration of Elaborated Equations of Motion // *Astron. Astrophys.* 2004. Vol. 420. P. 1171–1183.
- Lasagni F.M.* Canonical Runge–Kutta Methods // *ZAMP.* 1988. Vol. 39. P. 952–953.
- Lekien F., Ross S.D.* The Computation of Finite-Time Lyapunov Exponents on Unstructured Meshes and for Non-Euclidean Manifolds // *Chaos.* 2010. Vol. 20. P. 017505.
- Levenberg K.* A Method for the Solution of Certain Problems in Least Squares // *Quart. Appl. Math.* 1944. Vol. 2. P. 164–168.

- Levi-Civita T.* Traiettorie singolari ed urti nel problema ristretto dei tre corpi // Ann. Mat. Pura Appl. 1903. Vol. 9. P. 1–32.
- Lieske J.H.* Galilean Satellites Ephemerides E-5 // Astron. Astrophys. 1998. Vol. 129. P. 205–217.
- Lovett E.O.* The Theory of Perturbations and Lie's Theory of Contact Transformations // The Quart. J. of Pure and Appl. Math. 1899. Vol. 30. P. 47–149.
- Marsden B.G., Sekanina Z., Yeomans D.* Comets and Nongravitational Forces. V. // Astron. J. 1973. Vol. 78. P. 211–225.
- Marquardt D.* An Algorithm for Least-Squares Estimation of Non-linear Parameters // SIAM J. Appl. Math. 1963. Vol. 11. P. 431–441.
- Maxwell J.C.* A Treatise on Electricity and Magnetism. Oxford, 1873. Vol. 2. 518 p.
- Mikkola S.* A Practical and Regular Formulation of the N-body Equations // Mon. Notic. Roy. Astron. Soc. 1985. Vol. 215. P. 171–177.
- Mikkola S., Aarseth S.J.* A Chain Regularization Method for the Few-Body Problem // Celest. Mech. 1990. Vol. 47. P. 375–390.
- Milne W.E.* Numerical Integration of Ordinary Differential Equations // Amer. Math. Monthly. Vol. 33. No. 9. P. 455–460.
- Montenbruck O., Gill E.* Satellite Orbits: Models, Methods and Applications. Springer, 2000. 369 p.
- Muinsonen K., Bowell E.* Asteroid Orbital Error Analysis: Theory and Application // Asteroids, Comets, Meteors. Houston: Lunar and Planetary Inst., 1992. P. 429–432.
- Muinsonen K., Virtanen J., Granvik M., Laakso T.* Asteroid Orbits Using Phase-Space Volumes of Variation // Mon. Not. R. Astron. Soc. 2006. Vol. 368. P. 809–818.
- Murison M.A.* On an Efficient and Accurate Method to Integrate Restricted Three-Body Orbits // Astron. J. 1989. Vol. 97. I. 5. P. 1496–1509.
- Murray C.D., Dermott S.F.* Solar System Dynamics. Cambridge University Press, 1999. 592 p.

- Nacozy P.E.* The Use of Integrals in Numerical Integrations of the N-Body Problem // *Astrophys. Space Sci.* 1971. Vol. 14. P. 40–51.
- Nicholson P.D., Matthews K.* Near-infrared Observations of the Jovian Ring and Small Satellites // *Icarus.* 1991. Vol. 93. P. 331–346.
- Ortega J., Reinboldt W.* Iterative Solution of Nonlinear Equations in Several Variables. Philadelphia: Academic Press, 2000. 572 .
- Parker G.E., Sochacki J.S.* Implementing the Picard iteration // *Neural, Parallel, Sci. Comput.* 1996. Vol. 4. P. 97–112.
- Pines S.* Variation of Parameters for Elliptic and Near Circular Orbits // *Astron. J.* 1961. Vol. 66. No. 1. P. 5–7.
- Poynting J.H.* Radiation in the Solar System: its Effect on Temperature and its Pressure on Small Bodies // *Phil. Trans. Roy. Soc.* 1903. Vol. 202. P. 525–552.
- Quinlan G.D., Tremaine S.* Symmetric Multistep Methods for the Numerical Integration of Planetary Orbits // *Astron. J.* 1990. Vol. 100. P. 1694–1700.
- Robertson H.P.* Dynamical Effects of Radiation in the Solar System // *Mon. Not. Roy. Astron. Soc.* 1937. Vol. 97. P. 423–438.
- Rocher P., Chapront J.* Observations and Ephemerides of the Faint Satellites of Jupiter // *Astron. Astrophys.* 1996. Vol. 311. P. 710–714.
- Ross S.D., Tallapragada P.* Detecting and Exploiting Chaotic Transport in Mechanical Systems // *Applications of Chaos and Nonlinear Dynamics in Science and Engineering* / ed. by S. Banerjee et al. 2012. Vol. 2. P. 155–183.
- Runge C.* Über die numerische Auflösung von Differentialgleichungen // *Math. Ann.* 1895. Vol. 46. P. 167–178.
- Sanz-Serna J.M.* Runge–Kutta Schemes for Hamiltonian Systems // *BIT.* 1988. Vol. 28. P. 877–883.
- Shaikh N.A.* A New Perturbation Method for Computing Earth–Moon Trajectories // *Astronaut. Acta.* 1966. Vol. 12. P. 207–211.

- Silver M.* A Short Derivation of the Sperlring–Burdet Equations // *Celest. Mech.* 1975. Vol. 11. P. 39–41.
- Sperlring H.J.* The Collision Singularity in a Perturbed Two-Body Problem // *Celest. Mech.* 1969. Vol. 1. P. 213–221.
- Standish E.M.* JPL Planetary and Lunar Ephemerides, DE405/LE405 // *Interof. Memor.* 1998. Vol. 312. F-98-048. P. 1–18.
- Steffensen J.F.* On the Restricted Problem of Three Bodies // *K. Danske Vidensk. Selsk. Mat.-Fys. Meddr.* 1956. Vol. 30. No. 18. 17 p.
- Stiefel E., Rössler M., Waldvogel J., Burdet C.A.* Methods of Regularization for Computing Orbits in Celestial Mechanics. Washington, 1967. 124 p.
- Stiefel E.L., Scheifele G.* Linear and Regular Celestial Mechanics. Springer-Verlag, 1971. 301 p.
- Sudbury P.V.* The Motion of Jupiter’s Fifth Satellite // *Icarus.* 1969. Vol. 10. P. 116–143.
- Sundman K.* Mémoire sur le problème des trois corps // *Acta Math.* 1912. Vol. 36. P. 105–179.
- Szebehely V.* Linearization of Dynamical Systems Using Integrals of the Motion // *Celest. Mech.* 1976. Vol. 14. P. 499–508.
- Turner R.* Inverse of the Vandermonde Matrix with Applications // *NASA Technical Note.* Washington: NASA, 1966. NASA TN D-3547.
- Virtanen J., Muinonen K., Bowell E.* Statistical Ranging of Asteroid Orbits // *Icarus.* 2001. Vol. 154. P. 412–431.
- Williams I.P., Ryabova G.O., Baturin A.P., Chernitsov A.M.* Are asteroid 2003 EH1 and comet C/1490 Y1 dynamically related? // *Earth, Moon, and Planets.* 2005. Vol. 95. P. 11–18.
- Wisdom J., Holman M.* Symplectic Maps for the N-body Problem — Stability Analysis // *Astron. J.* 1992. Vol. 104. P. 2022–2029.
- Yoshida H.* Construction of Higher Order Symplectic Integrators // *Phys. Lett. A.* 1990. Vol. 150. P. 262–268.
- Yoshida H.* Recent Progress in the Theory and Application of Symplectic Integrators // *Celest. Mech.* 1993. Vol. 56. P. 27–43.

Zhu J., Wu X. A Comparison between the Stabilization Method and the Post-stabilization Method // *Chin. Astron. and Astrophys.* 2007. Vol. 31. P. 45–56.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- алиасинг, 239
- аномалия
 - истинная, 49, 304
 - средняя, 307
 - эксцентрическая, 49, 306
 - эллиптическая, 49, 102
- апоцентр, 306
- аргумент
 - перигея, 308
 - широты, 308
- астероид
 - Фаэтон, 114
 - Церера, 114
 - троянский, 56
- астрономическая единица, 30
- большая полуось, 306
- вектор
 - Лапласа, 304
 - динамического состояния, 17, 23, 232, 278, 297
 - момента количества движения, 303
 - положения, 17
- возможные значения параметров, 258, 267
- возмущение, 71–73, 98, 125, 270
 - вековое, 36, 109
 - короткопериодическое, 100
 - стабилизирующее, 61
- временной элемент, 43
- время, 17
 - всемирное, 217
 - звездное, 216
 - среднее солнечное, 216
 - фиктивное, 41, 48, 76, 92
 - эфемеридное, 217
- гамильтониан, 185
 - модифицированный, 202
 - разделенный, 198
- гармоника
 - зональная, 27, 78
 - секториальная, 27
 - тессеральная, 27
- гармонический осциллятор, 46, 76, 162, 185
- гауссово кольцо, 106
- главный член погрешности, 128
- гравитационный потенциал, 26
- дисперсия ошибок, 214, 259
- дифференциально-алгебраические уравнения, 130, 189
- дифференциальное уравнение
 - второго порядка, 126
 - первого порядка, 126
- дифференциальные уравнения
 - в вариациях, 221, 295
 - канонические, 185
- дифференциальные уравнения движения, 38
- Баумгарте, 59, 64

- Кустаанхеймо–Штифеля,
43
Роя, 68
Шперлинга–Боде, 43
Энке, 71
 в KS-переменных, 75
 классические, 73
в KS-элементах, 69
в SB-элементах, 69
в элементах, 66
в элементах Пайса, 70
жесткие, 63
классические, 17, 38
регулярные, 39
сглаженные, 48
синхронного слежения, 91
стабилизированные, 53
усредненные, 106
доверительная область, 262–
 264, 295
долгота, 26, 216
 восходящего узла, 308
 истинная, 68
 средняя, 68
дуга орбиты, 49
задача
 Гюльдена–Мещерского, 77
 Коши, 18, 90, 127, 152, 171
 двух неподвижных центров,
 78, 96
 двух тел, 80, 162
 круговая, 240
 многих тел, 96
 наименьших квадратов, 213,
 214, 259
 наименьших модулей, 215
 обратная, 18, 213
 овражная, 224, 227, 243
 прямая, 18
 трех тел (ограниченная),
 84, 96, 100
законы Ньютона, 18
изохронные производные, 220,
 294, 310
интеграл
 Лапласа, 43, 304
 канонической системы, 185
 площадей, 303
 эллиптический, 107
 энергии, 43, 303
интегратор Гаусса–Эверхарта,
 112, 152
интегрирование
 аналитическое, 124
 пошаговое, 128
 численное, 124
интерполяция
 Лагранжа, 147, 148, 159,
 172, 174, 311
 Ньютона, 152, 181
 Эйткена–Невилла, 168
 каноническая, 152
квадратура
 Гаусса, 150
 Ньютона–Котеса, 147
квантиль, 262, 263

- корректор, 175
 коэффициент гармонический, 26
 кривизна подпространства оценок, 266
 линеаризация, 294
 линия
 апсид, 306
 узлов, 308
 матрица
 Вандермонда, 157
 Гессе, 225
 Кустаанхеймо–Штифеля, 44
 Леви-Чивиты, 44
 Холецкого, 214, 264, 268
 весовая, 214, 272, 275
 вращения, 308
 идемпотентная, 265
 ковариационная, 260, 272, 274, 295
 нормальная, 225, 227
 перехода, 297
 машинный эpsilon, 81, 162
 метод
 Адамса, 173
 Адамса–Башфорта, 171
 Адамса–Мультона, 173, 205
 Гаусса, 150, 151, 201, 217
 Гаусса–Ньютона, 223, 225, 248, 260, 265
 демпфированный, 226, 233, 271
 Гаусса–Эверхарта, 205
 Гельфанда–Цетлина, 229
 Гира, 174
 Грэгга–Булирша–Штера, 170, 205
 Дормана–Принса, 142, 146, 205
 Инглэнда, 142
 Йошиды, 199
 Кахана–Ли, 200, 205
 Крамера, 220, 312
 Куинлэна–Тремейна, 201, 205
 Кунцмана–Бутчера, 146, 151
 Лагранжа, 66
 Лапласа, 217, 232, 311
 обобщенный, 218
 Левенберга–Марквардта, 226, 235
 Мерсона, 142
 Монте-Карло, 258
 Накози, 66, 189
 Ньютона, 225
 Паркера–Сочаки, 129
 Рунге–Кутты
 вложенный, 142
 классический, 134
 коллокационный, 146
 неявный, 144
 сопряженный, 193
 явный, 131
 Стеффенсена, 129
 Тейлора, 127
 Уиздома–Холмана, 198
 Фельдберга, 142

- Хойна, 132, 175
Штермера–Верле, 196
Эйлера
 геометрический, 188
 неявный, 144, 173, 186
 симплектический, 191
 явный, 129, 169, 172, 175, 186
Энке, 101, 197
аналитический, 16
бутстрэп, 274
возмущенных оценок, 269, 272
геометрический, 186
градиентного спуска, 224, 230
дифференциального исправления орбиты, 223
многошаговый, 171, 175, 201
овражный, 229
оптимизации, 224
проекторный, 189, 231
простых итераций, 217, 224
прыгающих лягушек, 197
симметричный, 169, 193
симплектический, 191
синхронного слежения, 91
составной, 198, 230, 236
средней точки, 144, 169, 170, 195
стохастического моделирования, 259
 линейный, 267
 нелинейный, 269
 теории специальных возмущений, 39
 трапеций, 132, 144, 169, 175, 195
 численный, 16
 экстраполяционный, 166
миделево сечение, 30, 34
многочлен
 Лагранжа, 147, 172
 Лежандра, 26, 151
 Ньютона, 152, 181
 Эйткена–Невилла, 168
 канонический, 152
моделирование
 математическое, 15
 стохастическое, 258
 численное, 16
модель
 Марсдена–Секанины, 32
 аналитическая, 18
 атмосферного торможения, 33
 гравитации кольца, 106
 гравитации материальной точки, 29
 гравитации протяженного тела, 26
 динамическая, 18
 кинематическая, 18
 математическая, 15
 многоточечная, 108
 плотности атмосферы, 34
 радиационной силы, 30
 реактивной силы, 32
 численная, 15, 18

- модификация Зейделя, 154
- момент прохождения через перигелий, 307
- наклонение орбиты, 308
- невязка, 20, 176, 214
- неустойчивость
 - метода, 178, 184
 - по Ляпунову, 39, 46, 53, 57, 72
- область-мишень, 299
- обратимость по времени, 185
- определение орбиты, 213
 - предварительное, 217
- орбита, 15
 - Херрика, 78
 - кеплеровская, 29, 57, 73, 303
 - номинальная, 294
 - опорная, 72, 73, 76, 124
 - оскулирующая, 70, 73
 - предварительная, 217
 - хаотичная, 296
 - целевая, 299
 - эллиптическая, 305
- орбитальные координаты, 305
- орбитальный период, 307
- ортогональный проектор, 265
- оценка дисперсии несмещенная, 262
- ошибка среднеквадратическая, 219, 255, 262
- ошибки
 - временные, 87
 - вычислительные, 20, 53, 71, 73, 79, 80, 161, 166, 183, 207, 223, 297
 - глобальные, 112, 134, 169, 178
 - интерполяции, 148
 - квадратурной формулы, 150
 - локальные, 102, 112, 134, 137, 142, 160, 178, 180, 182, 193, 194, 198, 199
 - методические, 20, 53, 56, 58, 71, 101, 103, 104, 128, 161, 183, 223
 - модельные, 20, 36, 97, 250
 - наблюдений, 215, 238, 259, 261, 269, 271, 272, 274
 - начальных данных, 57
 - параметрические, 57, 258, 269
 - перенесенные, 134
 - усечения, 128
- параметр
 - гравитационный, 26
 - модифицированный, 79, 106
 - модели, 17
 - стабилизирующий, 60, 63
 - фокальный, 304
- периодограмма, 243
- перигелий, 305
- планета
 - Меркурий, 114
 - Юпитер, 114

- подпространство оценок, 260,
264, 270, 273
- показатель
Ляпунова, 296
короткопериодичности, 102
нелинейности, 276, 279, 286
- порядок метода, 124, 125, 128,
133, 167, 177, 194
- последовательность
Ромберга, 169
гармоническая, 169
релаксационная, 224
- постньютоновские поправки,
35
- постоянная Липшица, 127, 163
- правило Рунге, 137
- предиктор, 175, 182
- преобразование
Кустаанхеймо–Штифеля,
44
Сундмана, 43
временное, 41, 48, 91
каноническое, 185, 191
координатное, 17, 41, 216
- проблема
короткопериодических воз-
мущений, 100
неоднозначного определе-
ния орбит, 237, 292, 313
разности близких величин,
72, 73
- пространство
наблюдений, 260, 272
параметрическое, 228, 260
плоское, 261
- прямое восхождение, 214, 216,
311
- радиус-вектор, 304
- разбиение
Гаусса–Лежандра, 151, 201
Гаусса–Лобатто, 151
Гаусса–Радо, 151, 155
- размерность оврага, 228
- разность разделенная, 153, 181
- распределение
 χ^2 , 262
Фишера, 263
Хотеллинга, 268
нормальное, 259
- регуляризация уравнений, 39,
41
- ряд
Лапласа, 26
Тейлора, 124
степенной, 124
- световое давление, 30
- сила
возмущающая, 38
инерциальная, 29
короткопериодическая, 100
нормальная, 33
радиальная, 33
трансверсальная, 33
центральная, 38
- симметризация, 297
- симплектичность, 185, 187
- сингулярность уравнений, 38
- система дифференциальных урав-
нений, 126

- система координат
 - барицентрическая, 103
 - гелиоцентрическая, 84
 - геоцентрическая, 216
 - гравитационная, 80
 - орбитальная, 32, 305
 - планетоцентрическая, 84
 - подвижная, 79
 - прямоугольная, 17
 - топоцентрическая, 216
- склонение, 214, 216, 311
- скобка Пуассона, 187
- спутник
 - S/2003 J04, 278
 - Адрастея, 232, 251, 286
 - Амальтея, 35, 113, 249, 286
 - Гималия, 113
 - Земли искусственный, 114
 - Метида, 252, 286
 - Теба, 286
 - Фобос, 113
 - геостационарный, 114
 - семейства S/2003, 256, 286
 - экваториальный, 78
- среднее движение, 307
- стабилизация уравнений, 46, 53
 - Баумгарте, 104
 - Накози, 66
 - Фукушимы, 66
 - диссипативная, 59
 - консервативная, 64
 - по времени, 65
 - по интегралу, 61
 - по энергии, 54, 60
- сублимация льда, 32
- сфера влияния, 85, 254, 301
- таблица
 - Бутчера, 133
 - Эйткена–Невилла, 168
- топоцентр, 216
- узел орбиты, 308
- уравнение
 - Кеплера, 74, 307
 - обобщенное, 309
 - Лагранжа, 254, 312
 - конического сечения, 304
 - орбиты, 308, 309
- условие
 - Липшица, 127
 - адекватности орбит, 92
 - коллокации, 146
 - начальное, 126
 - оскуляции, 73
 - порядка, 131, 177
 - согласованности, 177
 - экстремума, 214
- устойчивость
 - асимптотическая, 60
 - метода, 179, 184
 - орбитальная, 165
 - по Ляпунову, 53
- формула Милна, 180
- функция
 - Лагранжа, 300
 - производящая, 191
 - тени, 31
 - устойчивости, 165

- целевая, 214, 239, 241, 260, 272
- шаровая, 27
- хребет оврага, 228
- число
 - Стирлинга, 153
 - арифметического треугольника, 154
 - обусловленности, 158, 228
- член
 - Энке, 72
 - кеплеровский, 73, 81
- шаг интегрирования, 125, 129, 139, 171, 181, 203
- шаг метода спуска, 224, 226
- широта, 26, 216
- экстраполяция
 - Ричардсона, 139, 167
 - решения, 139, 166, 180
- эксцентриситет, 304
- элементы орбитальные, 303
- эллипсоид ошибок, 273
- элонгация, 253
- энергия, 303
- эпоха
 - начальная, 17, 229, 241, 242, 244, 279
 - оскуляции, 72, 85
 - стандартная, 216
- эфемерида, 19
 - DE405, 36, 88, 92, 96, 97, 99, 220
 - Лейни L1, 36
 - Лиске E5, 36
- эффект
 - Пойнтинга–Робертсона, 30
 - запаздывания света, 217
 - негравитационный, 32
 - релятивистский Шварцшильда, 35
- эффективность численного моделирования, 23, 112

Научное издание

АВДЮШЕВ Виктор Анатольевич

**ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ ОРБИТ
НЕБЕСНЫХ ТЕЛ**

*Издание подготовлено в авторской редакции
Дизайн обложки Л.Д. Кривцовой*

Подписано к печати 07.02.2015 г. Формат 60×84¹/₁₆.
Бумага для офисной техники. Гарнитура Georgia.
Усл. печ. л. 19,5.
Тираж 250 экз. Заказ № 1415.

Отпечатано на оборудовании
Издательского Дома
Томского государственного университета
634050, г. Томск, пр. Ленина, 36
Тел. 8+(382-2)-53-15-28
Сайт: <http://publish.tsu.ru>
E-mail: rio.tsu@mail.ru

ISBN 978-5-94621-519-0



9 785946 215190