



НЕКОТОРЫЕ АСПЕКТЫ СОВРЕМЕННЫХ ПРОБЛЕМ МЕХАНИКИ И ИНФОРМАТИКИ

СБОРНИК НАУЧНЫХ СТАТЕЙ

**ПОД РЕДАКЦИЕЙ
Р. Р. Назирова**

СЕРИЯ «МЕХАНИКА, ИНФОРМАТИКА И УПРАВЛЕНИЕ»

**Москва
2018**

Рецензент:
доктор технических наук, профессор *Вениаминов С. С.*

**P47 Некоторые аспекты современных проблем механики и информатики:
сб. науч. ст. М.: ИКИ РАН, 2018. 274 с.**

ISBN 978-5-00015-016-0

Представлены теоретические и прикладные исследования в широком спектре космической механики и информатики. Рассмотрены различные методы, используемые в данных отраслях науки. Показаны удивительные возможности применения методов обработки больших объёмов данных в практических предсказаниях поведения сложных глобальных систем.

Для специалистов в области космических исследований, научных работников, инженеров, а также аспирантов и студентов старших курсов.

Ключевые слова: механика, искусственный интеллект, муссоны, управление, астероид

Some aspects of contemporary problems of mechanics and computer science

Theoretical and applied research in a wide range of space dynamics and computer science is presented. The various methods used in these branches of science are considered. The surprising possibilities of applying the methods of big data processing in practical predictions of the behavior of complex global systems are shown.

For specialists in the field of space research, scientists, engineers, as well as for graduate and postgraduate students.

Keywords: Mechanics, artificial intelligence, monsoons, motion control, asteroid

Серия «Механика, информатика и управление»

Компьютерная вёрстка: *Комарова Н. Ю.*

Дизайн обложки: *Захаров А. Н.*

Мнение редакции не всегда совпадает с точкой зрения авторов статей.

Отдельные статьи даны в авторской редакции.

Качество иллюстраций соответствует предоставленному авторами материалу.

Электронная версия сборника размещена на сайтах ИКИ РАН <http://www.iki.rssi.ru/print.htm>
и Российской научной электронной библиотеки <http://elibrary.ru/>.

СОДЕРЖАНИЕ

Предисловие	5
<i>Ковалева А. С.</i> Авторезонанс в связанных осцилляторах с медленно убывающей жёсткостью	7
<i>Гришин В. А.</i> Классическая теория управления и методы искусственного интеллекта. Проклятие размерностей	19
<i>Алтайский М. В., Зольникова Н. Н.</i> Квантовые системы искусственного интеллекта	27
<i>Суровяткина Е. Д.</i> Индийский муссон: концепция критических элементов, прогноз начала и завершения	53
<i>Прохоренко В. И.</i> Об орбитальных торах во вращающихся системах координат	65
<i>Золотарёв В. В., Овечкин Г. В., Назиров Р. Р.</i> О передаче Оптимизационной Теории лидерства от прикладной классической теории помехоустойчивого кодирования	82
<i>Балтер Б. М., Балтер Д. Б., Егоров В. В., Стальная М. В., Фаминская М. В.</i> Программный комплекс «Геодиалог»: модель экологического управления	91
<i>Кузнецов Н. А., Золотарёв В. В., Овечкин Г. В., Назиров Р. Р., Аверин С. В.</i> Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона	99
<i>Маслов И. А., Гришин В. А., Николенко И. В.</i> Аппаратура для исследования положения линии морского горизонта в видимой области спектра с использованием поляризационных светофильтров	121
<i>Ивашкин В. В., Назиров Р. Р., Эйсмонт Н. А.</i> Динамика космического полёта, теоретическая механика, управление движением	131
<i>Буриков В. С., Буриков И. В., Буриков Ю. В.</i> Ядерная двигательная установка с замкнутым энергоблоком	143
<i>Константинов М. С.</i> Анализ возможностей космической транспортной системы при использовании электроракетной двигательной установки и гравитационного манёвра у Земли в начале межпланетного перелёта	153

<i>Аксенов С. А., Бобер С. А.</i>	
Управление движением космического аппарата на гало-орбите при наличии ограничений на направления корректирующих манёвров	166
<i>Лан А., Ивашкин В. В.</i>	
Исследование траекторий космического аппарата для экспедиции Земля – Апофис – Земля с учётом выбора орбит пребывания у астероида	175
<i>Крючков С. В., Маслов И. А., Николенко И. В., Угольников О. С.</i>	
Комплекс управления телескопом Zeiss-1000 с возможностью проведения наблюдений в режиме удалённого доступа	188
<i>Нароенков С. А., Шустов Б. М.</i>	
Астероидно-кометная опасность: потенциальная угроза	193
<i>Поль В. Г., Симонов А. В.</i>	
Астероидная опасность в 2018 году: текущие реалии и проблемы	203
<i>Гуо П., Ивашкин В. В.</i>	
Оценка точности определения столкновительной с Землёй орбиты опасного астероида Апофис по оптическим измерениям комплекса «Небосвод»	212
<i>Нароенков С. А., Шугаров А. С.</i>	
Исследование точности определения орбиты и места входа в атмосферу опасного небесного тела по наблюдениям КА СОДА	227
<i>Колюка Ю. Ф., Афанасьева Т. И.</i>	
Применение высокоэффективного метода численного интегрирования уравнений движения для решения задач динамики полёта астероидов, сближающихся с Землёй, и направляемых к ним космических аппаратов	236
<i>Сазонов В. С.</i>	
Применение кинетического удара для воздействия на малые опасные астероиды при их позднем обнаружении.	259
<i>Сизова М. Д., Ефремова Е.</i>	
Астероиды как ресурс на будущее	268

ПРЕДИСЛОВИЕ

Наука 21 века будет характеризоваться бурным развитием искусственного интеллекта (ИИ): как собственно развитием методов ИИ, так и проникновением его в различные сферы науки. Перефразируя известное изречение, суть ИИ можно отразить следующей фразой: ПРИШЁЛ – УВИДЕЛ – ОБЪЯСНИЛ. Иными словами, в соответствующей области науки должно появиться «НЕЧТО», обладающее огромным историческим опытом, включающим как описание объектов этой области, так и описание методов их взаимодействия, способное, опираясь на этот опыт, «объяснить» вновь увиденное. Чтó в современной науке является источником описания научных объектов и методов их познания? Безусловно, в первую очередь, это научные статьи. Бурно развивающиеся в настоящее время различные базы данных научных статей — путь в будущее, когда статьи смогут «общаться» друг с другом (автоматическое цитирование, классификация, поиск противоречий и др.). При этом очень важно наличие работ, в которых автор свободно, без каких-либо объективных или субъективных ограничений выражает свои мысли, свой предмет исследования.

Настоящий сборник является попыткой предоставить автору такую возможность. В сборнике представлены статьи очень большого спектра: от глубоких математических исследований отдельных объектов до широкого взгляда на различные проблемы в целом, от развития конкретных методов наблюдательной астрономии до изящной интерпретации больших массивов спутниковой информации, приводящей к уникальным результатам по прогнозированию развития сложных глобальных явлений. Замыкает сборник группа работ, посвящённых новому направлению механики космического полёта: УПРАВЛЯЕМОМУ ДВИЖЕНИЮ ИСКУССТВЕННЫХ НЕБЕСНЫХ ТЕЛ. Последнее открывает уникальные возможности человеку в познании и освоении новых космических далей.

Надеемся, что представленный сборник будет интересен как современным читателям, так и будущим «обучающим системам» неминуемо надвигающегося на нас Искусственного Интеллекта.

АВТОРЕЗОНАНС В СВЯЗАННЫХ ОСЦИЛЛЯТОРАХ С МЕДЛЕННО УБЫВАЮЩЕЙ ЖЁСТКОСТЬЮ

А. С. Ковалева

Институт космических исследований Российской академии наук (ИКИ РАН)

В работе исследуется резонансный перенос энергии от возбуждённого начальным импульсом нелинейного осциллятора к присоединённому линейному осциллятору с медленно убывающей жёсткостью. Показано, что даже при отсутствии внешнего периодического возбуждения в системе возникает резонансный процесс с возрастанием амплитуды колебаний линейного осциллятора и соответствующем убыванием амплитуды колебаний нелинейного возбудителя. Этот эффект назван «внутренним авторезонансом». Влияние параметров системы на возбуждение и устойчивость авторезонанса изучено численно и аналитически.

Ключевые слова: нелинейные системы, резонанс, перенос энергии

ВВЕДЕНИЕ

В работе исследуется перенос энергии в системе двух осцилляторов: нелинейного осциллятора с постоянными параметрами, возбуждённого начальным импульсом и связанного с ним линейного осциллятора с медленно убывающей во времени жёсткостью. Движение исследуется с помощью стандартного метода анализа резонансов [Arnold et al., 2006], использующего известное свойство нелинейного осциллятора изменять амплитуду и частоту колебаний при изменении частоты возбуждения. Способность нелинейного осциллятора поддерживать резонансный режим при изменении частоты возбуждения известна как *авторезонанс* (АР).

Первые исследования АР, возникающего при ускорении частиц [Bohm, Foldy, 1946; McMillan, 1945; Veksler, 1944] и в планетарной динамике [Greenberg, 1973; Sinclair, 1972], показали, что АР может рассматриваться как эффективный метод возбуждения высокоэнергетических режимов в системах различной физической природы; обзор результатов в этой области можно найти в работах [Fajans, Friedland, 2001; Kovaleva, 2017; Neishtadt et al., 2013]. Особое внимание уделялось *внешнему АР*, возбуждаемому внешней силой с медленно меняющейся частотой. Частный случай *параметрического АР* [Assaf, Meerson, 2005; Khain, Meerson, 2001] интерпретировали как расширение понятия классического параметрического резонанса [Fossen, Nijmeijer, 2012] на случай параметрического возбуждения с медленно меняющейся частотой.

В данной работе обсуждается *внутренний АР*, возникающий благодаря медленному изменению параметров системы при отсутствии быстрого периодического возбуждения. Показано, что сильно нелинейный осциллятор, не имеющий не зависящей от энергии движения собственной частоты, может быть захвачен в резонанс при изменении динамических параметров, соответствующих

медленному изменению частоты присоединённого линейного осциллятора. При этом возникает интенсивный перенос энергии от возбуждённого осциллятора к присоединённому, с резким уменьшением амплитуды колебаний нелинейного осциллятора.

С практической точки зрения, внутренний АР может рассматриваться как эффективный механизм гашения сейсмических вибраций, где присоединённый осциллятор играет роль «энергетической ловушки» [Vakakis et al., 2009]. При моделировании системы сейсмической защиты присоединённый линейный осциллятор описывает динамику одномассового изолятора с линейно убывающим во времени натяжением упругого элемента [Zhang et al., 2016] или упругой балки при линейно возрастающем осевом сжатии [Silina et al., 2015].

Уравнения движения представлены в следующей части работы. Предполагается, что оба осциллятора в начальный момент находятся в покое, но нелинейный осциллятор возбуждается начальным импульсом. Такие начальные условия определяют предельную фазовую траекторию (ПФТ), соответствующую максимальному переносу энергии от возбуждённого объекта к ловушке [Kovaleva, 2015; Kovaleva, Manevitch, 2013; Manevitch, Kovaleva, 2013; Manevitch et al., 2011].

Асимптотические решения резонансной системы получены с помощью метода многомасштабных разложений [Nayfeh, Mook, 2008; Sanders et al., 2007]. Построенные решения подтверждают перманентное возрастание амплитуды колебаний линейной ловушки при одновременном уменьшении амплитуды нелинейного возбуждённого осциллятора. Асимптотические решения показывают, что изменения амплитуд обоих осцилляторов адекватно описываются соответствующими скелетными кривыми, т.е. анализ резонансного переноса энергии может быть сведён к анализу скелетных кривых. Результаты численного моделирования иллюстрируют влияние параметров на интенсивность переноса энергии.

РЕЗОНАНСНАЯ ДИНАМИКА СВЯЗАННЫХ ОСЦИЛЛЯТОРОВ

Динамика связанных осцилляторов описывается уравнениями

$$\begin{cases} m_1 \frac{d^2 U_1}{dt^2} + \gamma U_1^3 + c_{12}(U_1 - U_2) = 0, \\ m_2 \frac{d^2 U_2}{dt^2} + c_2(1 - st)U_2 + c_{12}(U_2 - U_1) = 0. \end{cases} \quad (1)$$

где U_1, U_2 — абсолютные перемещения осцилляторов; m_1, m_2 — массы осцилляторов; $m_2 \leq m_1$; c_2 — жёсткость линейного осциллятора; γ — нелинейная жёсткость; c_{12} — жёсткость линейной связи между осцилляторами; $\omega = (c_2/m_2)^{1/2}$ — парциальная собственная частота линейного осциллятора при $c_{12} = 0, s = 0$. Начальные условия имеют вид $U_1 = U_2 = 0, V_1 = dU_1/dt = V_0, V_2 = 0$ при $t = 0_+$ (нижний индекс «+» в дальнейшем будет опущен).

Для упрощения численных процедур удобно привести систему (1) к безразмерной форме, уменьшив число независимых параметров. Для этой цели введём безразмерное «быстрое» и «медленное» время по формулам $\tau_0 = \omega t$

и $\tau = \varepsilon\tau_0$, соответственно. Учитывая слабую связь между осцилляторами, определим малый параметр системы $\varepsilon = c_{12}/2c_2 \ll 1$; малая скорость расстройки определяется формулой $s = 2\varepsilon^2\beta\omega$. Безразмерный начальный импульс имеет вид $v_0 = V_0/\omega$; соответствующие безразмерные переменные заданы соотношениями

$$u_r = \frac{U_r}{v_0}, \quad v_r = \frac{du_r}{d\tau_0} = \frac{V_r}{v_0}, \quad r = 1, 2. \quad (2)$$

Подставляя соотношения (2) в систему (1) и учитывая преобразования параметров, получим безразмерные уравнения движения в виде

$$\begin{cases} \frac{d^2 u_1}{d\tau_0^2} + \frac{8}{3}\alpha u_1^3 + 2\varepsilon M(u_1 - u_2) = 0, \\ \frac{d^2 u_2}{d\tau_0^2} + (1 - 2\varepsilon\xi_0(\tau))u_2 + 2\varepsilon(u_2 - u_1) = 0 \end{cases} \quad (3)$$

с начальными условиями

$$u_1(0) = 0, \quad v_1(0) = 1; \quad u_2(0) = 0, \quad v_2(0) = 0, \quad (4)$$

где

$$M = m_2/m_1 \leq 1, \quad \xi_0(\tau) = \beta\tau, \quad \alpha = 3V_0^2\gamma/8m_1\omega^4. \quad (5)$$

Уравнения (3) включают четыре независимых параметра вместо восьми параметров в системе (1). Напомним, что начальные условия (4) определяют ПФТ системы (3), соответствующую максимальному переносу энергии от возбуждённого осциллятора.

При условии резонанса 1:1 в системе уравнения (3) могут быть записаны в виде

$$\begin{cases} \frac{d^2 u_1}{d\tau_0^2} + u_1 + \varepsilon\mu \left[\frac{8}{3}\alpha u_1^3 - u_1 \right] + 2\varepsilon M(u_1 - u_2) = 0, \\ \frac{d^2 u_2}{d\tau_0^2} + (1 - 2\varepsilon\xi_0(\tau))u_2 + 2\varepsilon(u_2 - u_1) = 0, \end{cases} \quad (6)$$

Параметр $\mu = 1/\varepsilon$ сохраняет формальную эквивалентность уравнений (3) и (6) при условии резонанса 1:1. Отметим, что линейная жёсткость присоединённого осциллятора остаётся положительной и система устойчива на интервале времени I_0 : $\tau_0 \in [0, 1/(2\varepsilon^2\beta)]$ (см. рис. 1). Следовательно, интервал I_0 может рассматриваться как допустимый интервал исследования резонансных колебаний.

Асимптотические решения уравнений (6) ищутся в виде многомасштабных разложений [Nayfeh, Mook, 2008; Sanders et al., 2007]. На первом шаге делается замена переменных

$$\Psi_r = (v_r + iu_r)e^{-i\tau_0}, \quad \Psi_r^* = (v_r - iu_r)e^{i\tau_0}, \quad r = 1, 2, \quad (7)$$

где знак * обозначает комплексно-сопряжённую переменную. Подставив выражения (7) в соотношения (6), получим следующие уравнения:

$$\begin{cases} \frac{d\Psi_1}{d\tau_0} = i\varepsilon \left[\mu \left(\alpha |\Psi_1|^2 - \frac{1}{2} \right) \Psi_1 + M(\Psi_1 - \Psi_2) + G_1 \right], & \Psi_1(0) = 1, \\ \frac{d\Psi_2}{d\tau_0} = -i\varepsilon \left[\xi_0(\tau) \Psi_2 + (\Psi_1 - \Psi_2) + G_2 \right], & \Psi_2(0) = 0. \end{cases} \quad (8)$$

и аналогичные уравнения для сопряжённых переменных. Слагаемые $G_{1,2}$ включают высшие гармоники по быстрой переменной, но их точные выражения не существенны для дальнейшего анализа.

Решение уравнений (8) ищется в виде асимптотического разложения:

$$\Psi_r(\tau_0, \tau, \varepsilon) = \psi_r(\tau) + \varepsilon \psi_r^{(1)}(\tau_0, \tau) + O(\varepsilon^2), \quad r = 1, 2. \quad (9)$$

Уравнения для медленно меняющихся огибающих $\psi_r(\tau)$ могут быть получены усреднением (8) по быстрой переменной τ_0 . Стандартная процедура [Sanders et al., 2007] приводит к следующим уравнениям для медленных переменных $\psi_r(\tau)$:

$$\begin{cases} \frac{d\psi_1}{d\tau} = i \left[\mu \left(\alpha |\psi_1|^2 - \frac{1}{2} \right) \psi_1 + M(\psi_1 - \psi_2) \right], & \psi_1(0) = 1, \\ \frac{d\psi_2}{d\tau} = -i \left[\xi_0(\tau) \psi_2 + (\psi_1 - \psi_2) \right], & \psi_2(0) = 0. \end{cases} \quad (10)$$

и аналогичным уравнениям для сопряжённых медленных переменных. Если огибающие $\psi_r(\tau)$ найдены, то главное приближение к решению порождающей системы (6) имеет вид

$$u_{r_0}(\tau_0, \tau) = |\psi_r(\tau)| \sin(\tau_0 + \gamma_r(\tau)), \quad v_{r_0}(\tau_0, \tau) = |\psi_r(\tau)| \cos(\tau_0 + \gamma_r(\tau)), \quad r = 1, 2, \quad (11)$$

где $\gamma_r(\tau_1) = \arg \psi_r(\tau)$. Известно, что процедура усреднения даёт ошибку аппроксимации $O(\varepsilon)$ на интервале времени $\tau_0 \sim O(1/\varepsilon)$. Однако более тонкий анализ показал, что интервал сходимости зависит от свойств приближённого решения [Arnold et al., 2006; Sanders et al., 2007]. Более того, относительно большие значения жёсткости связи ε , использующиеся в практических задачах, не обязательно приводят к возрастанию ошибки на большом интервале времени (см., например, [Vakakis et al., 2009]). Точность полученных приближений должна дополнительно проверяться численным моделированием.

Легко проверить, что система (10) сохраняет следующий интеграл движения:

$$|\psi_1(\tau)|^2 + M |\psi_2(\tau)|^2 = |\psi_1(0)|^2 = 1. \quad (12)$$

Из соотношений (10), (12) следует, что $|\psi_2(\tau)|^2 \approx \tau^2$, но $|\psi_1(\tau)|^2 \approx 1 - M\tau^2$ на начальном интервале времени. Замена переменных

$$\psi_1 = a \cdot e^{i\delta_1}, \quad \sqrt{M} \psi_2 = b \cdot e^{i\delta_2}, \quad \Delta = \delta_2 - \delta_1, \quad a = \cos \theta, \quad b = \sin \theta \quad (13)$$

приводит комплексные уравнения (10) к двум уравнениям

$$\begin{cases} \frac{d\theta}{d\tau} = -\sqrt{M} \cdot \sin \Delta, \\ \sin 2\theta \frac{d\Delta}{d\tau} = - \left[\mu \left(\alpha \cos^2 \theta - \frac{1}{2} \right) + \xi(\tau) \right] \sin 2\theta - 2\sqrt{M} \cos 2\theta \cdot \cos \Delta. \end{cases} \quad (14)$$

где $\zeta(\tau) = \xi_0(\tau) - (1 - M)$. Как и в квазилинейной теории [Manevitch, Kovaleva, 2013], начальные условия $\theta(0) = 0$, $\Delta(0) = \pi/2$ определяют ПФТ медленного движения в координатах θ , Δ .

Из выражений (13), (14) следует, что переменные $y = a^2$, $z = b^2$ описывают состояния нелинейного и линейного осцилляторов, соответственно. Рассматривая $z = \sin^2\theta$ как новую независимую переменную, получим следующие уравнения:

$$\begin{cases} \frac{dz}{d\tau} = -2\sqrt{Mz(1-z)} \cdot \sin \Delta, \\ \frac{d\Delta}{d\tau} = -\left[\mu\left(\alpha(1-z) - \frac{1}{2}\right) + \zeta(\tau)\right] - \sqrt{M} \frac{1-2z}{\sqrt{z(1-z)}} \cos \Delta. \end{cases} \quad (15)$$

с начальными условиями $z(0) = 0$, $\Delta(0) = \pi/2$. Адиабатические (квазистационарные) решения уравнений (15) ищутся из соотношений

$$\frac{dz}{d\tau} = 0, \quad \frac{d\Delta}{d\tau} = 0. \quad (16)$$

при «замороженном» параметре $\zeta(\tau)$. Из первого уравнения системы (16) следует, что $\sin \Delta = 0$, причём фаза $\Delta = 0$ устойчива [Manevitch, Kovaleva, 2013]. При $\Delta = 0$ второе уравнение системы (16) приводится к виду

$$F(z) = -\mu\left[\alpha(1-z) - \frac{1}{2}\right] - \sqrt{M} \frac{1-2z}{\sqrt{z(1-z)}} = \zeta(\tau). \quad (17)$$

При $\xi_0 = 0$, $\zeta = -(1 - M)$ решение \bar{z}_0 уравнения (17) определяет устойчивое положение равновесия консервативной системы. Если $\xi_0(\tau) \neq 0$, то решение $\bar{z}(\tau)$ при «замороженном» параметре $\zeta(\tau)$ определяет *скелетную кривую* линейного осциллятора, для которой $\bar{z}(0) = \bar{z}_0$. Напомним, что линейный осциллятор включён в нелинейную систему, т.е. его скелетная кривая имеет более сложную форму, чем вертикальная прямая осциллятора в линейной системе.

Подставляя $y = 1 - z$ в соотношения (15), получим уравнение для квадрата амплитуды $y = a^2$; соответствующее адиабатическое решение $\bar{y}(\tau) = 1 - \bar{z}(\tau)$ определяет скелетную кривую нелинейного осциллятора в связанной системе.

Учитывая положительность производной

$$F'(z) = \mu\alpha + \frac{0,5\sqrt{M}}{[z(1-z)]^{3/2}} > 0, \quad 0 < z < 1, \quad (18)$$

можно утверждать, что функция $F(z)$ возрастает при увеличении аргумента z и, соответственно, корень $\bar{z}(\tau)$ уравнения (17) растёт при увеличении τ ; соответственно, решение $\bar{y}(\tau) = 1 - \bar{z}(\tau)$ уменьшается при росте τ . Это означает, что резонансное взаимодействие, приводящее к уменьшению энергии возбуждённого осциллятора, можно рассматривать как результат необратимого переноса энергии от возбуждённого осциллятора к присоединённому.

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

В этой части работы приведены результаты численного моделирования, иллюстрирующие возникновение резонансного переноса энергии. Напомним, что безразмерная медленная система (10) включает четыре независимых коэффициента, но эти коэффициенты зависят от всех восьми параметров исходной системы, включая начальный импульс (см. уравнение (5)). Следовательно, влияние каждого исходных параметров должно исследоваться независимо. В последней части работы показано, что сильно нелинейные системы с коэффициентом $\alpha > 1$ можно исключить из рассмотрения, так как малые колебания присоединённого осциллятора в такой системе слабо влияют на динамику нелинейного возбудителя. Предсказанный эффект подтверждён численным моделированием системы (3) с параметрами $M = 1$, $\varepsilon = 0,1$, $\beta = 0,05$ и различными коэффициентами нелинейности (рис. 1).

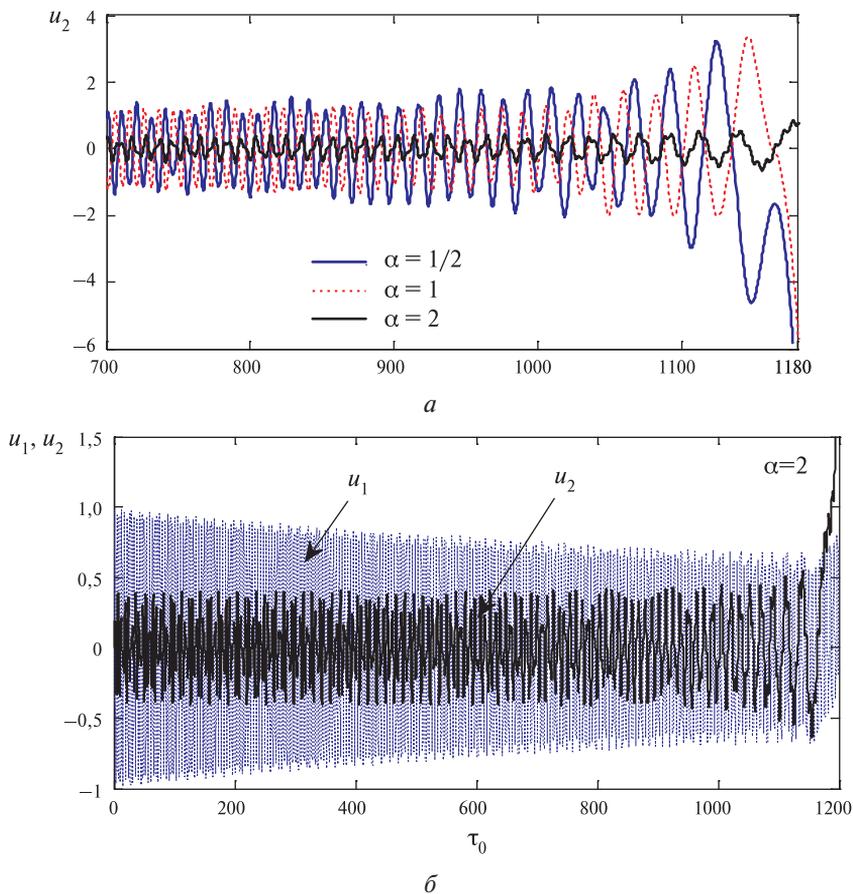


Рис. 1. Динамика системы (3) с параметрами $M = 1$, $\varepsilon = 0,1$, $\beta = 0,05$ и различными коэффициентами нелинейности: a — перемещения u_2 линейного осциллятора в системе с различными коэффициентами нелинейности; b — перемещения обоих осцилляторов в системе с нелинейностью $\alpha = 2$

Рисунок 1а иллюстрирует динамику линейного осциллятора в системе (3) с коэффициентами нелинейности $\alpha = 0,5$, $\alpha = 1$, $\alpha = 2$. Наблюдается усиление колебаний при $\alpha = 0,5$ и $\alpha = 1$ и малые колебания при $\alpha = 2$. Заметим, что колебания теряют устойчивость при $\tau_0 \approx 1/(2\varepsilon^2\beta) = 1000$. Рисунок 1б иллюстрирует колебания линейного и нелинейного осцилляторов в системе с сильной нелинейностью $\alpha = 2$. Можно отметить сходство резонансных колебаний в системах с коэффициентами нелинейности $\alpha = 0,5$ и $\alpha = 1$ и слабое влияние малых колебаний линейного осциллятора на динамику возбудителя при сильной нелинейности $\alpha = 2$.

Учитывая подобие динамических реакций при $\alpha = 1$ и $\alpha = 0,5$, ограничимся изучением систем с коэффициентом нелинейности $\alpha = 0,5$. Рисунок 2 иллюстрирует перенос энергии в базовой системе с параметрами $M = 1$, $\alpha = 0,5$, $\varepsilon = 0,1$ ($\mu = 10$), $\beta = 0,05$. В этом случае уравнение (17) принимает вид

$$F(z) = \frac{\mu}{2}z - \frac{1-2z}{\sqrt{z(1-z)}} = \beta\tau. \quad (19)$$

Система устойчива на интервале $\tau \in [0, 1/(2\varepsilon\beta)]$, или $\tau \in [0, 100]$. На рис. 2а видно, что уменьшение амплитуды колебаний нелинейного осциллятора происходит одновременно с возрастанием амплитуды колебаний присоединённой линейной ловушки. Рисунок 2б описывает функцию $z(\tau) = b^2(\tau)$, соответствующую скелетную кривую $\bar{z}_l(\tau)$ и её линейную аппроксимацию $\bar{z}_l(\tau)$ (см. выражение (20)). На рис. 2б видим, что скелетная кривая адекватно описывает возрастание квадрата амплитуды линейного осциллятора. Следовательно, изучение резонанса можно свести к анализу скелетных кривых.

В общем случае уравнение (19) решается численно. Но из рис. 2б следует, что кривая $\bar{z}(\tau)$ хорошо аппроксимируется линейной функцией $\bar{z}_l(\tau) = \bar{z}_0 + \xi(\tau)$, удовлетворяющей уравнению

$$F(\bar{z}_0 + \xi) = F(\bar{z}_0) + F'(\bar{z}_0)\xi = \beta\tau - (1 - M). \quad (20)$$

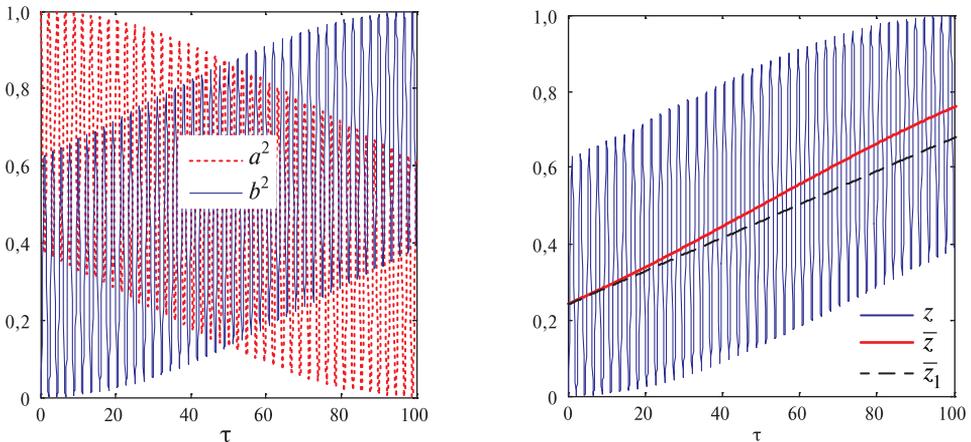


Рис. 2. Динамика осцилляторов в базовой системе с параметрами $M = 1$, $\alpha = 0,5$, $\varepsilon = 0,1$, $\beta = 0,05$: a — квадраты амплитуд a и b ; b — функция $z(\tau) = b^2(\tau)$, скелетная кривая $\bar{z}(\tau)$ (сплошная линия) и её линейная аппроксимация $\bar{z}_l(\tau)$ (штрих-пунктир) для линейного осциллятора

Учитывая, что $F(\bar{z}_0) = -(1 - M)$, получим уравнение для переменной ξ

$$F'(\bar{z}_0)\xi = \beta\tau. \quad (21)$$

Рисунок 3 демонстрирует скелетные кривые (19) в системах с различными параметрами. Сплошная линия (1) описывает скелетную кривую $\bar{z}(\tau)$ базовой системы. Точечный пунктир соответствует скелетной кривой $\bar{z}(\tau)$ в системе с параметром $\varepsilon = 0,1$ и скоростью $\beta = 0,01$. В этой системе масса и жёсткость совпадают с базовыми параметрами, но скорость β падает в силу уменьшения параметра $s = 2\varepsilon^2\beta\omega$. Очевидно, что при $\beta = 0,01$ скелетная кривая возрастает медленно, и результирующий процесс не может рассматриваться как локализация энергии на линейной ловушке. Штрих-пунктир (3) соответствует скелетной кривой $\bar{z}(\tau)$ в системе со слабой связью $\varepsilon = 0,05$, полученной при уменьшении жёсткости связи c_{12} . Как и ожидалось, уменьшение жёсткости связи приводит к падению амплитуды колебаний по сравнению с базовой системой. Наконец, точечный пунктир (4) описывает усиленный перенос энергии в системе с относительной массой $M = 0,25$. В этой системе $m_1 = 4m_2$; коэффициент нелинейности γ также увеличивается в четыре раза, сохраняя значение коэффициента α (см. выражения (5)).

В случае (4) квазистационарное решение может быть найдено из уравнения

$$F(z) = \frac{\mu}{2}z - \frac{1-2z}{2\sqrt{z(1-z)}} = \beta\tau - \frac{3}{4}. \quad (22)$$

Хотя начальное значение скелетной кривой $\bar{z}^{(4)}(0) = \bar{z}_0^{(4)} = 0,106$ намного меньше, чем значение $\bar{z}_0^{(1)} = 0,24$ для скелетной кривой в базовой системе, её конечное значение $\bar{z}_T^{(4)} = 0,74$ при $\tau = 100$ близко к соответствующему значению $\bar{z}_T^{(1)} = 0,76$ в системе с равными массами. Этот результат можно объяснить с помощью аппроксимаций (18), (20), (21). Из соотношений (18), (20) следует, что начальное состояние \bar{z}_0 при $\tau = 0$ убывает при уменьшении относительной массы M . В то же время, из соотношения (18) следует, что производная $F'(\bar{z}_0)$ возрастает при уменьшении параметров M и \bar{z}_0 , тем самым компенсируя разрыв между начальными точками при различных значениях M .

Полученные аналитические аппроксимации подтверждаются результатами численного моделирования для системы (3) с параметрами $\alpha = 0,5$, $\varepsilon = 0,1$, $\beta = 0,05$ и различными значениями коэффициента M (рис. 4, см. с. 15).

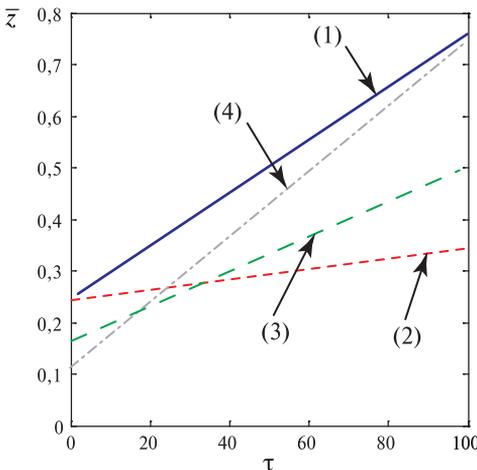


Рис. 3. Скелетные кривые $\bar{z}(\tau)$ в системах с различными параметрами: (1) $\varepsilon = 0,1$, $M = 1$, $\beta = 0,05$ (базовая система); (2) $\varepsilon = 0,1$, $M = 1$, $\beta = 0,01$; (3) $\varepsilon = 0,05$, $M = 1$, $\beta = 0,05$; (4) $\varepsilon = 0,1$, $M = 0,25$, $\beta = 0,05$

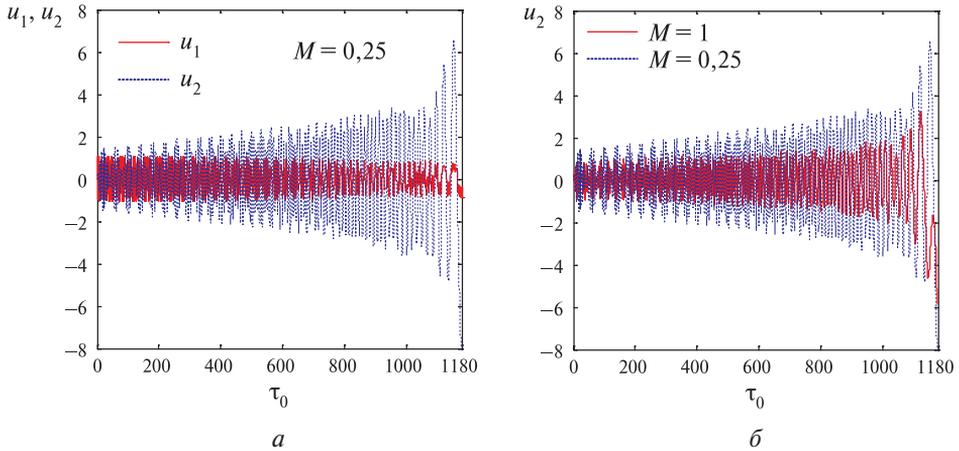


Рис. 4. Перемещения u_1, u_2 в системах с параметрами $\alpha = 0,5$, $\varepsilon = 0,1$, $\beta = 0,05$ и различными значениями относительной массы M : a — перемещения обоих осцилляторов при $M = m_2/m_1 = 0,25$; b — перемещения линейного осциллятора при $M = 0,25$ и $M = 1$

УСЛОВИЯ ВОЗБУЖДЕНИЯ РЕЗОНАНСА

Важно помнить, что собственная частота сильно нелинейного осциллятора не фиксирована как в линейном случае, но явно зависит от энергии колебаний, определяемой как нелинейностью, так и начальными условиями. Следовательно, резонанс 1:1 в стационарном аналоге системы (3) может появиться, если собственная частота нелинейного осциллятора

$$\frac{d^2 u_1}{d\tau_0^2} + \frac{8}{3} \alpha \cdot u_1^3 = 0 \quad (23)$$

с начальными условиями $u_1(0) = 0$, $v_1(0) = 1$ близка к 1. Частота нелинейных колебаний приближённо определяется с помощью метода гармонического баланса [Nayfeh, Moork, 2008], при котором периодическое решение ищется в виде

$$u_1(\tau_0) = A \sin[\Omega(A)\tau_0], \quad \frac{du_1}{d\tau_0} = \Omega(A)A \cdot \cos[\Omega(A)\tau_0] \quad (24)$$

с частотой $\Omega(A) = \alpha^{1/2}A$. С учётом начальных условий получим

$$\begin{cases} \Omega(A)A = \alpha^{1/2}A^2 = 1, & A = \alpha^{-1/4}, \\ \Omega(A) = \alpha^{1/4}, \end{cases} \quad (25)$$

и, соответственно, $\Omega = 1$ при $\alpha = 1$. Отметим, что приближённый анализ справедлив для различных значений коэффициента нелинейности $\alpha \sim 1$. В частности, параметрам $\alpha = 2$ и $\alpha = 0,5$ соответствуют частоты $\Omega = 1,18$ и $\Omega = 0,84$, близкие к требуемому резонансному значению $\Omega = 1$, но формула (25) показывает, что амплитуда A падает при возрастании α . Этот эффект ярче проявляется в связанных осцилляторах (см. рис. 1).

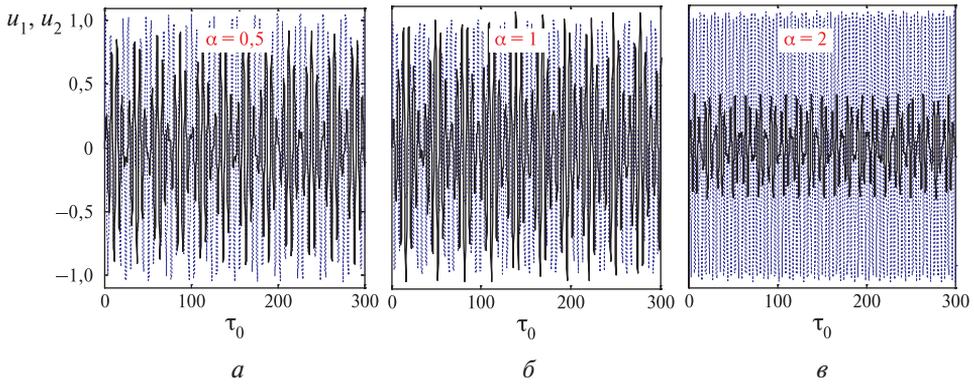


Рис. 5. Перемещения u_1, u_2 консервативных осцилляторов с параметрами $M = 1, \varepsilon = 0,1, \beta = 0$ и различными коэффициентами нелинейности; u_1 — штрихпунктирная линия, u_2 — сплошная линия

Условия возбуждения резонансного переноса энергии для нелинейных систем можно формально получить так же, как и для квазилинейных осцилляторов [Manevitch, Kovaleva, 2013]. Однако зависимость коэффициентов от параметров $\mu = 1/\varepsilon, \alpha$ и M усложняет теоретический анализ. Для упрощения задачи предположим, что $|z| \ll 1$. Подставляя $b = \sqrt{z}$ в выражения (15) и учитывая, что $|b| \ll 1, \zeta_0 = 0$, приведём соотношения (15) к виду

$$\begin{cases} \frac{db}{d\tau} = -\sqrt{M} \cdot \sin \Delta, \\ b \frac{d\Delta}{d\tau} = -\left[\mu \left(\alpha(1-b^2) - \frac{1}{2} \right) - (1-M) \right] b - \sqrt{M} \cdot \cos \Delta, \end{cases} \quad (26)$$

Легко убедиться, что уравнения (26) совпадают с усреднёнными уравнениями для консервативного осциллятора Дуффинга. Используя результаты, ранее приведённые для осциллятора Дуффинга [Manevitch et al., 2011], получим, что линейный осциллятор совершает малые колебания, если выполняется условие

$$27\mu\alpha M < [\mu\alpha - 0,5\mu - (1-M)]^3. \quad (27)$$

Пусть $M = 1, \mu = 10$ ($\varepsilon = 0,1$). В этом случае условие (27) выполняется при $\alpha > 1,2$; при $\alpha < 1,2$ имеем $27\mu\alpha > (\mu\alpha - 0,5\mu)^3$. Это означает, что при $\alpha < 1,2$ в системе возникает энергообмен, характеризующийся большими колебаниями обоих осцилляторов, но при $\alpha > 1,2$ энергия локализуется на возбуждённом нелинейном осцилляторе, а линейный осциллятор совершает малые колебания (рис. 5).

Аналогичный результат для адиабатической системы с медленно убывающей жёсткостью линейного осциллятора показан на рис. 1. В частном случае $M = 1/4$ получим, что энергия остаётся локализованной на возбуждённом осцилляторе при $\alpha > 0,98$.

Полученные результаты имеют простой физический смысл: энергообмен в консервативной системе может возникнуть при умеренной нелинейности, удовлетворяющей условиям (25); рост нелинейности ведёт к фактическому

ослаблению связи между частями системы и локализации энергии на возбуждённом нелинейном осцилляторе. В то же время, уменьшение относительной массы M также эквивалентно снижению жёсткости связи. При этом уменьшается перенос энергии от возбудителя к ловушке и падает амплитуда колебаний присоединённого осциллятора.

ЛИТЕРАТУРА

- [Arnold et al., 2006] *Arnold V. I., Kozlov V. V., Neishtadt A. I.* Mathematical Aspects of Classical and Celestial Mechanics. N. Y.: Springer, 2006.
- [Assaf, Meerson, 2005] *Assaf M., Meerson B.* Parametric autoresonance in Faraday waves // *Physical Review E*. 2005. V. 72. P. 011810.
- [Bohm, Foldy, 1946] *Bohm D., Foldy L.* The theory of the synchrotron // *Physical Review*. 1946. V. 70. P. 249.
- [Fajans, Friedland, 2001] *Fajans J., Friedland L.* Autoresonant (nonstationary) excitation of pendulums, Plutinos, plasmas, and other nonlinear oscillators // *American J. Physics*. 2001. V. 69. P. 1096–1102.
- [Fossen, Nijmeijer, 2012] *Fossen T., Nijmeijer H.* Parametric Resonance in Dynamical Systems. N. Y.: Springer, 2012.
- [Greenberg, 1973] *Greenberg R. J.* Evolution of satellite resonances by tidal dissipation // *The Astronomical J.* 1973. V. 78. P. 338–346.
- [Khain, Meerson, 2001] *Khain E., Meerson B.* Parametric autoresonance // *Physical Review E*. 2001. V. 64. P. 036169.
- [Kovaleva, 2015] *Kovaleva A.* Capture into resonance of coupled Duffing oscillators // *Physical Review E*. 2015. V. 92. P. 022909.
- [Kovaleva, 2017] *Kovaleva A.* Control of autoresonance in mechanical and physical models // *Philosophical Transactions of the Royal Society*. 2017. V. 375. Iss. 2088P. 20115213.
- [Kovaleva, Manevitch, 2013] *Kovaleva A., Manevitch L. I.* Limiting phase trajectories and emergence of autoresonance in nonlinear oscillators // *Physical Review E*. 2013. V. 88. P. 024901.
- [Manevitch, Kovaleva, 2013] *Manevitch L. I., Kovaleva A.* Nonlinear energy transfer in classical and quantum systems // *Physical Review E*. 2013. V. 87. P. 022904.
- [Manevitch et al., 2011] *Manevitch L. I., Kovaleva A., Shepelev D. S.* Non-smooth approximations of the limiting phase trajectories for the Duffing oscillator near 1:1 resonance // *Physica D: Nonlinear Phenomena*. 2011. V. 240, P. 1–12.
- [McMillan, 1945] *McMillan E. M.* The synchrotron — a proposed high energy particle accelerator // *Physical Review*. 1945. V. 68. P. 144.
- [Nayfeh, Mook, 2008] *Nayfeh A. H., Mook D. T.* Nonlinear Oscillations. N. Y.: John Wiley and Sons, 2008.
- [Neishtadt et al., 2013] *Neishtadt A. I., Vasiliev A. A., Artemyev A. V.* Capture into resonance and escape from it in a forced nonlinear pendulum // *Regular and Chaotic Dynamics*. 2013. V. 18. P. 691–701.
- [Sanders et al., 2007] *Sanders J. A., Verhulst F., Murdock J.* Averaging Methods in Nonlinear Dynamical Systems. N. Y.: Springer, 2007.
- [Silina et al., 2015] *Silina K. G., Kikot I. P., Manevich L. I.* Energy exchange and localization in the planar motion of a weightless beam carrying two discrete masses // *Regular and Chaotic Dynamics*. 2015. V. 20. P. 109–122.
- [Sinclair, 1972] *Sinclair A. T.* On the origin of commensurabilities among the satellites of Saturn // *Monthly Notices of the Royal Astronomical Society*. 1972. V. 160. Iss. 2. P. 169–187.
- [Vakakis et al., 2009] *Vakakis A. F., Gendelman O. V., Bergman L. A., McFarland D. M., Kerschen G., Lee Y. S.* Passive Nonlinear Targeted Energy Transfer in Mechanical and Structural Systems. N. Y.: Springer, 2009.

- [Veksler, 1944] *Veksler V. I.* A new method of accelerating relativistic particles // Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS. 1944. V. 43. P. 329–331.
- [Zhang et al., 2016] *Zhang Z., Koroleva I., Manevitch L. I., Bergman L. A., Vakakis A. F.* Nonreciprocal acoustics and dynamics in the in-plane oscillations of a geometrically nonlinear lattice // Physical Review E. 2016. V. 94. P. 032214.

AUTORESONANCE IN COUPLED OSCILLATORS WITH SLOWLY-DECAYING STIFFNESS

A. S. Kovaleva

Space Research Institute of the Russian Academy of Sciences (IKI RAN)

In this work, we study resonance energy transfer from an impulsively loaded strongly nonlinear oscillator to a weakly coupled linear attachment with a slowly time-decaying stiffness. It is shown that even in the absence of external periodic forcing both oscillators may exhibit the resonance phenomenon, with the permanent response enhancement of the linear oscillator and the corresponding response reduction of the nonlinear actuator. This effect is said to be internal autoresonance. The influence of the system parameters on the emergence and stability of autoresonance is investigated both analytically and numerically.

Keywords: nonlinear systems, resonance, energy transfer

Kovaleva Agnesa Solomonovna, Leading Researcher, Doctor of Science, agnessa_kovaleva@hotmail.com

КЛАССИЧЕСКАЯ ТЕОРИЯ УПРАВЛЕНИЯ И МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА. ПРОКЛЯТИЕ РАЗМЕРНОСТЕЙ

В. А. Гришин

Институт космических исследований РАН (ИКИ РАН)
Московский государственный технический университет
имени Н. Э. Баумана (МГТУ)

Искусственный интеллект всё шире вторгается во все стороны жизни. Этот процесс естественный и его невозможно игнорировать. Статья является попыткой хоть в какой-то степени приблизиться к решению большой и сложной задачи — тесной интеграции классической теории управления и методов искусственного интеллекта. Рассматривается размерность пространств фазовых и управляющих параметров.

Ключевые слова: теория систем управления, искусственный интеллект, проклятие размерности

ВВЕДЕНИЕ

Мы будем рассматривать методы искусственного интеллекта (ИИ) с позиции создания мобильных систем высокой степени автономности. То есть речь будет идти о системах автономного ИИ, способных анализировать текущую обстановку, самостоятельно формулировать цели, ставить конкретные задачи и выбирать методы их решения (достижения). Причём указанные задачи должны решаться как в случае функционирования в естественной (природной) среде, так и в случае конфликтного и/или кооперативного взаимодействия с другими системами, как искусственного, так и естественного интеллекта. Подобные системы часто называют *physically embodied rational intelligent agents* [Embodied, 2004; Hopgood, 2001; Jones, 2008; Luger, 2002; Russell, Norvig, 2003]. Очевидно, что все живые организмы с достаточно высокой степенью организации (включая человека) также полностью соответствуют этому определению. Поэтому, если это не будет специально указано, рассуждения будут относиться в равной степени к системам ИИ, высшим животным или человеку.

Такая постановка задачи сразу же отсекает подавляющее большинство методов ИИ, известных в настоящее время [Grishin, 2012]. Кроме этого возникает весьма нетривиальная задача развития эффективных методов рефлексивного управления, которые позволяют манипулировать поведением систем как искусственного, так и естественного интеллекта (противников и союзников), с которыми взаимодействует наша система, а также выявления таких попыток манипулирования, как со стороны противников, так и со стороны союзников. Следует также отметить, что динамический выбор противников

Гришин Владимир Александрович — старший научный сотрудник, кандидат технических наук, доцент, vgrishin@iki.rssi.ru

и союзников сам по себе требует глубокого анализа и прогноза развития ситуаций. Насколько известно автору, это направление сейчас находится в самом зачаточном состоянии применительно к системам ИИ. Некоторые попытки формализации процессов рефлексивного управления можно найти в работе [Adversarial..., 2007]. Нельзя, однако, не упомянуть про работы Дружинина и Конторова (1982) и [Дружинин и др., 1989], в которых вопросы рефлексивного управления рассмотрены достаточно подробно. В человеческом обществе рефлексивное управление распространено чрезвычайно широко [Lefebvre, 1999; Shemayev, 2007; Thomas, 2002, 2004] и столь же широко отражено в литературе, например у У. Шекспира в пьесе «Отелло»(1603) (*Shakespeare W. Othello*. URL: <http://shakespeare.mit.edu/othello/full.html>).

Необходимость анализа методов ИИ с позиций классической теории управления диктуется тем, что создание роботов под управлением автономного ИИ невозможно без высокой, а скорее, высочайшей, степени интеграции классических методов управления и методов ИИ. И такая интеграция может быть выполнена только на основе методов и подходов, а также в системе понятий, используемых в классических системах управления (СУ). Другое дело, что, скорее всего по прошествии некоторого времени как классические СУ, так и интеллектуальные СУ станут частными случаями общих методов управления.

1. РАЗЛИЧИЯ МЕЖДУ КЛАССИЧЕСКИМИ И ИНТЕЛЛЕКТУАЛЬНЫМИ СИСТЕМАМИ УПРАВЛЕНИЯ

Чем же различаются классические и интеллектуальные СУ, где проходит водораздел между ними? Системы управления на основе автономного ИИ отличаются от классических СУ по целому ряду параметрам. В частности, можно указать на следующие различия:

- большие размерности пространства информационных и управляющих параметров;
- сложная структура многомерных целевых функций;
- существенно меньшая определённость состояния системы и окружающей среды;
- функционирование в условиях явного или скрытого противодействия (состояние конфликта, рефлексивное управление и т. д.);
- большая роль управлений по возмущению, а также по прогнозируемому возмущению;
- многопетлевые обратные связи, структура которых может меняться динамически.

По-видимому, всё-таки ключевым вопросом, по которому происходит водораздел между классическими методами управления и ИИ, является размерность фазового пространства, описывающего состояние системы, а также размерность пространства возможных управлений. При увеличении размерности пространства происходит целый ряд крайне неприятных изменений, часть которых можно проследить, например, в такой классической задаче, как поиск экстремума. Поэтому рассмотрим более подробно размерность области функционирования.

2. РАЗМЕРНОСТЬ ФАЗОВОГО ПРОСТРАНСТВА

Классические СУ функционируют в пространствах сравнительно невысокой размерности. Иногда в этих пространствах могут быть сложные, а в некоторых случаях динамически меняющиеся системы нелинейных ограничений (например, как в случае управления полётом маневрирующих гиперзвуковых беспилотных летательных аппаратов в атмосфере), тем не менее, фазовые траектории системы являются достаточно плотным множествами. И в этих пространствах есть метрики, определено понятие окрестности, понятие сходимости, предела, а также различные дифференциальные и интегральные операторы.

Автономные системы искусственного интеллекта — это по определению весьма сложные системы, которые описываются весьма большим количеством переменных. И это влечёт за собой целый ряд последствий. Вообще говоря, пространства высокой размерности сами по себе обладают весьма специфическими свойствами, которые совершенно не характерны для пространств низкой размерности. Некоторые из них представлены в работе [Zorich, 2011]. Поясним некоторые особенности пространств высокой размерности на примере задач оптимизации.

В классических СУ часто ставится (и решается) задача оптимизации. Однако при увеличении размерности пространства, хотя возможность постановки такой задачи сохраняется, но возможности по её решению падают катастрофически. Задачи оптимизации в пространствах большой размерности приобретают следующие особенности:

1. Целевые функции сложных систем в пространствах высокой размерности носят векторный характер. Хотя чисто теоретически возможно сведение векторной целевой функции к скалярной функции, но такое преобразование становится настолько ситуационно неустойчивым с ростом размерности, что просто теряет смысл.
2. Объём области притяжения экстремума с ростом размерности становится исчезающе малым по сравнению с объёмом всего пространства с учётом физических ограничений. Вероятность найти такой экстремум (глобальный) за приемлемое время (время функционирования системы) очень быстро стремится к нулю, особенно с учётом п. 1.
3. Ограничения на переменные сохраняются, более того, их роль значительно возрастает. По этой причине понятие «оптимальное решение» в пространствах высокой размерности просто теряет смысл и трансформируется в понятие «приемлемое решение».

Получается, что с ростом размерности пространства поиск экстремума становится всё более и более трудной задачей. Пространства высокой размерности для большинства практически полезных случаев не могут быть достаточно подробно исследованы. Отсюда теряется смысл понятия «оптимальности».

Ограниченные ресурсы СУ не позволяют непосредственно работать в пространстве управлений очень высоких размерностей. Очень специфические свойства пространств высокой размерности, тем не менее, позволяют осуществлять такое управление при существенно меньшей размерности пространства управляющих параметров, при котором не происходит катастрофического

ухудшения качества управления. Рассмотрим, например, задачи управления движением человека. Эти задачи сводятся к небольшому числу образцов (их число, вероятно не превышает нескольких сотен). Причём эти образцы зависят от небольшого числа параметров. То есть являются многообразиями небольшой размерности. Эти многообразия погружены в пространство с размерностью порядка 640–850 (это число мышц у человека). Из них непосредственно в процессе движения функционирует порядка 200 мышц, а в процессе стояния (задача сохранения равновесия) — около 400. Если эти многообразия дополнить некоторой малой окрестностью (обусловленной неточностью управления и внешними возмущениями), то суммарный объём всех этих многообразий будет ничтожно мал, по сравнению с объёмом пространства с размерностью 640–850 с учётом всех физических ограничений. То есть фазовые подпространства (многообразия), используемые для управления, расположены очень неплотно в объёмлющем пространстве. И вот на этих многообразиях и происходит функционирование организма в течение всей жизни.

Указанные особенности справедливы и для более сложных форм поведения. Например — движения по горным склонам, когда требуется тесное взаимодействие со зрительной системой, анализ устойчивости камней и осыпей, планирование относительно безопасного маршрута и т. д. Другой пример — движение в пустыне, где необходимо минимизировать потери воды, которая становится основным условием выживания. То есть система многообразий, которой располагает организм, «настроена» на вполне конкретную среду обитания и в иной среде будет, в лучшем случае, неэффективной.

Пространство управлений системы ИИ небольшой размерности вложено в пространство возможных управлений очень большой размерности. И, если в пространстве очень высокой размерности есть гладкость и сходимости, то в пространстве малой размерности совершенно необязательно будет иметь место дифференцируемость. Кроме того, может сильно отличаться и топологические свойства пространства. Могут исчезнуть непрерывность, сходимости, появляются новые, не характерные для обычного физического мира метрики. И здесь появляются объекты из «чужого мира» для классических СУ, а именно — сценарии, фреймы, графы, семантические сети и т. д.

В качестве примера можно привести задачу о семи кёнигсбергских мостах (можно пройти по всем семи мостам Кёнигсберга, не проходя ни по одному из них дважды) отображается в виде графа, который содержит несравненно меньше информации, чем реальный город. При этом, естественно, теряется непрерывность, дифференцируемость отображения и другие приятные свойства классических фазовых пространств. Другой пример — реализация инструкций по движению к кладу в классической книге Р.Л. Стивенсона «Остров сокровищ» 1883 г. (*Stevenson R. L. Treasure Island*. URL: http://www.planetpublish.com/wp-content/uploads/2011/11/Treasure_Island_NT.pdf). Данное описание состоит из направлений, расстояний и ориентиров.

В настоящее время развивается дифференциально-геометрический подход к разработке математических моделей. Этот подход включает замену фазового пространства дифференцируемыми многообразиями [Lewis, 2018]. Преимуществом данного подхода является унификация методов, используемых для анализа и синтеза систем управления для различных систем (линейных, нелинейных, пространственно-распределённых и т. д.). Наш случай отличается тем, что пространство управлений совершенно необязательно является дифференцируемым и имеет существенно более сложную структуру.

3. КЛАССИЧЕСКИЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ В ПРОСТРАНСТВАХ ВЫСОКОЙ РАЗМЕРНОСТИ (ИЛИ СО СЛОЖНЫМИ ПЕРЕСТРАИВАЕМЫМИ ОГРАНИЧЕНИЯМИ)

В ИИ весьма распространены задачи поиска в пространстве состояний, при решении которых используется перебор с различными методами сокращения вычислительной сложности этого процесса. Это игры, доказательства теорем, задачи планирования, задачи логического вывода и др. [Jones, 2008; Luger, 2001; Russell, Norvig, 2003]. Известно также, что человек решает подобные задачи без перебора, по крайней мере, полного. При этом число сравниваемых человеком вариантов очень невелико. Это достигается за счёт использования всего прошлого опыта человека плюс аналогии, иногда очень далёкие.

Отказываясь от предыдущего опыта, абстрагируясь от него, получаем вырожденную систему с множеством альтернатив, которые приходится перебирать. Даже придумали такое понятие — «проклятие размерности». Но это проклятие не размерности, а проклятие формализации, формальных систем. Как попытка использовать опыт человека должно рассматриваться использование различных приёмов эвристического программирования.

Многочисленные варианты случайного поиска в задачах ИИ (например, генетические алгоритмы и т.п.) — это попытка механического распространения свойств пространств низкой размерности на пространства высокой размерности. А это совсем иной мир со своими, часто очень странными и непривычными, законами.

Реально, в живой природе перебор не используется, поскольку это слишком «дорого» — система реального интеллекта заведомо погибнет раньше, чем сделает полный перебор. Тем более, что «плохой» выбор, сделанный вовремя, оказывается зачастую с точки зрения выживания организма лучше «хорошего», но сделанного с небольшим опозданием. Поэтому реально производится выбор из 2-3 альтернатив, имеющих смысл и апробированных практически — т.е. тех же прецедентов. При этом большая размерность реальных пространств исключает вырождение задачи, поскольку все варианты разные, а целевая функция тем более исключает равные оценки для рассматриваемых вариантов (которых и так немного).

Легендарный Буриданов осёл, умерший от голода потому, что никак не мог решить, какую охапку травы ему следует съесть, в природе невозможен в принципе. Поскольку в пространствах высокой размерности две охапки травы не могут быть одинаковыми, а также потому, что целевая функция сложна и нестабильна. Это тем более верно, что и сам поиск оптимума в реальности крайне затруднён и вместо оптимальности используется принцип приемлемости, т.е. решение должно быть удовлетворительным, и принято вовремя.

Более того, в случае каких-то проблем именно повышение размерности может спасти ситуацию. На протяжении всей истории человечество только тем и занято, что повышает размерность пространства — создаёт и вовлекает в оборот всё новые и новые инструменты, машины, приспособления, приборы.

4. ПРЕЦЕДЕНТЫ КАК РЕАЛИЗАЦИЯ ВЛОЖЕНИЙ

Что же помогает бороться реальным системам естественного интеллекта с «проклятием размерности»? А помогают бороться прецеденты. Прецеденты — это траектории движения в многомерном пространстве, а также порождающие их управления, которые сохраняются в той или иной форме в памяти системы. Это удачная реализация отображения пространства низкой размерности в пространство высокой размерности. Они сохраняются в памяти по той простой причине, что они не являются фатальными для её существования. Если прецедент является фатальным, то естественно, он не может быть использован, поскольку система больше не существует. Прецеденты существуют в пространствах очень высокой размерности и обладают тем свойством, что могут быть описаны параметрами, принадлежащими к пространствам гораздо меньшей размерности.

Не фатальными в большинстве случаев будут и небольшие вариации, что позволяет уже решать задачи *локальной* оптимизации. Хотя, естественно, существуют некие границы, нарушение которых переводит к катастрофической перестройке и фатальному исходу. В пространствах высокой и сверхвысокой размерности все процессы обработки информации и управления происходят в окрестности некоторого множества (возможно и достаточно большого) прецедентов.

Таким образом, прецеденты — это такие «оазисы» в пустыне, в которых только и может кипеть жизнь. Отыскание новых «оазисов» возможно, но крайне затруднено, поскольку большинство таких попыток закончится фатально. В этом причина того, что все древние сообщества людей были крайне консервативны в своей жизни, поскольку только такая консервативность и давала возможность выжить сообществу.

При этом на основе обработки прецедентов могут быть реализован весь комплекс задач, связанный с управлением, включая обучение и самообучение. Располагая некоторым множеством прецедентов, решающих одну задачу, можно осуществить локальную оптимизацию, путём обобщения формировать эвристические и продукционные правила (для экспертных систем), а также формировать программы управления. Сходные задачи могут решаться по аналогии с уже имеющимися прецедентами. Человек для решения задач широко использует аналогии, иногда весьма далёкие.

Следует также учитывать, что для прецедентов может использоваться и реально используется механизмы обобщения, т.е. отображения на более высокие, более абстрактные уровни представления информации. Таким образом реализуется механизм извлечения знаний, который позволяет существенно «сжать» необходимую для управления информацию и знания. При этом механизм аналогий позволяет для новых прецедентов использовать ссылки на фрагменты уже известных прецедентов, что стабилизирует объём баз данных и знаний и позволяет избежать их неконтролируемого роста в процессе функционирования системы. Это справедливо, естественно, для более или менее стабильной среды.

Вообще говоря, в последние годы происходит повышение интереса к «сильным» методам решения проблем (т.е. методам, основанным на знаниях предметной области, включая продукционные экспертные системы) и уменьшение интереса к «слабым» методам, решения проблем, основанным на формальных

логических системах. Это связано, в частности, с успешным использованием экспертных систем, основанных на знаниях.

Что касается прецедентов, то сейчас активно развиваются такие направления, как Case-based control (управление), Case-based reasoning (вывод), Case-based planning (планирование) [Aamodt, Plaza, 1994; Bao, 2008; Kolodner, 1992; Marling et al., 2002; Pal, Shiu, 2004]. При этом знания экспертов, которые аккумулируются в экспертных системах, являются ничем иным, как результатом обработки, обобщения, верификации некоторого множества прецедентов.

ЗАКЛЮЧЕНИЕ

Мир искусственного интеллекта — мир высоких размерностей. Таким образом, управление в пространстве высоких размерностей решается путём формирования набора прецедентов. Причём, объём фазового пространства, занимаемого этими прецедентами, очень мал по сравнению со всем возможным пространством возможных управлений, в которое при нормальном функционировании объект никогда не попадёт. Это является одним из существенных отличий управление в пространстве больших и сверхбольших размерностей. Тем не менее, большинство явлений, характерных для классических СУ, находят прямые аналогии и в системах ИИ. В статье кратко рассмотрен вопрос размерности задач управления, решаемых ИИ. Решение задач тесной интеграции классической теории СУ и методов ИИ требует анализа целого комплекса вопросов, который невозможно провести в рамках одной статьи.

ЛИТЕРАТУРА

- [Дружинин, Конторов, 1982] Дружинин В. В., Конторов Д. С. Конфликтная радиолокация. М.: Радио и связь, 1982.
- [Дружинин и др., 1989] Дружинин В. В., Конторов Д. С., Конторов М. Д. Введение в теорию конфликта. М.: Радио и связь, 1989.
- [Aamodt, Plaza, 1994] Aamodt A., Plaza E. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches // AI Communications. 1994. V. 7. Iss. 1. P. 39–59.
- [Adversarial..., 2007] Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind / eds. Kott A., McEneaney W. M. Ser.: Computer And Information Science. Boca Raton, FL: Chapman and Hall/CRC, Taylor and Francis Group, 2007. 97 p.
- [Bao, 2008] Bao D. A Survey on Similarity-based Reasoning. 2008. URL: <https://pdfs.semanticscholar.org/7800/f05fa44fda7a015bd74045845cd8317ce4ae.pdf>.
- [Embodied..., 2004] Embodied Artificial Intelligence / eds. Iida F., Pfeifer R., Steels L., Kuniyoshi Y. Intern. Seminar, Dagstuhl Castle, Germany, July 7–11, 2003. Revised Selected Papers; Ser.: Lecture Notes in Artificial Intelligence 3139; Subseries: Lecture Notes in Computer Science; Springer-Verlag: Berlin, 2004.
- [Grishin, 2012] Grishin V. Application of Artificial Intelligence Methods in Unmanned Aerial Vehicle Control // J. Information, Intelligence and Knowledge (JIИK). 2012. V. 3. Iss. 4.
- [Hopgood, 2001] Hopgood A. A. Intelligent Systems for Engineers and Scientists. 2nd ed. Boca Raton, FL: CRC Press LLC, 2001.
- [Jones, 2008] Jones M. T. Artificial Intelligence: A Systems Approach. Hingham, MA: Infinity Science Press LLC, 2008.

- [Kolodner, 1992] *Kolodner J. L.* An Introduction to Case-Based Reasoning // Artificial Intelligence Review. 1992. V. 6. P. 3–34.
- [Lefebvre, 1999] *Lefebvre V.* Sketch of Reflexive Game Theory // Proc. Workshop on Multi-Reflexive Models of Agent Behavior. Los Alamos, NM. 1999. P. 11–40.
- [Lewis, 2018] *Lewis A. D.* The bountiful intersection of differential geometry, mechanics, and control theory // Annual Review of Control, Robotics, and Autonomous Systems. 2018. V. 1. P. 135–158.
- [Luger, 2002] *Luger G. F.* Artificial Intelligence. Structures and Strategies for Complex Problem Solving. 4th ed. L.: Addison Wesley, 2002.
- [Marling et al., 2002] *Marling C., Sqalli M., Risland E., Muñoz-Avila H., Aha D.* Case-Based Reasoning Integrations // AI Magazine. 2002. V. 23. No. 1. P. 69–86.
- [Pal, Shiu, 2004] *Pal S. K., Shiu S. C. K.* Foundations of Soft Case-Based Reasoning. Hoboken, NJ: John Wiley and Sons, Inc., 2004.
- [Russell, Norvig, 2003] *Russell S. J., Norvig P.* Artificial Intelligence: A modern Approach. 2nd ed. Upper Saddle River, NJ: Prentice Hall, Pearson Education Inc., 2003.
- [Shemayev, 2007] *Shemayev V. N.* Cognitive Approach to Modeling Reflexive Control in Socio-Economic Systems // Information and Security. 2007. V. 22. P. 28–37.
- [Thomas, 2002] *Thomas T. L.* Reflexive Control in Russia: Theory and Military Applications: Reflexive Processes and Control // Intern. Interdisciplinary Scientific and Practical J. 2002. V. 1. No. 2. P. 60–76.
- [Thomas, 2004] *Thomas T. L.* Russia's Reflexive Control Theory and the Military // J. Slavic Military Studies. 2004. V. 17. Iss. 2. P. 237–256.
- [Zorich, 2011] *Zorich V. A.* Mathematical Analysis of Problems in the Natural Sciences. N. Y.: Springer Verlag, 2011.

CLASSICAL CONTROL THEORY AND METHODS OF ARTIFICIAL INTELLIGENCE AND THE CURSE OF DIMENSIONALITY

V. A. Grishin

Space Research Institute of the Russian Academy of Sciences (IKI RAN)
Bauman Moscow State Technical University (BMSTU)

Artificial intelligence increasingly intrudes into many aspects our lives. This process is natural and can be ignored. This article is an attempt to approach the solution of the complicated problem of the close integration of classical control theory and methods of artificial intelligence. Dimensions of the phase space and space of control parameters are considered.

Keywords: control system theory, artificial intelligence, curse of dimensionality

Grishin Vladimir Aleksandrovich — senior scientist, PhD, associate professor, vgrishin@iki.rssi.ru

КВАНТОВЫЕ СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

М. В. Алтайский, Н. Н. Зольникова

Институт космических исследований Российской академии наук (ИКИ РАН)

В работе рассматриваются современные направления развития квантовых систем искусственного интеллекта, сочетающих в себе концепцию машинного обучения на основе искусственных квантовых нейронных сетей и идеи квантовых вычислений, обеспечивающие экспоненциальное ускорение решения задач оптимизации за счёт квантового параллелизма. В работе рассмотрены как уже существующие адиабатические квантовые компьютеры (нейронные сети) на основе сквидов и оптических сетей, так и оригинальные результаты авторов в области разработки и создания элементной базы для адиабатических квантовых компьютеров будущего. Рассматриваются перспективы создания портативных квантовых систем искусственного интеллекта, годных для бортового применения, которые в будущем могли бы быть использованы для управления автономными космическими аппаратами, а также для высокопроизводительной обработки изображений в космических исследованиях и задачах мониторинга Земли из космоса.

Ключевые слова: квантовое обучение, нейронные сети, искусственный интеллект

ВВЕДЕНИЕ

Понятие квантового искусственного интеллекта (ИИ) выросло из описания систем искусственного интеллекта, реализуемых программным образом на обычных компьютерах, или выполненных аппаратно, на базе специализированной микроэлектроники, когда стало ясно, что возможности электронных вычислительных устройств при обработке больших данных существенно ограничены быстродействием этих компьютеров при выполнении стандартных операций булевой алгебры. К этому времени уже были сформулированы основные идеи квантовых вычислений, открывающих возможность решения экспоненциально сложных задач за полиномиальное время за счёт квантового параллелизма [Deutsch, Jozsa, 1992].

С точки зрения практической полезности, повышение скорости вычислений при выполнении тех или иных алгоритмов не всегда является самоцелью. Значительно более важной задачей может быть замена человеческого интеллекта системами искусственного интеллекта в различных технических системах. Это может быть управление автономными летательными аппаратами, контроль производственных процессов, распознавание лиц и поведения в системах безопасности.

Действия систем ИИ часто не являются алгоритмическими, а представляют собой некоторое отображение «выход = f (вход)», в котором функция $f(\cdot)$ формируется путём обучения. Простейшая модель обучающейся системы —

Алтайский Михаил Викторович — ведущий научный сотрудник, доктор физико-математических наук, altaisky@iki.rssi.ru

Зольникова Надежда Николаевна — научный сотрудник, кандидат физико-математических наук, nzolnik@iki.rssi.ru

модель нейронной сети — была построена на основе физиологических исследований Хебба [Hebb, 1949]. Нейронная сеть представляет собой массив одинаковых элементов (нейронов), связанных между собой настраиваемыми связями (весами). Каждый нейрон может находиться либо в основном, либо в возбуждённом состоянии, что определяется суммой всех сигналов от других нейронов, подаваемых на его вход. При достаточном числе нейронов в массиве подбор связей, называемый процессом обучения, может обеспечить любое желаемое отображение [Haykin, 1999]. Подбор весов путём предъявления сети пар вход-выход из обучающей выборки называется обучением с учителем. Примерно так работает нейронная сеть мозга. Если подобным образом организован массив одинаковых элементов, создаваемый искусственно, в целях получения желаемого отображения, то такую сеть называют искусственной нейронной сетью (ИНС).

Хотя ИНС вполне успешно справляются со многими простыми задачами, такими как, например, распознавание изображений и сигналов, при решении более сложных задач, время обучения экспоненциально растёт с размером обучающей выборки, что может сделать применение сетей практически неприемлемым. Вполне логичным здесь представляется использование квантового параллелизма для ускорения процесса обучения. Однако, в силу того, что работа нейронной сети, в большинстве случаев, не является алгоритмической, использование квантового параллелизма для увеличения её быстродействия, требует других подходов, отличных от алгоритмических методов квантовых вычислений. Необходимость новых методов связана как со случайным характером обучения, так и с открытостью самой сети. Последнее особенно важно в том смысле, что в отличие от сетевого квантового компьютера, представляющего собой устройство, изолированное от окружающей среды на время вычислений, нейронная сеть обучается путём взаимодействия с окружением, и следовательно не может быть изолирована.

В данном обзоре представлены как результаты, полученные в данном направлении за последние десять лет, так и результаты исследований авторов, направленные на создание квантовых нейронных сетей, работающих при обычных температурах.

1. Алгоритмические и неалгоритмические системы

Понятие искусственного интеллекта появилось в связи с попытками перенести часть когнитивных функций человека на искусственно создаваемые устройства — роботы, *intelligent agents*, — устройства, которые воспринимают (отображают) своё окружение и, исходя из полученных отображений, могут производить такие действия, которые максимизируют вероятность достижения поставленных целей. Системы ИИ предназначены для решения конкретных задач. Ставить цели и задачи системы ИИ не могут. Примером системы такого типа является автопилот, задачей которого является обеспечение безопасного передвижения воздушного судна из точки P_1 в точку P_2 , при заданных параметрах полёта.

Традиционными задачами систем ИИ, в широком смысле, являются восприятие, распознавание сообщений и объектов, обучение, представление знаний, логический вывод. В узком смысле системы ИИ предназначены для решения задач оптимизации в присутствии случайных параметров.

Существуют алгоритмические и неалгоритмические методы решения задач оптимизации. Так, для нахождения основного состояния гамильтоновой системы с потенциалом взаимодействия $U(x)$, формально просто найти минимум этого потенциала путём удовлетворения условий $U'(x_*) = 0, U''(x_*) > 0$, или применить метод градиентного спуска. Аналогичным образом, используя уравнения механики, являющиеся следствием принципа наименьшего действия, можно рассчитать траекторию космического аппарата при заданных начальных условиях.

Полностью алгоритмические методы не могут быть применены для исчерпывающего решения любой практической задачи по нескольким причинам:

1. Невозможность полностью изолировать систему от меняющегося окружения. Эволюция системы определяется не только гамильтонианом самой системы, но и влиянием на неё окружения.
2. Существование динамического хаоса даже в относительно простых системах уравнений, когда малое отклонение начальных условий приводит к большому разбеганию траекторий.
3. Даже если предположить существование уравнений или функционалов, полностью описывающих эволюцию системы, их вид может оказаться настолько сложным, что аппроксимация отображения «вход → выход», с фиксированной точностью, может оказаться алгоритмически намного более эффективной, чем прямое вычисление на классическом компьютере, длительность которого может сравняться с возрастом Вселенной.

Развитие систем ИИ шло путём подражания поведению живых существ, которые, естественно, не производят алгоритмических вычислений, а действуют по принципу обучения. При этом их действия, например по распознаванию лиц, часто оказываются более эффективными чем любые алгоритмы, выполняемые на современных компьютерах. Математическую модель таких действий, возникшую в результате нейрофизиологических исследований Хебба [Hebb, 1949], называют нейронной сетью.

2. Машинное обучение

Реализованная программным образом экспертная система может выполнять долгие и трудоёмкие вычисления для решения различных проблем. Но, в отличие от человека, если дать ей такую же или подобную проблему второй раз, она не «вспомнит» решение, а каждый раз вновь будет выполнять те же вычисления. Применение большинства экспертных систем ограничено негибкостью их стратегии принятия решений и трудностью модификации больших объёмов кода.

Очевидное решение этих проблем состоит в создании систем, которые могли бы учиться на своём опыте, аналогиях или примерах. Таким образом, возникает *машинное обучение* — процесс, в результате которого машина (компьютер) способна показывать поведение, которое в неё не было явно заложено (запрограммировано). Основные типы задач машинного обучения — это обучение с учителем (классификация, регрессия и др.), обучение без учителя (кластеризация, визуализация данных, понижение размерности и др.), и обучение с подкреплением [Mitchell, 1997].



Рис. 1. Схема машинного обучения

Обучение с учителем. Пусть имеется множество входных данных X . Это могут быть различные свойства, характеристики и признаки объектов. Пусть Y — множество выходных данных (ответов, откликов, меток и др.). При этом каждому элементу $x \in X$ по некоторому (детерминированному или вероятностному) закону ставится в соответствие элемент $y \in Y$. Задача обучения с учителем — аппроксимировать эту зависимость, т.е. по данным из обучающей выборки с элементами $(x^{(i)}, y^{(i)})$, где $i = 1, \dots, M$, выработать некое решающее правило $y = h(x)$, по которому для каждого нового элемента $x \in X$ будет

найдено соответствующее значение $y \in Y$. Пара $(x^{(i)}, y^{(i)})$ называется обучающим примером, а вся совокупность обучающих примеров — обучающей выборкой.

Линейная регрессия. Если по условиям задачи решающая функция $h(x)$ может быть выбрана в виде линейной комбинации входных данных

$$h(x) = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n,$$

то задача машинного обучения сводится к задаче линейной регрессии, так что весовые коэффициенты w_i могут быть найдены минимизацией квадратичного функционала ошибки

$$J(w) = \frac{1}{2M} \sum_{i=1}^M \left(h(x^{(i)}) - y^{(i)} \right)^2. \quad (1)$$

Поиск минимума функции $J(w)$ является линейной задачей, однако при больших размерностях вектора входных данных $n \gtrsim 1000$ более эффективным может оказаться использование метода градиентного спуска. Алгоритм градиентного спуска заключается в следующем. Задаются случайным образом начальные значения весов, а затем производится их одновременное итеративное обновление по формуле

$$w_j^{(t+1)} = w_j^{(t)} - \alpha \frac{\partial J \left\{ w_j^{(t)} \right\}}{\partial w_j}, \quad (2)$$

где t — номер итерации. От выбора параметра α зависит то, как быстро будет сходиться градиентный спуск, и будет ли сходиться вообще. Если α слишком мало, то процесс сходимости будет довольно длительным, а если велико, то можно проскочить минимум, и процесс либо будет расходящимся, либо сходиться, но также очень медленно. Для контроля за правильностью работы алгоритма надо следить, чтобы на каждой итерации функция стоимости убывала. Градиентный спуск носит пакетный характер, т.е. на каждой итерации в обновлении весов принимают все обучающие примеры. Существует вариант стохастического градиентного спуска, когда веса обновляются последовательно для каждого обучающего примера, или для их небольших выборок (mini batch). В общем случае оптимизационные задачи очень чувствительны к выбору начальных условий, и численный метод легко может застрять в локальном

минимуме, но в случае линейной регрессии функция стоимости $J(w)$ является выпуклой и имеет единственный, глобальный, минимум. В случае, если входных признаков больше одного, необходимо вначале провести их предварительную обработку, т.е. нормировать их значения таким образом, чтобы они находились в диапазоне $[-1, 1]$. Это надо сделать для того, чтобы избежать овражности целевой функции, что очень резко замедляет сходимость численных методов.

Задача классификации. В задаче классификации нужно по признакам объекта определить его принадлежность одному из заданных классов. Причём названия классов очень часто носят категоричный характер. В этом случае каждой категории присваивают целое число. Например,

- классификация электронных писем по принципу спам/не спам, здесь $y = 1, 0$;
- классификация опухолей — злокачественная/доброкачественная, $y = 1, 0$;
- предсказание погоды на следующий день — дождь/облачно/туман/солнце, $y = 1, 2, 3, 4$.

Таким образом, в отличие от задачи регрессии, выходные значения задачи классификации принадлежат множеству целых чисел.

К задачам классификации относятся задачи распознавания текста, распознавание типов самолётов по изображениям, распознавание речи и т.п.

3. НЕЙРОННЫЕ СЕТИ

3.1. Биологический нейрон

Понимание того, что мышление каким-то образом связано с мозгом, существовало в течение тысяч лет, однако только во второй половине 19-го века появился метод, позволяющий наблюдать за отдельными нейронами мозга. Каждый нейрон состоит из **тела клетки** (или **сомы**), которое содержит ядро клетки. От тела клетки ответвляется множество коротких волокон, называемых **дендритами**, и одно длинное волокно, называемое **аксоном**. Вид биологического нейрона под электронным микроскопом показан на рис. 2. Аксон растягивается на большое расстояние, которое может достигать одного метра. Нейрон создаёт соединения с другими нейронами, количество которых может составлять от 10 до 100 000 в точках сопряжения, называемых **синапсами**. Сигналы распространяются от одного нейрона к другому с помощью сложной электрохимической реакции.

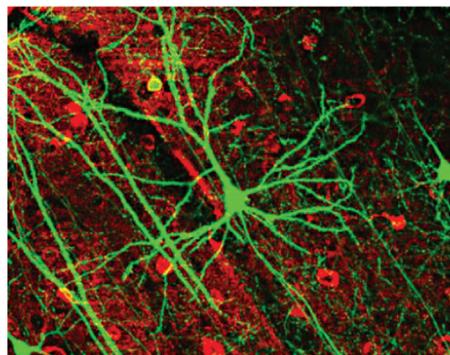


Рис. 2. Пирамидный нейрон коры головного мозга мыши, экспрессирующий зелёный флуоресцентный белок (из работы [Lee et al., 2006])

Эти сигналы управляют активностью мозга в течение короткого интервала времени, а также становятся причиной долговременных изменений состояния самих нейронов и их соединений. Импульсы, поступившие к нейрону одновременно по нескольким дендритам, суммируются. Если суммарный импульс превышает некоторый порог, нейрон возбуждается, формирует собственный импульс и передаёт его далее по аксону. Важно отметить, что веса синапсов могут изменяться со временем, а значит, меняется и поведение соответствующего нейрона. Считается, что эти механизмы служат мозгу основой для обучения. Обработка информации главным образом происходит в коре головного мозга.

3.2. Математическая модель нейрона

МакКаллош и Питс предложили следующую модель нейрона [McCulloch, Pitts, 1943]. Пусть к m синапсам нейрона поступают входные сигналы x_1, \dots, x_m , соответственно. Тогда, после прохождения синапсов и дендритов, к нейрону поступают сигналы $w_j x_j$. Они суммируются, сдвигаются на постоянную величину b , и преобразуются с помощью некоторой функции активации (передаточной функции) $\phi(\cdot)$. Таким образом выходной сигнал нейрона, выдаваемый в аксон, может быть записан в виде:

$$y_k = \phi(u_k), \text{ где } u_k = \sum_{j=1}^m w_{kj} x_j + b_k, \tag{3}$$

здесь k — метка нейрона, а сумма u_k называется потенциалом активации нейрона.

Блок-схема классического нейрона приведена на рис. 3.

Традиционно, в качестве модельной функции активации используется либо ступенчатая функция

$$\phi(z) = \begin{cases} 0 & z < 0 \\ 1 & z \leq 0, \end{cases}$$

либо функция с насыщением, типа логистической функции

$$f(z) = \frac{1}{1 + \exp(-az)}, \quad a > 0.$$

Современные исследования, на компьютерных моделях многослойных нейронных сетей с большим числом нейронов, показывают, что весьма эффективным является использование кусочно-линейной функции $\phi(z) = \max(0, az)$ [LeCun et al., 2015].

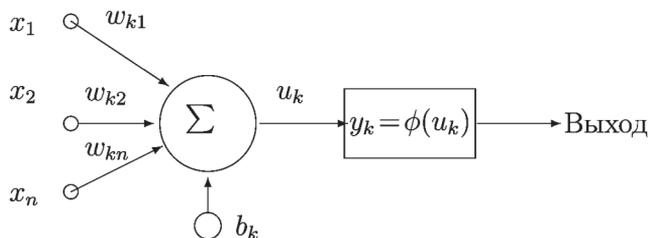


Рис. 3. Математическая модель нейрона

Математическая модель (3) является весьма упрощённой. Она не учитывает стохастических аспектов работы нейрона, к тому же выход реального биологического нейрона кодируется по частоте, а не по амплитуде. Тем не менее, эта простая математическая модель позволила Розенблату построить реально работающую компьютерную модель классификатора, способного распознавать, например, буквы латинского алфавита [Rosenblatt, 1958]. В качестве алгоритма обучения при этом было использовано простое правило: при неправильном выходе веса сети изменялись на величину входного сигнала, взятого с соответствующим знаком. Начальные значения весов при этом задаются случайным образом в диапазоне $(-1, 1)$.

3.3. Типы нейронных сетей

Аксон биологического нейрона имеет многочисленные ответвления — синаптические терминалы, — с помощью которых выходной сигнал одного нейрона может быть подан на входы многих других нейронов. Так образуется *нейронная сеть*. Согласно физиологическому правилу Хебба, если возбуждения двух нейронов часто коррелируют, то синаптическая связь между ними усиливается $\Delta w_{ij} = \eta y_i y_j$. Подобным образом формируются и искусственные нейронные сети. Для определения ИНС нужно задать:

- какова размерность вектора входных данных — число входных нейронов,
- какова размерность вектора выходных данных — число выходных нейронов,
- сколько нейронов задействовано в сети и как они соединены между собой — топология нейронной сети.

В простых моделях считается, что информация распространяется в нейронной сети в одном направлении — от входа к выходу. В этом случае нейронная сеть представляет собой направленный граф, в котором нейроны представлены вершинами графа, а синаптические связи — рёбрами графа. После того, как выбрана топология ИНС, т.е. соответствующий ей граф, приступают к обучению нейронной сети, т.е. к подбору весов w_{ij} , оптимальным образом обеспечивающих решение поставленной задачи классификации или регрессии. Топологию ИНС обычно изображают планарным графом (рис. 4).

Слой нейронной сети, на которые не подаются входные данные и не снимается выход, называют *скрытыми*.

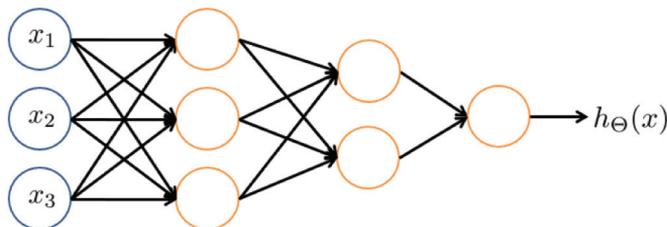


Рис. 4. Прямоточная нейронная сеть типа перцептрон, содержащая три входных нейрона и один выходной. Скрытый слой сети содержит два нейрона

Существует несколько десятков различных архитектур нейронных сетей, причём эффективность многих из них доказана математически. Наиболее популярные и изученные архитектуры — это:

- однослойные сети прямого распространения — состоят только из входов и выходов, информация передаётся от входов к выходам,
- многослойные сети прямого распространения — кроме входов и выходов есть один или несколько скрытых слоёв, информация передаётся в направлении от входов к выходам,
- рекуррентные сети — отличаются от предыдущих наличием обратной связи, передающей информацию к предыдущим слоям.

Рекуррентные сети отличаются от однонаправленных тем, что сигнал с выходных нейронов может подаваться на вход одного из предыдущих слоёв. Это во многих аспектах более адекватно описывает динамику реальных биологических сетей нейронов, чем прямоточные сети, и может приводить к весьма сложному динамическому поведению. Если в такую систему введено некое правило изменения весов при предъявлении каждого нового вектора данных (например, winner takes all), то такая сеть может быть использована для классификации поступающих входных данных без привлечения обучающих примеров, т. е. для спонтанной кластеризации данных.

В сложных рекуррентных сетях может не быть выходных нейронов как таковых. Вместо этого существует часть интерфейсных нейронов посредством которых происходит обмен информацией между нейронной сетью и её окружением, как в ту, так и в другую сторону, и часть внутренних (скрытых) нейронов, испытывающих сложное динамическое поведение под воздействием как интерфейсных нейронов, так и динамики остальной сети.

Схематически, динамика такой сети может быть описана динамическими уравнениями типа

$$C_j u_j(t) = -\frac{u_j(t)}{R_j} + \sum_{i,j=1}^N w_{ji} \Phi_j(u_i(t)) + I_j, \quad (4)$$

где C_j , R_j , I_j — ёмкость, сопротивление и ток, связанные с j -м нейроном. Исследование динамического поведения в большой системе уравнений типа (4) может быть заменено термодинамическим описанием системы векторов $\{x_j^{(k)}\}_k$ с ляпуновской функцией энергии

$$E = -\frac{1}{2} \sum_{i,j=1}^N w_{ij} x_i x_j + \sum_{j=1}^N \frac{1}{R_j} \int_0^{x_j} \Phi_j^{-1}(x) dx - \sum_{j=1}^N I_j x_j.$$

Такая сеть, например, может быть использована для «запоминания» входных данных путём подбора весов $w_{ij} = w_{ji}$ таким образом, чтобы бoльцмановская функция вероятности

$$P(x) \propto \exp\left\{-\frac{\sum_{ij} w_{ij} x_i x_j}{2T}\right\}$$

выдавала различные вектора данных x с теми же вероятностями, с какими они встречались в представленной выборке. Состояния, реализующие минимум энергии будут при этом появляться с максимальной вероятностью. Это так называемая машина Больцмана.

3.4. Обучение нейронных сетей с учителем

Для обучения ИНС с учителем нужно таким образом подобрать веса нейронной сети w , чтобы, с одной стороны, минимизировать ошибку на обучающей выборке

$$\min \frac{1}{2M} \sum_{i=1}^M (y(x^{(i)}) - d^{(i)})^2,$$

а с другой — обеспечить должную степень обобщения, чтобы малые отклонения во входных данных не приводили к ошибкам классификации. Обучению ИНС посвящена обширная литература, см., например [Осовский, 2004; Haykin, 1999].

В случае одного нейрона (3), простейшим методом обучения является итерационное правило

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta [d(n) - y(n)] \mathbf{x},$$

последовательно применяемое к каждому вектору \mathbf{x} из обучающей выборки.

В случае сложной многослойной нейронной сети минимизацию можно проводить методом градиентного спуска

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta \frac{\partial \mathcal{E}(\mathbf{w})}{\partial \mathbf{w}},$$

где \mathbf{w} — вся совокупность весов сети.

4. КВАНТОВЫЕ ВЫЧИСЛЕНИЯ

4.1. Квантовые компьютеры

Классические компьютеры работают с регистрами *битов* — физических систем, которые могут находиться в одном из двух взаимоисключающих состояний, обозначаемых, для определённости, «1» и «0». Обычно этими состояниями являются наличие и отсутствие тока в транзисторе, в случае электронной схемы, и две взаимно ортогональные поляризации фотонов, в случае оптической схемы. Всякая неопределённость состояния отдельного бита в регистре может привести к ошибке вычисления, и должна быть устранена либо путём дублирования, либо с помощью исправляющих ошибки кодов.

Переход элементной базы вычислительных устройств на нанометровые масштабы привёл к тому, что физическая система, содержащая один бит информации, скажем одноэлектронный транзистор, уже не может быть описана классически. Для её описания необходимо привлекать законы квантовой механики. Это означает, что каждый бит может находиться состояниях «1» и «0» с отличными от нуля и единицы вероятностями. Для классических микропроцессоров, функционирующих по правилам булевой логики, отличие вероятности выходного состояния от единицы является паразитным эффектом. Архитектура микропроцессоров разрабатывается таким образом, чтобы это отличие было пренебрежимо мало. Этого можно достичь при технологиях в сотни и даже десятки нанометров.

На масштабах в единицы нанометров квантовой неопределённостью, являющейся следствием принципа неопределённости Гейзенберга, уже нельзя пренебречь. Состояние каждого бита приходится описывать с помощью вектора квантового состояния:

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle, \quad |0\rangle, |1\rangle \in \mathcal{H}, \quad c_0, c_1 \in \mathbb{C}, \quad |c_0|^2 + |c_1|^2 = 1, \quad (5)$$

где гильбертово пространство \mathcal{H} — комплексное линейное пространство со скалярным произведением $\langle \cdot | \cdot \rangle$. Вероятность обнаружения бита в одном из двух ортогональных квантовых состояний, $\langle 0|0\rangle = 1 = \langle 1|1\rangle$, $\langle 0|1\rangle = 0$, определяется квадратами модулей амплитуд состояния:

$$P(0') = |c_0|^2, \quad P(1') = |c_1|^2.$$

Такой бит, эволюция которого подчиняется уравнению Шрёдингера

$$i \frac{\partial}{\partial t} |\psi\rangle = \hat{H} |\psi\rangle, \quad (6)$$

называют квантовым битом, или *кубитом*.

В реальности, дело обстоит ещё сложнее: наряду с квантовой природой кубита приходится учитывать его взаимодействие с флуктуирующим окружением (термостатом) в терминах матрицы плотности, согласно методам теории открытых квантовых систем [Breuer, Petruccione, 2002; Feynman, Vernon, 1963]. Однако уже представление (5) позволило расширить известную теорию информации до *теории квантовой информации* [Nielsen, Chuang, 2000].

Основная идея квантовой теории информации состоит в том, что большой класс задач, требующий для своего решения на классическом компьютере ресурсов/времени, экспоненциальных по размеру входных данных, может быть решён значительно быстрее (за полиномиальное время), если вместо эволюции классического регистра по правилам булевой логики использовать эволюцию квантового регистра, т. е. *регистра, состоящего из кубитов*, по правилам квантовой механики. Такой компьютер называют квантовым компьютером [Stolze, Suter, 2008].

4.2. Квантовые алгоритмы

Квантовые компьютеры, равно как и классические компьютеры, могут быть построены по принципу машины Тьюринга и выполнять любую предписанную последовательность операций с битами регистра, классическими или квантовыми [Deutsch, 1985]. Если при этом используется булева алгебра операций, то алгоритм остаётся классическим, независимо от того, какие биты при этом используются, классические или квантовые. При этом операции, проводимые над регистром, не являются обратимыми в смысле квантовой механики, а время и ресурсы, необходимые для выполнения алгоритма, имеют примерно один и тот же порядок для квантовых и классических регистров.

Получить превосходство в эффективности вычислений над классическими алгоритмами можно в том случае, если вместо классических булевских операций использовать квантомеханическую эволюцию регистра — для этого необходимо состоящего из квантовых битов. Унитарная эволюция квантового регистра (если нет взаимодействия с окружением) является обратимой во времени, а все возможные квантовые состояния регистра могут обрабатываться

параллельно, если для этого используются специализированные устройства, называемые *квантовыми гейтами*, или другое специализированное оборудование, функционирующее по законам квантовой механики. Таким образом, квантовый алгоритм — это такой алгоритм, для выполнения которого существенно требуется наличие квантового компьютера.

Наиболее известными к настоящему времени квантовыми алгоритмами являются алгоритм поиска в неупорядоченной базе данных (алгоритм Гровера) [Grover, 1997], и алгоритм разложения числа на простые множители с помощью квантового преобразования Фурье (алгоритм Шора) [Shor, 1994]. Оба этих алгоритма являются идеализацией. В них приготовленное начальное состояние регистра подвергается унитарной эволюции с помощью применяемых к наборам кубитов унитарных операций, осуществляемых квантовыми гейтами.

Унитарность всего процесса вычислений, от приготовления начального состояния до измерения конечного состояния — получения результата, — обеспечивается изоляцией квантового компьютера от флуктуирующего окружения. Последнее является лишь теоретической идеализацией: обеспечить полную изоляцию реальной физической системы от окружения невозможно. По этой причине, в дополнение к устоявшимся и вошедшим в учебники квантовым схемам [Nielsen, Chuang, 2000], необходимо искать новые схемы, в которых учитывалось бы взаимодействие квантового компьютера с окружением. Делается это на основе теории открытых квантовых систем [Breuer, Petruccione, 2002; Feynman, Vernon, 1963].

4.2.1. Алгоритм Гровера

При поиске в неупорядоченной базе данных длиной в N записей, например поиске фамилии абонента по номеру в телефонном справочнике, классическими методами в среднем требуется $N/2$ просмотров записей. На квантовом языке это означает, что к каждой из $N = 2^n$ записей, индексируемых с помощью метки $|x\rangle$, записанной в n -кубитовом регистре, применяется некоторый оператор \mathbf{O} (от англ. *Oracle*):

$$\mathbf{O}|x\rangle = (-1)^{f(x)}|x\rangle,$$

где функция $f(x)$ принимает значение 1 если x удовлетворяет условию поиска, и значение 0 в противном случае. Благодаря квантовой механике, можно приготовить равновзвешенную квантовую суперпозицию всех меток $|x\rangle$

$$|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle, \quad (7)$$

и применять квантовые гейты не к каждой из записей, а к линейной суперпозиции (7). С помощью оператора \mathbf{O} можно менять знак у тех членов линейной суперпозиции, которые удовлетворяют условию поиска $f(x)=1$. Используя единичную нормировку векторов состояния, алгоритм можно построить таким образом, чтобы на каждой итерации возрастала норма члена, удовлетворяющего условию поиска, а норма остальных членов падала. Отсылая читателя к стандартной литературе, касающейся деталей алгоритма [Grover, 1997; Nielsen, Chuang, 2000; Stolze, Suter, 2008], отметим лишь, что для обнаружения необходимой записи в среднем оказывается достаточно $\sim \sqrt{N}$ итераций, вместо $N/2$ итераций классического поиска.

4.2.2. Алгоритм Шора

Алгоритм разложения большого натурального числа N на простые множители является ключевым алгоритмом шифрования, применяемым в различных системах передачи данных. С формальной точки зрения, задача факторизации числа N может быть сведена к нахождению периода функции модулярной экспоненты $F_N(x+r) = F_N(x)$, где

$$F_N(x) := a^x \bmod N, \quad (8)$$

т.е. такого минимального целого числа r , что $a^r \bmod N = 1$. Используя бинарное разложение $k = \sum_{j=0}^{K-1} k_j 2^j$ и базис квантовых состояний $(|0\rangle, |1\rangle)$ любое натуральное число можно представить в виде состояния квантового регистра при достаточно большом числе кубитов (K). Период же функции удобно искать с помощью преобразования Фурье: при наличии периодичности Фурье образ исследуемой функции содержит пики, отвечающие периодам функции.

Для поиска периода модулярной экспоненты (8) на квантовом компьютере Шором был предложен следующий алгоритм [Shor, 1994]. Исследуемой функции $f(x)$ ставится в соответствие унитарный оператор $U_f |x, y\rangle := |x, y \oplus f(x)\rangle$. Для этого используются два регистра, обозначаемые, для определённости, x и y : $|x, y\rangle \equiv |x\rangle \otimes |y\rangle$. В случае задачи поиска периода модулярной экспоненты длина регистра x составляет K кубитов, $N^2 \leq 2^K < 2N^2$, длина регистра y должна быть $n \geq \log_2 N$.

1. Готовится нулевое начальное состояние регистров $|\psi_1\rangle = |0, 0\rangle$.
2. К регистру x применяется преобразование Фурье:

$$\hat{F}|x\rangle := \frac{1}{\sqrt{2^K}} \sum_{k=0}^{2^K-1} e^{i2\pi(kx/2^K)} |k\rangle,$$

которое переводит любое состояние регистра в суперпозицию всех возможных состояний с различными фазовыми множителями. Поскольку регистр x находился в состоянии $|0\rangle$, то все фазовые множители оказываются равными 1, и мы получаем равновзвешенную суперпозицию всех состояний:

$$|\psi_2\rangle = (\hat{F} \otimes \mathbf{1})|\psi_1\rangle = \frac{1}{\sqrt{2^K}} \sum_{k=0}^{2^K-1} |k\rangle |0\rangle.$$

3. Преобразованием, отвечающим модулярной экспоненте, действуем на регистр y , не изменяя состояния регистра x :

$$|\psi_3\rangle = (\mathbf{1} \otimes U_{F_N})|\psi_2\rangle = \frac{1}{\sqrt{2^K}} \sum_{k=0}^{2^K-1} |k\rangle |a^k \bmod N\rangle.$$

4. Применяем к регистру x преобразование Фурье:

$$|\psi_4\rangle = (\hat{F} \otimes \mathbf{1})|\psi_3\rangle = \frac{1}{2^K} \sum_{k,q=0}^{2^K-1} e^{i2\pi(kq/2^K)} |q\rangle |a^k \bmod N\rangle.$$

5. Вероятность обнаружения различных состояний пары регистров $|x\rangle|y\rangle$ теперь определяется квадратом модуля вектора $|\psi_4\rangle$. Измеряя вероятность различных состояний теперь можно найти период r как период повторяемости пиков вероятности.

5. КВАНТОВЫЕ НЕЙРОННЫЕ СЕТИ

Идея создания искусственной квантовой нейронной сети, предложенная в работе [Как, 1995], состоит в замене классических сигналов, поступающих на вход нейронов, на квантовые состояния, обладающие амплитудой и фазой. При этом на выходе нейрона также должно формироваться квантовое состояние, зависящее от линейной суперпозиции входящих состояний. Веса, в случае квантовой нейронной сети являются комплексными числами (которые изменяются в процессе обучения сети), так что каждое входное квантовое состояние не только взвешивается по амплитуде, но и сдвигается по фазе. Использование фазы сигнала в нейросетевых моделях является красивой математической находкой, и связано с обобщением метода обратного распространения ошибки, используемого для обучения обычного классического перцептрона, на нейронные сети с комплексными весовыми коэффициентами [Nitta, 1997]. В этом смысле, без учёта нелинейных эффектов, квантовые нейронные сети (КНС) идентичны оптическим нейронным сетям [Shariv, Friesem, 1989]. Тем не менее, фазовый множитель может влиять и на вероятность подбарьерного туннелирования ионов в некоторых моделях биологических нейронов [Altaisky, Rao, 2009]. Следует отметить, что идея того, что именно фазовые соотношения между квантовомеханическими волновыми функциями отдельных элементов нервных волокон лежат в основе работы мозга человека была впервые высказана в работах В. Чавчанидзе [Чавчанидзе, 1970], задолго до появления самой концепции квантовых вычислений.

В последнее время, благодаря возросшим возможностям квантовой оптики, КНС на оптической основе получили новый импульс развития [Inagaki et al., 2016; McMahon et al., 2016].

Ключевым элементом функционирования нейронных сетей, как классических, так и квантовых, является *процесс обучения*. Различают обучение самоорганизацией, с учителем и с подкреплением. При обучении без учителя предъявление обучающей выборки приводит к кластеризации данных путём оптимизации некоторого функционала качества представления данных. Обычно это происходит путём конкуренции нейронов за получение большего сигнала.

В основе обучения с учителем лежит *намеренное* сопоставление каждому входному вектору из обучающей выборки желаемого вектора результата. При этом производится настройка весов сети — последовательность действий, направленных на то, чтобы при каждом предъявлении сети вектора из обучающей выборки выходной вектор был максимально близок к вектору желаемого результата. В известных моделях квантовых нейронных сетей процедура настройки сети состоит из последовательности классических операций измерения и приготовления состояний. Вопрос о том, может ли настройка сети производиться на чисто квантовом уровне связан с ролью сознания в квантовой механике и пока является открытым [Hagan et al., 2002].

Квантовые методы машинного обучения обобщают методы обучения классических нейронных сетей таким образом, чтобы можно было использовать квантовый параллелизм для ускорения работы сети [Ezhov, Ventura, 2000]. В соответствии с этим, существующие модели КНС можно разделить на несколько типов:

1. Сети типа Хопфилда, минимизирующие энергию системы изинговских спинов [Johnson et al., 2011].

2. Квантовая ассоциативная память [Ezhov et al., 2000; Ventura, Martinez, 2000; Vlasov, 1997].
3. Прямоточные нейронные сети, типа многослойного персептрона [Altaisky, 2001; Kouda et al., 2002; Wan et al., 2017; Zhou et al., 2006].
4. Линейные классификаторы, основанные на методе опорных элементов (SVM) [Rebentrost et al., 2014].
5. Безвесовые квантовые сети [da Silva et al., 2012].
6. Квантовые машины Больцмана [Amin et al., 2018].
7. Квантовые симуляторы [Feynman, 1982; Georgescu et al., 2014].
8. Квантовые блуждания на графах [Schuld et al., 2014].

Здесь мы не уточняем, на каких именно данных, классических или квантовых, учиться сеть — это привело бы к дальнейшему расширению классификации.

Практическое применение к настоящему времени получили КНС хопфилдовского типа, называемые также адиабатическими квантовыми компьютерами, или машинами квантового отжига.

6. АДИАБАТИЧЕСКИЕ КВАНТОВЫЕ КОМПЬЮТЕРЫ

Основной причиной, препятствующей созданию масштабируемой технологии производства квантовых компьютеров является *проблема декогерентности*: вследствие взаимодействия с окружающей средой нарушаются фазовые соотношения между квантовыми состояниями отдельных кубитов и эволюция многокубитовой системы теряет унитарность. Проблема декогерентности существует для всех известных реализаций кубитов: сквидов, ионов в ловушках, межуровневых переходов в молекулах, квантовых точек и других систем. Основной способ снижения декогерентности — понижение температуры системы — приводит к существенному удорожанию сетевых квантовых компьютеров. К настоящему времени квантовые компьютеры сетевого типа содержат не более сотни кубитов.

Наряду с сетевой обработкой квантовых регистров с помощью системы квантовых гейтов, существует, однако, и другой способ использования квантового параллелизма для ускорения вычислений, не связанный непосредственно с сохранением фазовых соотношений между отдельными кубитами. Этот способ, часто называемый *методом квантового отжига*, состоит в решении задачи оптимизации путём квантовой эволюции спинового стекла из специально подготовленного начального состояния в состояние, отвечающее минимуму квадратичного гамильтониана

$$\mathcal{H}_P = -\sum_i h_i \sigma_i^z - \sum_{i < j} J_{ij} \sigma_i^z \sigma_j^z. \quad (9)$$

Кубиты в данном случае реализованы в виде изинговских спинов, принимающих значения $\sigma_i^z = \{-1, +1\}$ после измерения. Спины расположены на d -мерной решётке \mathbb{Z}^d , а условия задачи оптимизации задаются матрицей связей J_{ij} , элементы которой в квантовом случае принимают только дискретные значения $J_{ij} \in \{0, 1\}$. Решение задачи оптимизации в данном случае сводится

к нахождению такой конфигурации спинов $\{\sigma_i^z\}_{i \in \mathbb{Z}^d}$, которая обеспечивает минимум гамильтониану \mathcal{H}_p при заданных значениях параметров $\{h_i, J_{ij}\}$,

Решение задачи поиска минимума энергии системы спинов известно и в случае классических нейронных сетей: гамильтониан (9), при дискретных значениях $\sigma_i^z = \{-1, +1\}$, описывает так называемую сеть Хопфилда, или спиновое стекло. Для достаточно сложного рельефа функции $\mathcal{H}_p(J, h)$ решение путём перебора всех возможных конфигураций спинов, или градиентными методами, является экспоненциально сложной задачей. По этой причине, для её решения на классическом компьютере часто используют эвристические методы, или используют классический метод отжига [De Vicente et al., 2003], время исполнения которого существенно возрастает при наличии у исследуемой функции большого числа *локальных* минимумов.

Преимущество системы квантовых спинов, принимающих наряду с двумя классическими значениями все возможные квантовые суперпозиции значений

$$|\psi\rangle = c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle, \quad (10)$$

состоит в существовании возможности квантового туннелирования системы спинов из локального минимума энергии в другой минимум, имеющий меньшую энергию. Это избавляет от больших затрат времени на выход из локального минимума, неизбежных для любого классического алгоритма. Адиабатические квантовые компьютеры, со «спинами» реализованными в виде сверхпроводящих квантовых интерференционных устройств (СКВИД, SQUID — Superconducting Quantum Interference Device), являются к настоящему времени единственным видом коммерчески доступных квантовых компьютеров. Число кубитов в них превосходит 1024. Эти адиабатические квантовые компьютеры выпускаются компанией D-Wave Systems Inc. Высокая цена изделия, $\sim 10^7$ дол., объясняется низкой рабочей температурой сквидов, менее 0,1 К, требующей больших затрат энергии в системе охлаждения.

Принцип работы адиабатического квантового компьютера состоит в следующем. Система спинов, описываемая гамильтонианом (9), реализуется в виде набора сверхпроводящих квантовых интерференционных устройств SQUID. Кольцевой ток в сверхпроводящей петле, разделённой джозефсоновским контактом, порождает существование квантованного магнитного потока, в направлении вверх или вниз, в зависимости от направления тока. Квантовые состояния магнитного потока описываются теми же правилами суперпозиции (10), что и состояния отдельных спинов, а время жизни когерентных состояний существенно выше. Константы связи J_{ij} реализуются с помощью подключения/отключения индуктивных связей между отдельными кубитами [Johnson et al., 2011].

Поскольку сверхпроводящий кубит на основе SQUID имеет классические размеры (от микрометров до миллиметров), а вероятность самопроизвольного изменения направления магнитного потока на противоположное исчезающе мала, система кубитов на основе SQUID с индуктивными связями является надёжной реализацией масштабируемого квантового регистра. Для приведения такого регистра в основное состояние используют метод квантового отжига, сочетающий в себе обычный метод отжига и квантовое туннелирование [Martoňák et al., 2002].

Для этого, в дополнение к обычному взаимодействию «спинов», описываемому гамильтонианом (9), перпендикулярно к магнитному полю h_z , в направлении

которого производится измерение проекций спинов в конечном состоянии, включается дополнительное магнитное поле Δ в направлении x , величина которого монотонно убывает со временем. Результирующий гамильтониан системы взаимодействующих спинов в магнитном поле приобретает вид

$$\mathcal{H}(t) = -\Gamma(t) \sum_{i=1}^n \Delta_i \sigma_i^x + \Lambda(t) H_P, \quad (11)$$

где $\Gamma(t)$ монотонно убывает со временем от 1 до 0, а $\Lambda(t)$ возрастает от 0 до 1. В начальный момент времени $t=0$, $\Lambda(0)=0$, основным состоянием системы является ориентация всех спинов вдоль оси x . При достаточно медленном уменьшении $\Gamma(t)$, с одновременным увеличением $\Lambda(t)$, система спинов адиабатически переходит в основное состояние гамильтониана H_P [Fahri et al., 2001].

Если система находится в области локального минимума энергии, отделённого потенциальным барьером, то вероятность перехода через потенциальный барьер высотой ΔE для классического метода отжига пропорциональна $\exp(-\Delta E/T)$, а для квантового отжига вероятность прохождения барьера высотой ΔE и шириной L пропорциональна $\exp(-L\sqrt{\Delta E}/\Gamma)$, где Γ — постоянная квантового отжига. Это означает, что алгоритм квантового отжига будет превосходить по эффективности алгоритм классического отжига для барьеров, обладающих большой высотой, но малой шириной [Das, Chakrabarti, 2008]. Доказательство фактической эквивалентности метода квантового отжига и обычного сетевого квантового компьютеринга было дано в работе [Mizel et al., 2007].

7. КВАНТОВАЯ АССОЦИАТИВНАЯ ПАМЯТЬ

В задаче минимизации энергии изинговского гамильтониана (9) при заданной матрице J_{ij} неизвестной является конфигурация спинов $\{\sigma_i^z\}_{i=1, \overline{M}}$, соответствующая минимуму энергии. Обратная задача — поиск матрицы J_{ij} по набору векторов $\{\sigma_i^z\}_{i=1, \overline{M}}$ — возникает при обработке и классификации изображений. Если необходимо запомнить M изображений размером $N = m \times n$ каждое, то для классической сети Хопфилда в качестве матрицы Хейбба J_{ij} , можно выбрать корреляционную матрицу

$$J_{ij} = \frac{1}{M} \sum_{l=1}^M \xi_i^{(l)} \xi_j^{(l)}, \quad 1 \leq i, j \leq N. \quad (12)$$

Переход от классических битов к квантовым превращает матрицу Хейбба (12) в проекционный оператор

$$\hat{J} = \sum_{l=1}^M |l\rangle \langle l|, \quad (13)$$

что фактически сводит модель сети Хопфилда к оптической нейронной сети, в которой исходное оптическое изображение разделяется на M каналов, на каждом из которых установлена маска, фильтр, отвечающий изображению $|l\rangle \langle l|$, так что задача классификации сводится к определению того канала, в котором больше амплитуда выходного сигнала (рис. 5, см. с. 43).

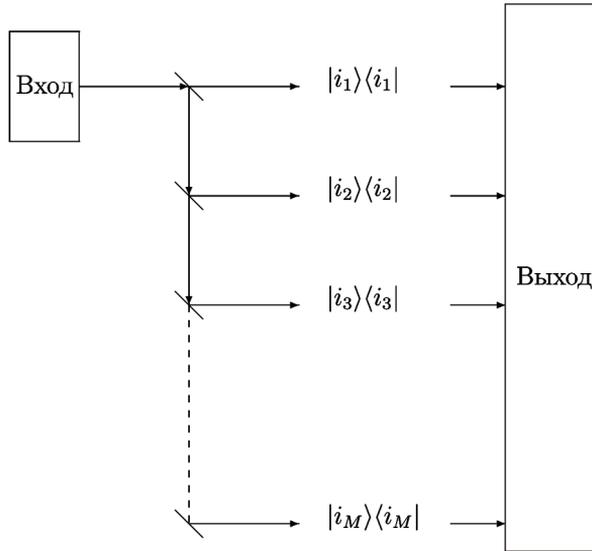


Рис. 5. Реализация квантовой ассоциативной памяти на основе оптической интерференции (схема из работы [Vlasov, 1997]). Входной сигнал — квантовое состояние $|\chi\rangle$, которое необходимо отнести к одному из M классов, — расщепляется с помощью системы полупрозрачных зеркал на M каналов, в которых установлены проекционные фильтры $\hat{J}_i = |i\rangle\langle i|$. Отнесение изображения к одному из классов производится по максимуму выходной интенсивности светового потока $I_i = (I_0/M) \cdot |\langle \chi | i \rangle|^2$

Все каналы в оптической нейронной сети, равно как и в квантовой сети, обрабатываются одновременно.

8. КЛАССИФИКАЦИЯ БОЛЬШИХ ДАННЫХ

Одним из методов машинного обучения с учителем, предназначенным для классификации данных на два класса $\{-1, +1\}$, является *метод опорных векторов* (SVM, Support Vector Machine). Задачей SVM является классификация предъявляемых произвольных N -мерных векторов на основании имеющейся обучающей выборки из M эталонных векторов:

$$\left\{ (\mathbf{x}_j, y_j) : \mathbf{x}_j \in \mathbb{R}^N, y_j = \pm 1 \right\}_{j=1, \overline{M}}.$$

Для выполнения такой классификации находят оптимальную гиперплоскость, определяемую вектором \mathbf{w} в пространстве данных, которая делит \mathbb{R}^N на два искомых класса:

$$\mathbf{w} \cdot \mathbf{x}_j + b \geq +1 \quad \text{для класса } +1,$$

$$\mathbf{w} \cdot \mathbf{x}_j + b \leq -1 \quad \text{для класса } -1,$$

где $b/|\mathbf{w}|$ — смещение гиперплоскости. Для нахождения параметров \mathbf{w} и b , определяющих гиперплоскость, классическим алгоритмам линейной алгебры

необходимо $O(\log(1/\epsilon) \text{poly}(N, M))$ шагов, при точности ϵ [Boyd, Vandenberghe, 2004]. При этом нахождение разделяющей гиперплоскости сводится к минимизации нормы $|\mathbf{w}|^2/2$ с ограничением $y_j(\mathbf{w} \cdot \mathbf{x}_j + b) \geq 1$.

Аналогично методу главных компонент, для построения разделяющей гиперплоскости удобно использовать корреляционную матрицу векторов обучающей выборки

$$J_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j. \quad (14)$$

Определив эту матрицу, решают задачу максимизации для функции Каруша – Куна – Такера

$$L(\vec{\alpha}) = \sum_{j=1}^M y_j \alpha_j - \frac{1}{2} \sum_{j,k=1}^M \alpha_j J_{jk} \alpha_k, \quad (15)$$

с ограничениями $\sum_{j=1}^M \alpha_j = 0$ и $y_j \alpha_j \geq 0$. После чего можно положить

$$\mathbf{w} = \sum_{j=1}^M \alpha_j \mathbf{x}_j$$

и $b = y_j - \mathbf{w} \cdot \mathbf{x}_j$, для тех j , для которых $\alpha_j \neq 0$. Векторы \mathbf{x}_j , для которых $\alpha_j \neq 0$, называют *опорными векторами*. Поскольку вычисление скалярного произведения требует $O(N)$ шагов, выполнение всего алгоритма с точностью ϵ требует $O(\log(1/\epsilon) M^2 \times (M + N))$ шагов.

В квантовом случае для построения корреляционных матриц можно ввести дополнительный регистр, отвечающей номеру вектора в обучающей выборке. В этом случае, вся обучающая выборка будет представлена вектором

$$|\chi\rangle = \frac{1}{\sqrt{\sum_{i=1}^M |\mathbf{x}_i|^2}} \sum_{i=1}^M |\mathbf{x}_i\rangle |\mathbf{x}_i\rangle, \quad (16)$$

где

$$|\mathbf{x}_i\rangle = \frac{1}{|\mathbf{x}_i|} \sum_{k=1}^N (\mathbf{x}_i)_k |k\rangle.$$

Корреляционная матрица (14) становится следом проекционного оператора $P_\chi \equiv |\chi\rangle\langle\chi|$ по второму аргументу:

$$\text{Tr}_2 P_\chi = N_\chi^{-1} \sum_{i,j=1}^M \langle \mathbf{x}_j | \mathbf{x}_i \rangle |\mathbf{x}_i\rangle\langle j| = \frac{J}{\text{Tr} J},$$

где $N_\chi = \sum_{i=1}^M |\mathbf{x}_i|^2$, и, с учётом квантового параллелизма, может быть вычислена на сетевом квантовом компьютере за $O(\log NM)$ шагов. Таким образом, процесс обучения бинарного классификатора может быть экспоненциально ускорен, если описанный алгоритм классификации выполнять на сетевом квантовом компьютере [Rebentrost et al., 2014].

С другой стороны, нахождение корреляционной матрицы векторов обучающей выборки есть не что иное, как решение задачи об ассоциативной квантовой памяти.

9. ЗАДАЧИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АЭРОКОСМИЧЕСКОЙ ОТРАСЛИ

Будущее исследования космоса неразрывно связано с развитием систем искусственного интеллекта. Автономные планетоходы, предназначенные для исследования планет, и автономные космические аппараты должны принимать решения по осуществлению тех или иных действий самостоятельно, или с минимальным участием человека: большие расстояния, и соответствующие времена задержки сигнала, не позволяют проводить обработку поступающей от датчиков информации на Земле. Искусственные нейронные сети, как и биологические, осуществляют управление автономными аппаратами по схеме, изображённой на рис. 6.

В значительной степени это же относится и к пилотируемым миссиям, где задачи корректировки орбиты, стыковки и посадки, требующие значительного объёма расчётов, до сих пор решаются эвристическими методами.

Широкое использование методов искусственного интеллекта в аэрокосмической отрасли достигнуто при обработке изображений, где методы применения искусственных нейронных сетей, в том числе методы глубокого обучения, используются напрямую.

Большинство упомянутых выше задач может быть сведено к решению задачи оптимизации для функции большого числа переменных. Так, задача корректировки орбиты космического аппарата есть задача минимизации расхода топлива при известных параметрах начальной и конечной орбит. В отсутствие специальных ограничений, объём ресурсов, необходимых для решения задачи оптимизации, экспоненциально зависит от размера данных, т.е., фактически, от числа параметров, определяющих состояние космического аппарата и его окружения. Поскольку число этих параметров весьма велико, решение задачи оптимизации на классическом компьютере часто оказывается невозможным. В силу возможности ускорения вычислений за счёт квантового параллелизма большие надежды здесь возлагаются на адиабатические квантовые компьютеры, которые позволяют решать задачи оптимизации путём квантового туннелирования системы кубитов из начального состояния в оптимальное.

Кроме того, если для какой-то задачи размерности входных и выходных данных не велики, а время счёта алгоритмическим путём велико, то аппроксимация решения задачи отображением «вход → выход» на основе обучения бывает эффективнее, чем проведение вычислений. Именно так поступает человек при управлении летательным аппаратом, где параметрами движения являются координаты и скорости, и при решении задачи оптимизации задействовано большое число нейронов мозга.

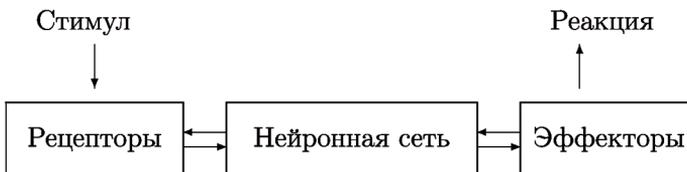


Рис. 6. Включение нейронной сети в систему искусственного интеллекта (из работы [Altaisky et al., 2014])

Решение задачи оптимизации с таким же числом нейронов искусственной нейронной сети на классическом компьютере было до настоящего времени не реальным (до распространения методов глубокого обучения и появления высокопроизводительных микропроцессоров). В квантовом же случае минимизация энергии в системе из 50 кубитов, в каком-то смысле эквивалентно работе с матрицами размером $2^{50} \times 2^{50}$, что нереально для классических компьютеров [Castelvecchi, 2017; Dumitrescu et al., 2018].

Одним из наиболее перспективных приложений высокопроизводительных систем искусственного интеллекта являются задачи зондирования Земли и других планет с помощью космических аппаратов. Здесь главной проблемой является уже не время обработки отдельного изображения, а большой объём данных, получаемых космическим аппаратом в видимом или ином спектральном диапазоне. Просмотреть такое количество изображений человеку просто невозможно. По этой причине, необходимо симулировать классификационные процессы, производимые нейронной сетью мозга, с помощью искусственных нейронных сетей. Отчасти это уже делается с использованием методов глубокого обучения. Использование квантовых нейронных сетей, обеспечивающих квантовое ускорение процесса обучения, должно существенно повысить эффективность классификации в больших потоках данных.

10. ЭЛЕМЕНТНАЯ БАЗА КВАНТОВЫХ НЕЙРОННЫХ СЕТЕЙ

Используемая в квантовых нейронных сетях, работающих по принципу машины квантового отжига, модель спинового стекла, описываемая квадратичным гамильтонианом изинговского типа (9), является достаточно общей математической моделью, используемой в различных областях физики неупорядоченных систем, от металлургии до биологической физики и космологии.

Первой реализацией квантовой нейронной сети хопфилдовского типа явилась система кубитов на основе квантованного магнитного потока SQUID [Johnson et al., 2011]. Матрица J_{ij} в такой системе реализована в виде индуктивных связей между отдельными сквидами. Достоинствами такой системы являются простота управления индуктивными связями J_{ij} , и наличие планарной технологии, позволяющей размещать на плате нужное число сквидов, и тем самым обеспечивая масштабируемость. Фотография процессора компании D-wave на 1000 кубитов приведена на рис. 7.

Существенным недостатком производимых на этой основе процессоров D-Wave является низкая рабочая температура сквидов, лежащая в милликельвиновом диапазоне, что требует громоздкой системы охлаждения.

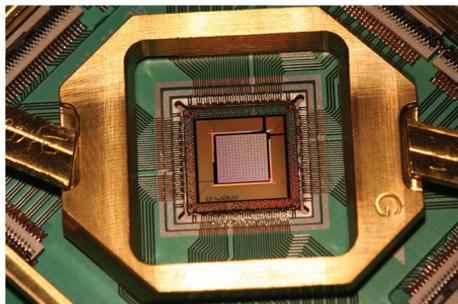


Рис. 7. Процессор компании D-Wave Systems на 1000 кубитов на основе SQUID (с сайта www.dwavesys.com)

Большие затраты на охлаждение, $\sim 14,2$ кВт для выпускаемых сейчас процессоров D-Wave 2000Q, приводят к их удорожанию и препятствуют миниатюризации.

В качестве альтернативной модели квантовой нейронной сети Хопфилда японскими учёными была построена оптическая реализация гамильтониана (9) на основе оптических параметрических осцилляторов (ОРО, optical parametric oscillator), где каждый «спин» представлен возможными значениями фазы $\{0, \pi\}$ по отношению к фазе накачивающего лазера [Inagaki et al., 2016]. Набор осцилляторов циркулирует в замкнутом оптоволоконном волноводе, причём расстояния между осцилляторами (пакетами) поддерживаются одинаковыми. Матрица связей J_{ij} реализуется с помощью интерферометра с временной задержкой [Inagaki et al., 2016]. Полная длина оптического волновода может достигать нескольких километров, а число эмулируемых спинов — десятков тысяч. Вычисление завершается когда система «спинов», после прохождения определённого числа замкнутых циклов в волноводе ($\sim 10^2$), приходит в равновесное состояние. Типичное время вычислений при решении комбинаторных задач в системе из 100 изинговских спинов составляет менее 1 мс [McMahon et al., 2016]. Очевидным достоинством данной схемы является использование стандартной оптоволоконной технологии, обеспечивающей работоспособность изделия при комнатных температурах и низкую себестоимость. К потенциальным недостаткам следует отнести использование в качестве вычислительного пространства квазиодномерной системы — оптического волновода. В сочетании с использованием классического электронного контроллера, использование одномерного массива делает не вполне понятными свойства масштабируемости данной схемы и их влияние на скорость вычислений.

Для решения задач комбинаторной оптимизации методов квантового отжига, в существующих магнитных [Johnson et al., 2011] и оптических [McMahon et al., 2016] моделях, использовался модельный гамильтониан изинговского типа (9). Нечто аналогичное можно построить и на других квантовых системах с двумя ортогональными квантовыми состояниями. Перспективным здесь представляется использование основного и возбуждённого состояния квантовых точек в качестве кубитов. Идея использования квантовых точек в качестве кубитов в сетевом квантовом компьютере была предложена Лоссом и Ди-Винченцо [Loss, DiVincenzo, 1998], а квантовая нейронная сеть на основе квантовых точек в работе [Behrman et al., 2000].

Квантовые точки («искусственные атомы») (КТ) представляют собой малые, от нескольких нанометров до микрометра, участки полупроводниковой гетероструктуры с точно контролируемым числом электронов [Kastner, 2005; Reimann, Manninen, 2002]. Потенциальными достоинствами квантовой нейронной сети на основе квантовых точек являются миниатюрность отдельного кубита, менее 100 нм, масштабируемость технологии производства массивов квантовых точек, и возможность управления параметрами квантовых точек с помощью внешних электрических и магнитных полей. Управлять корреляциями между состояниями отдельных квантовых точек можно устанавливая между ними электрические или оптические связи.

Взаимодействие между отдельными точками в массиве, выполненном на основе полупроводниковой гетероструктуры, обычно из арсенида галлия, осуществляется путём формирования экситонов и их взаимодействия с фононами гетероструктуры. Модельный гамильтониан такого массива, в приближении вращающейся волны, может быть записан в виде

$$H = \sum_i \frac{h_i}{2} \sigma_i^z + \sum_i \frac{K_i}{2} \sigma_i^x + \sum_{i \neq j} J_{ij} \sigma_i^+ \sigma_j^- + \sum_{\alpha, i} g_{\alpha} x_i |X_i\rangle \langle X_i| + H_{ph}, \quad (17)$$

где все матрицы Паули действуют в базисе основного (0) и возбуждённого (X) состояний квантовой точки ($\sigma^{\pm} \sigma^x \pm i \sigma^y$); g_{α} — общая константа взаимодействия фононной моды x_{α} с возбуждённым состоянием квантовой точки (экситоном).

Численные исследования поведения систем КТ в присутствии общего термостата акустических фононов, проведённые интегрированием уравнения фон Неймана

$$\dot{\rho} = \text{Tr}_{ph} \left[-\frac{i}{\hbar} [H, \rho_{tot}] \right]$$

с последующим усреднением по состояниям фононов, показывают, что квантовые осцилляции в массиве квантовых точек выживают до температуры порядка сотен градусов Кельвина [Altaisky et al., 2016a, b, 2017] (рис. 8). Аналогия с уже созданными оптическими системами изинговского типа [Inagaki et al., 2016; McMahon et al., 2016] позволяет надеяться, что диполь-дипольный аналог (17) спинового стекла (9), построенный на основе массива квантовых точек, может быть использован для построения квантовых нейросетевых процессоров на основе массивов квантовых точек работающих при обычной температуре.

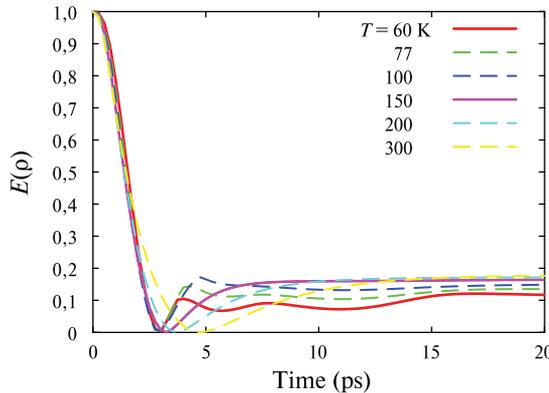


Рис. 8. Запутанность формирования пары квантовых точек в модели нейронной сети (17) при различной температуре. Наличие квантовых осцилляций прослеживается до температуры порядка 100 К (из работы [Altaisky et al., 2017])

Возможным недостатком здесь является проблема соединения всех со всеми и чувствительность системы к внешним электронным шумам.

ЗАКЛЮЧЕНИЕ

Идея квантового параллелизма, позволившая успешно решать модельные комбинаторные задачи экспоненциальной сложности за полиномиальное время с использованием квантовой суперпозиции состояний вычислительного регистра, в последние 10 лет получила новый импульс развития. Этим импульсом стало появление квантовых нейронных сетей.

Если в традиционном квантовом компьютеринге основной проблемой была защита от декогеренции, что сильно ограничивало число кубитов, употребляемых для вычислений, то объединяя идею квантовых вычислений [Deutsch, 1985] с парадигмой искусственных нейронных сетей, можно отчасти снять эту проблему, применяя оптические технологии, работающие при комнатной температуре [Inagaki et al., 2016], а также используя свойства симметрии открытых квантовых систем, составляющих нейронную сеть. В результате мы ожидаем в ближайшем будущем получить нейроморфные квантовые компьютеры для решения широкого класса задач, входящих в функции искусственного интеллекта: комбинаторной оптимизации, классификации больших данных, распознавания образов. В значительной степени такие нейроморфные системы смогут заменить интеллектуальную деятельность человека при решении задач оптимизации и классификации.

В данной работе был рассмотрен лишь достаточно узкий класс задач, наиболее близкий к практической реализации с помощью квантовых нейронных сетей. За пределами работы остались методы обучения с учителем, применяемые для квантовых нейронных сетей, и так называемые квантово-инспирированные нейронные сети [Menneer, Narayanan, 1995], в которых квантовая суперпозиция состояний или весов сети отдельным образом моделируется на обычном компьютере с помощью классических алгоритмов. Некоторые направления развития таких методов, которые могут оказаться весьма полезными при разработке различных тренажеров и симуляторов, можно найти, например, в работах [Алтайский и др., 2014; Schuld et al., 2014; Wan et al., 2017].

ЛИТЕРАТУРА

- [Алтайский и др., 2014] *Алтайский М. В., Капуткина Н. Е., Крылов В. А.* Квантовые нейронные сети: Современное состояние и перспективы развития // *Физика элементарных частиц и атомного ядра.* 2014. Т. 45. С. 1824–1864.
- [Осовский, 2004] *Осовский С.* Нейронные сети для обработки информации: пер. с польского. М.: Финансы и статистика, 2004.
- [Чавчанидзе, 1970] *Чавчанидзе В. В.* К вопросу о пространственно-временных волновых процессах в нервных сетях // *Сообщения АН Грузинской ССР.* 1970. Т. 59. № 1. С. 37–40.
- [Altaisky, 2001] *Altaisky M. V.* Quantum neural network. Technical report, arXiv.org:quant-ph/0107012. 2001.
- [Altaisky, Rao, 2009] *Altaisky M. V., Rao V. S. H.* Inverted mexican hat potential in activation of receptor cells // *Nonlinear Analysis B.* 2009. V. 10. Iss. 5. P. 2961–2970.
- [Altaisky et al., 2016a] *Altaisky M. V., Zolnikova N. N., Kaputkina N. E., Krylov V. A., Lozovik Y. E., Dattani N. S.* Towards a feasible implementation of quantum neural networks using quantum dots // *Applied Physics Letters.* 2016. V. 108. Iss. 10. P. 103–108.
- [Altaisky et al., 2016b] *Altaisky M. V., Zolnikova N. N., Kaputkina N. E., Krylov V. A., Lozovik Yu. E., Dattani N. S.* Decoherence and entanglement simulation in a model of quantum neural network based on quantum dots // *EPJ Web of Conferences.* 2016. V. 108. P. 02006.
- [Altaisky et al., 2017] *Altaisky M. V., Zolnikova N. N., Kaputkina N. E., Krylov V. A., Lozovik Yu. E., Dattani N. S.* Entanglement in a quantum neural network based on quantum dots // *Photonics and Nanostructures – Fundamentals and Applications.* 2017. V. 24. P. 24–28.

- [Amin et al., 2018] *Amin M. H., Andriyash E., Rolfe J., Kulchytsky B., Melko R.* Quantum Boltzmann Machine // *Phys. Rev. X*. 2018. V. 8. Iss. 2. P. 021050. doi: 10.1103/PhysRevX.8.021050.
- [Behrman et al., 2000] *Behrman E. C., Nash L. R., Steck J. E., Chandrashekar V. G., Skinner S. R.* Quantum dot neural networks // *Inf. Sci.* 2000. V. 128. P. 257.
- [Boyd, Vandenberghe, 2004] *Boyd S., Vandenberghe L.* Convex Optimization. Cambridge University Press, 2004.
- [Breuer, Petruccione, 2002] *Breuer H. P., Petruccione F.* The theory of open quantum systems. Oxford University Press, 2002.
- [Castelvecchi, 2017] *Castelvecchi D.* Quantum cloud goes commercial // *Nature*. 2017. V. 543. P. 159.
- [da Silva et al., 2012] *da Silva A. J., de Oliveira W. R., Ludermir T. B.* Classical and superposed learning for quantum weightless neural networks // *Neurocomputing*. 2012. V. 75. Iss. 52. P. 52–60.
- [Das, Chakrabarti, 2008] *Das A., Chakrabarti B. K.* Colloquium: Quantum annealing and analog quantum computation // *Rev. Mod. Phys.* 2008. V. 80. P. 1061–1081.
- [De Vicente et al., 2003] *De Vicente J., Lanchares J., Hermida R.* Placement by thermodynamic simulated annealing // *Physics Letters A*. 2003. V. 317. P. 415–423.
- [Deutch, 1985] *Deutch D.* Quantum theory, the Church-Turing principle and the universal quantum computer // *Proc. Roy. Soc. Lond. A*. 1985. V. 400. P. 97–117.
- [Deutsch, Jozsa, 1992] *Deutsch D., Jozsa R.* Rapid solution of problems by quantum computation // *Proc. Roy. Soc. Lond. A*. 1992. V. 439. P. 553–558.
- [Dumitrescu et al., 2018] *Dumitrescu E. F., McCaskey A. J., Hagen G., Jansen G. R., Morris T. D., Papenbrock T., Pooser R. C., Dean D. J., Lougovski P.* Cloud quantum computing of an atomic nucleus // *Phys. Rev. Lett.* 2018. V. 120. P. 210501.
- [Ezhov et al., 2000] *Ezhov A. A., Nifanova A. V., Ventura D.* Quantum associative memory with distributed queries // *Inf. Sci.* 2000. V. 128. Iss. 3–4. P. 271–293.
- [Ezhov, Ventura, 2000] *Ezhov A., Ventura D.* Quantum neural networks // *Future Directions for Intelligent Systems and Information Science* / ed. N. Kasabov. Physica-Verlag, 2000.
- [Fahri et al., 2001] *Fahri E., Goldstone J., Gutmann S., Lapan J., Lundgren A., Preda D.* A quantum adiabatic evolution algorithm applied to random instances of NP-complete problem // *Science*. 2001. V. 292. P. 472–475.
- [Feynman, 1982] *Feynman R.* Simulating physics with computers // *Intern. J. Theor. Phys.* 1982. V. 21. Iss. 467–488.
- [Feynman, Vernon, 1963] *Feynman R. P., Vernon Jr. F. L.* The theory of a general quantum system interacting with a linear dissipative system // *Ann. Phys.* 1963. V. 24. P. 118–173.
- [Georgescu et al., 2014] *Georgescu I. M., Ashhab S., Nori F.* Quantum simulation // *Rev. Mod. Phys.* 2014. V. 86. P. 153–185.
- [Grover, 1997] *Grover L. K.* Quantum mechanics helps in searching for a needle in a haystack // *Phys. Rev. Lett.* 1997. V. 79. P. 325–328.
- [Hagan et al., 2002] *Hagan S., Hameroff S. R., Tuszynski J. A.* Quantum computation in brain microtubules: Decoherence and biological feasibility // *Phys. Rev. E*. 2002. V. 65. P. 061901.
- [Haykin, 1999] *Haykin S.* Neural Networks. Pearson Education, 1999.
- [Hebb, 1949] *Hebb D. C.* Organization of behavior. N. Y.: Wiley, 1949.
- [Inagaki et al., 2016] *Inagaki T., Inaba K., Hamerly R., Inoue K., Yamamoto Y., Takesue H.* Large-scale Ising spin network based on degenerate optical parametric oscillators // *Nat. Photonics*. 2016. V. 10. P. 415–419.
- [Johnson et al., 2011] *Johnson M. W. et al.* Quantum annealing with manufactured spins // *Nature*. 2011. V. 473. P. 194–198.
- [Johnson et al., 2011] *Johnson M. W., Amin M. H. S., Gildert S., Lanting T., Hamze F., Dickson N., Harris R., Berkley A. J., Johansson J., Bunyk P., Chapple E. M., Enderud C., Hilton J. P., Karimi K., Ladizinsky E., Ladizinsky N., Oh T., Perminov I., Rich C., Thom M. C.,*

- Tolkacheva E., Truncik C.J.S., Uchaikin S., Wang J., Wilson B., Rose G.* Quantum annealing with manufactured spins // *Nature*. 2011. V. 473. Iss. 7346. P. 194–198.
- [Kak, 1995] *Kak S.* On quantum neural computing // *Inf. Sci.* 1995. V. 83. P. 143–160.
- [Kastner, 2005] *Kastner M.A.* Prospects for quantum dot implementation of adiabatic quantum computers for intractable problems // *IEEE Proc.* 2005 V. 93. Iss. 10. P. 1765–1771.
- [Kouda et al., 2002] *Kouda N., Matsui N., Nishimura H.* Image compression by layered quantum neural networks // *Neural Processing Letters*. 2002. V. 16. Iss. 1. P. 67–80.
- [LeCun et al., 2015] *LeCun Y., Bengio Y., Hinton G.* Deep learning // *Nature*. 2015. V. 521. P. 436–444.
- [Lee et al., 2006] *Lee W.-C.A., Huang H., Feng G., Sanes J.R., Brown E.N., So P.T., Nedivi E.* Dynamic remodeling of dendritic arbors in GABAergic interneurons of adult visual cortex // *PLoS Biology*. 2006. V. 4. Iss. 2. P. e29.
- [Loss, DiVincenzo, 1998] *Loss D., DiVincenzo D.P.* Quantum computation with quantum dots // *Phys. Rev. A*. 1998. V. 57. Iss. 1. P. 120–126.
- [Martoňák et al., 2002] *Martoňák R., Santoro G.E., Tosatti E.* Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model // *Phys. Rev. B*. 2002. V. 66. P. 094203.
- [McCulloch, Pitts, 1943] *McCulloch W.S., Pitts W.* A logical calculus of the ideas immanent in nervous activity // *Bulletin of Mathematical Biophysics*. 1943. V. 5. P. 115–133.
- [McMahon et al., 2016] *McMahon P.L., Marandi A., Haribara I., Hamerly R., Langrock C., Tamate S., Inagaki T., Takesue H., Utsunomiya S., Aihara K., Byer R.L., Fejer M.M., Mabuchi H., Yamamoto Y.* A fully programmable 100-spin coherent Ising machine with all to all connections // *Science*. 2016. V. 354. Iss. 6312. P. 614–617.
- [Menner, Narayanan, 1995] *Menner T., Narayanan A.* Quantum-inspired neural networks // *Neural Information Processing Systems 1995*. Denver, Colorado, 1995. 8 p.
- [Mitchell, 1997] *Mitchell T.M.* *Machine Learning*. McGraw-Hill, 1997.
- [Mizel et al., 2007] *Mizel A., Lidar D.A., Mitchell M.* Simple proof of equivalence between adiabatic quantum computation and the circuit model // *Phys. Rev. Lett.* 2007. V. 99. P. 070502.
- [Nielsen, Chuang, 2000] *Nielsen M.A., Chuang I.L.* *Quantum computation and quantum information*. N.Y.: Cambridge University Press, 2000.
- [Nitta, 1997] *Nitta T.* An extension of the back-propagation algorithm to complex numbers // *Neural Networks*. 1997. V. 10. Iss. 8. P. 1397–1415.
- [Rebentrost et al., 2014] *Rebentrost P., Mohseni M., Lloyd S.* Quantum support vector machine for big data classification // *Phys. Rev. Lett.* 2014. V. 113. Iss. 13. P. 130503.
- [Reimann, Manninen, 2002] *Reimann S.M., Manninen M.* Electronic structure of quantum dots // *Rev. Mod. Phys.* 2002. V. 74. P. 1283–1342.
- [Rosenblatt, 1958] *Rosenblatt F.* The perceptron: a probabilistic model for information storage and organization of brain // *Psychological Review*. 1958. V. 65. No. 6. P. 386–408.
- [Schuld et al., 2014] *Schuld M., Sinayskiy I., Petruccione F.* Quantum walks on graphs representing the firing patterns of a quantum neural network // *Phys. Rev. A*. 2014. V. 89. P. 032333.
- [Schuld et al., 2014] *Schuld M., Sinayskiy I., Petruccione F.* The quest for a Quantum Neural Network // *Quantum Information Processing*. 2014. V. 13. Iss. 11. P. 2567–2586.
- [Shariv, Friesem, 1989] *Shariv I., Friesem A.A.* All-optical neural network with inhibitory neurons // *Opt. Lett.* 1989. V. 14. Iss. 10. P. 485–487.
- [Shor, 1994] *Shor P.* Polynomial-time algorithm for prime factorization and discrete logarithms on a quantum computer // *Proc. 35th Annual Symposium on Foundations of Computer Science*. Piscataway, NJ, 1994.
- [Stolze, Suter, 2008] *Stolze J., Suter D.* *Quantum computing: A short course from theory to experiment*. Wiley-VCH, 2008.
- [Ventura, Martinez, 2000] *Ventura D., Martinez T.* Quantum associative memory // *Inf. Sci.* 2000. V. 124. Iss. 1–4. P. 273–296.

- [Vlasov, 1997] *Vlasov A. Y.* Quantum computations and images recognition. arXiv.org:quant-ph/9703010. 1997.
- [Wan et al., 2017] *Wan K. H., Dahlsten O., Hlér K., Gardner R., Kim M. S.* Quantum generalisation of feedforward neural networks // npj Quantum information. 2017. V. 3. Iss. 36.
- [Zhou et al., 2006] *Zhou R., Qin L., Jiang N.* Quantum perceptron network. In Artificial Neural Networks Proc // Int. Conf. ICANN 2006. 2006. V. 4131. P. 651–657.

QUANTUM ARTIFICIAL INTELLIGENCE SYSTEMS

M. V. Altaisky, N. N. Zolnikova

Space Research Institute of Russian Academy of Sciences (IKI RAN)

The modern trends in the development of quantum artificial intelligence systems, which combine machine learning methods and the ideas of quantum parallelism, enabling the exponential quantum speedup of the solution of optimisation problems are considered in this paper. The paper presents the review of both the existing adiabatic quantum computers (quantum neural networks) based on SQUIDs and optical networks, and the original results of the authors in the field of the perspective element base for future compact quantum computers. We concentrate on the perspective of the development of the on-board quantum artificial intelligence systems aimed for the control of autonomus space missions as well as for the high-performance image processing systems for space research and the Earth monitoring.

Keywords: quantum learning, neural networks, artificial intelligence

Mikhail V. Altaisky — leading scientist, Dr. Sc. (Math. Phys), altaisky@iki.rssi.ru
Nadezhda N. Zolnikova — scientist, PhD, nzolnik@iki.rssi.ru

ИНДИЙСКИЙ МУССОН: КОНЦЕПЦИЯ КРИТИЧЕСКИХ ЭЛЕМЕНТОВ, ПРОГНОЗ НАЧАЛА И ЗАВЕРШЕНИЯ

Е. Д. Суровяткина

Институт космических исследований Российской академии наук (ИКИ РАН), Москва

Индийский муссон — важный элемент климатической системы Земли. Причины возникновения муссона связывают с двумя основными факторами: локальные сезонные изменения атмосферной циркуляции, вызванные асимметричным нагревом суши и моря, и глобальное сезонное перемещение внутритропической зоны конвергенции. Характерной чертой муссона является внезапность начала и завершения, даты которых могут варьироваться от года к году в течение месяца. Поэтому прогноз муссона представляет собой весьма сложную научную проблему. Долгосрочный прогноз муссона крайне важен для Индии, 70 % населения которой заняты в сельскохозяйственном производстве, которое чрезвычайно зависит от сезона дождей.

Представлен новый подход к исследованию явления муссона, основанный на теории критических переходов. Показано, что в пространственной организации муссона существуют такие регионы, в которых критические условия для его возникновения зарождаются раньше, чем в других. Эти регионы являются системообразующими элементами системы муссона. Локализация критических элементов обусловлена формой внутритропической зоны конвергенции и орографией Индостана. Таким образом, концепция критических элементов объединяет два различных описания явления муссона, основанных на глобальных и локальных причинах.

Полученный результат положен в основу нового метода прогноза муссона, в котором сравнительный анализ данных внутри областей критического поведения позволяет определять даты наступления и завершения муссона в центральной части Индии. В настоящее время официальный прогноз муссона осуществляется только для южной оконечности Индостана за 15 дней до средней даты начала муссона, а завершение муссона вообще не прогнозируется. Предложенный метод имеет ряд преимуществ по сравнению с существующими подходами: позволяет осуществлять долгосрочный прогноз начала муссона (за 40 дней) и завершение муссона (за 70 дней), впервые прогноз осуществляется для региона, расположенного в центральной части Индии.

Предложенный подход прогноза основан на анализе данных, когда модель процесса недостаточно развита для прогноза или ещё не существует. Подход не ограничивается приложением к Индийскому муссону, а может быть применён для прогнозов других физических систем, в которых происходит внезапный критический переход.

Ключевые слова: индийский муссон, критические явления, долгосрочный прогноз

ВВЕДЕНИЕ

Слово «муссон» происходит от арабского “mawsim”, которое означает время года или сезон дождей. Муссонами также называют ветры, вызывающие сезоны дождей или экстремальные осадки на территории большей части тропиков в Азии, Северной Америке, Южной Америке, Австралии и Африке.

Суровяткина Елена Дмитриевна — ведущий научный сотрудник, доктор физико-математических наук, selena@iki.rssi.ru

Основной механизм, ответственный за наступление муссонов или сезонов дождей, — сезонные изменения атмосферной циркуляции и осадков, связанные с асимметричным нагревом суши и моря (суша нагревается сильнее и быстрее, чем океан). Важнейшим управляющим фактором тропических муссонов является глобальное сезонное перемещение так называемой внутритропической зоны конвергенции (ВЗК) [Gadgil, 1998; Saha, Saha, 1980] — зоны низкого давления вдоль экватора, где встречаются тропические ветры — пассаты Северного и Южного полушарий. В этой зоне выпадает максимальное количество осадков. Поскольку ВЗК следует за максимумом солнечной радиации, она перемещается вместе с сезонным изменением склонения солнца (в пределах $\pm 23^{\circ}27'$). В большинстве случаев она смещается в то полушарие, в котором лето. Так, с наступлением лета в северном полушарии она смещается к северу, а зимой — к югу. Летнее перемещение ВЗК приносит юго-западный или летний муссон в Индию, в течение которого выпадает вплоть до 80 % годовой нормы осадков. Зимний или северо-восточный муссон затрагивает только несколько южных штатов Индии расположенных на побережье Бенгальского залива. В этой работе основное внимание уделено летнему Индийскому муссону.

Юго-западный муссон приходит на территорию Индии в виде двух основных ветвей — ветви Бенгальского залива и ветви Аравийского моря. Обычно около 25 мая он впервые достигает территории Индии — Андаманских и Никобарских островов в Бенгальском заливе. Около 1 июня он обрушивается на территорию Малабарского берега в штате Керала, расположенном на южной оконечности Индостана. А затем муссон на время «уходит» в Бенгальский залив, чтобы вернуться уже с востока. Этот момент чрезвычайно важен, потому что когда именно придёт муссон в центральную часть Индии через неделю, две или три — неизвестно. Принимая во внимание, что началу муссона предшествует три месяца засухи при температуре достигающей 44°C , а в 2016 г. до 51°C , становится понятно, почему вся страна замирает в ожидании этого события. Более того, доходы бедных (это около 770 миллионов человек или 70 % населения), продовольственная безопасность и запасы питьевой воды зависят от сезона дождей. В период сезона муссона страна запасает воду на два будущих года. Поэтому прогноз муссона для центральной Индии представляет собой насущную необходимость.

В настоящее время Департамент метеорологии Индии (Indian Meteorological Department — IMD) публикует прогноз начала сезона дождей 15 мая только для штата Керала за две недели до усреднённой даты его начала. О приходе муссона в других регионах страна узнаёт только из прогноза погоды от 2–5 дней до начала, что, конечно, недостаточно для планирования сельскохозяйственных работ, энергоресурсов, запаса воды и т.д. Таким образом, начало муссона для всей страны, исключая штат Керала, официально не прогнозируется, а лишь констатируется после того, как 60 % метеостанций региона сообщат о 48-часовом ливне.

Основная проблема прогноза муссона — внезапность его начала и завершения, причём даты могут варьироваться от года к году в течение месяца (рис. 1). Эта особенность характерна для всех регионов в Индии. Поэтому при решении задачи прогноза муссона приходится иметь дело с пространственной и временной вариабельностью. По этой причине прогноз распространения муссона по всей территории п-ва Индостан представляет сложнейшую научную проблему [Ananthakrishnan, Soman, 1990].

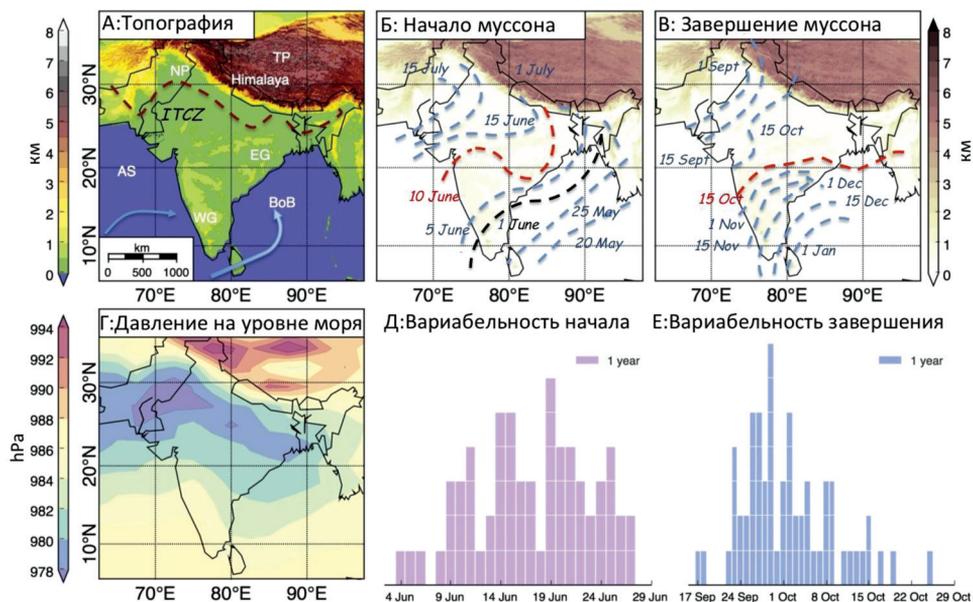


Рис. 1. Топография полуострова Индостан с ключевыми особенностями индийского летнего муссона (А): Гималаи, Тибетское плато (ТР), Северный Пакистан (NP), Восточные Гаты (EG), Западные Гаты (WG), Аравийское море (AS), Бенгальский залив (BoB) — внутритропической зоны конвергенции (ITCZ) (данные топографии — etopo1); синие стрелки указывают направление ветра на уровне моря. Схематическое изображение долгосрочного среднего распространения (с 1951 г., основано на IMD, доступно в интернете, 2015 г.) о начале и завершении муссона над индийским субконтинентом (северная граница муссона) (Б, В); пунктирная чёрная линия показывает усреднённое наступление муссонов для региона Кералы, прогнозируемое IMD, пунктирная красная линия — для Восточных Гат (область, представляющая большой интерес в этом исследовании). Распределение среднего давления на уровне моря (июнь — сентябрь 1951–2014 гг. по данным NCEP/NCAR) (Г); гистограммы дат начала и завершения для региона Восточных Гат (1951–2014) (Д, Е)

В отличие от метеорологических методик, основанных на численном моделировании прогноза погоды, в данной работе описан новый подход к прогнозированию Индийского муссона, в котором используются принципы анализа критических переходов. Впервые этот подход был описан в нашей статье в *Geophysical Research Letters* в 2016 году [Stolbova et al., 2016]. Основная идея заключается в том, чтобы рассматривать начало муссона, как время наступления срыва или внезапного перехода от состояния пре-муссонной засухи к состоянию муссона. Применение нового подхода позволило выявить в пространственной организации муссона такие регионы, в которых зарождаются критические условия, необходимые для его возникновения. Эти регионы и являются системообразующими компонентами системы муссона.

Полученный результат положен в основу нового метода предсказания муссона, в котором сравнительный анализ данных внутри областей критического поведения позволяет определять даты наступления и завершения Индийского муссона. Предложенный метод позволяет осуществлять долгосрочный (за 40 дней) прогноз начала муссона (вместо 15 дней существующими методами),

и завершение муссона за 70 дней (завершение муссона ранее никогда не прогнозировалось, а только декларировалось постфактум). Впервые прогноз осуществляется для региона, расположенного в центральной части Индостана.

В разделе 1 приведено краткое описание данных используемых для анализа. Далее, в разделе 2, изложен новый метод определения критических элементов и в разделе 3 — результаты его применения. В разделе 4 описан метод прогноза начала и завершения муссона. В разделе 5 представлены результаты тестирования метода для прогнозов будущих муссонов в 2016–2018 гг. В заключении подведены итоги работы.

1. ДАННЫЕ

Исследования муссона проводились на основе данных ре-анализа (ассимиляции наземных и спутниковых данных) NCEP/NCAR (Национальный центр океанических и атмосферных исследований, NOAA — США) [Kalnay et al., 1996] и ERA-40 (Европейский центр среднесрочных прогнозов, ECMWF) [Urrala et al., 2005] о среднесуточной приповерхностной температуре, давлении и относительной влажности с пространственным разрешением $2,5 \times 2,5^\circ$ с 1951 г. по настоящее время.

2. МЕТОД ОПРЕДЕЛЕНИЯ КРИТИЧЕСКИХ ЭЛЕМЕНТОВ

Новый подход к предсказуемости начала и завершения муссона основан на принципах анализа критических переходов. Рассматривая начало и завершение муссона, как время наступления срыва или внезапного перехода системы от одного состояния к другому, прогнозируется время, когда критические условия для перехода к муссону (и обратно) возникнут в Центральной Индии. Этот подход включает в себя два новых метода. Сначала применяется метод определения критических элементов и затем, на основании полученных результатов, применяется метод прогноза начала и завершения муссона.

Метод определения критических элементов основан на явлении предбифуркационного роста флуктуаций накануне критических переходов. Главные положения метода были разработаны в цикле работ [Kravtsov, Surovyatkina, 2003; Majumdar et al., 2013; Surovyatkina, 2004; Surovyatkina et al., 2005]. Усиление роста флуктуаций возникает на пороге потери устойчивости (что соответствует жёсткому срыву стационарного режима) или обмена устойчивостью (мягкой смене режима). Явления роста флуктуаций наблюдают вблизи критических точек жидкостей и растворов, а также вблизи точек фазовых переходов 2-го рода в теории Ландау [Ландау, Лифшиц, 1976; Физическая..., 1988]. Возникновение этих явлений связано с аномальным ростом флуктуации и их взаимодействием — корреляцией. Области параметров состояния, в которой флуктуации велики, называют критическими областями. Сходство критических явлений в объектах различной природы позволяет рассматривать их с единой точки зрения, используя аппарат теории бифуркаций и теории катастроф [Арнольд, 1990]. Обнаружение подобных явлений при анализе данных указывает

на возможную близость системы к критическому порогу, что важно в технических системах для предотвращения аварий. Однако в природных системах, когда необходимо предсказать конкретное время перехода (в частности, начало муссона), это можно было сделать только ретроспективно [Keif et al., 2014], т. е. в прошлом по уже имеющимся данным наблюдений. В новом методе предлагается альтернатива — использовать критические флуктуации для обнаружения не времени, а места зарождения критического перехода, которые являются областями критического поведения. Далее мы будем называть эти области критическими элементами. Как же определить положение критических элементов?

Выделим три основных этапа в предлагаемом методе:

1. Выявление областей критического поведения. Проводится анализ флуктуаций наблюдаемых переменных как вдали от критического порога, так и в непосредственной близости к точке критического перехода, и выявляются регионы с максимальной дисперсией флуктуаций σ_f^2 .
2. Установление связи между критическими областями. Определяются физические основы возникновения критических условий в регионах с усиленными флуктуациями. При наличии физического обоснования такие регионы могут называться критическими элементами.
3. Определение механизмов распространения критических условий в пространстве и направления распространения критических условий.

Метод применён для определения областей критического поведения в системе Индийского муссона. Предложенная концепция не ограничивается приложением к Индийскому муссону, а может быть применена для анализа других физических систем испытывающих критические переходы.

3. РЕЗУЛЬТАТЫ ОПРЕДЕЛЕНИЯ КРИТИЧЕСКИХ ЭЛЕМЕНТОВ

Проведены исследования распределения полей характеристик приповерхностного слоя атмосферы (по данным ре-анализа NCEP/NCAR и ERA40) на полуострове Индостан. Анализом флуктуаций в распределении температуры и относительной влажности выявлены области, в которых возникает значительное усиление флуктуаций в преддверии Индийского муссона. Оказалось, что в пространственно-временной системе муссона существует два критических элемента, которые обеспечивают установление сезона дождей в центральной части Индии: один расположен в районе горного хребта Восточные Гаты (Eastern Ghats, EG), другой — в Северном Пакистане (North Pakistan, NP) (рис. 2, см. с. 58).

Восточные Гаты, а именно точка с координатами 20°N, 80°E, — это, фактически, центральный элемент образования муссона в Индии. Однако интересно, что вторым критическим элементом оказался Северный Пакистан, где муссона как такового не бывает. Второй критический элемент задаёт направление распространения муссона.

Какие физические условия возникают в этих областях? Как видно на рис. 2, накануне муссона в центре Индии возникают два кластера (EG и NP), в которых наблюдается усиление флуктуаций приповерхностной температуры,

которое обусловлено следующими причинами. Во-первых, в Восточных Гатах в этот момент Солнце — в зените (находится прямо над головой): эта область широт — самое горячее место на индийском Полуострове. Во-вторых, географическое положение Восточных Гат способствует зарождению депрессии (зоны пониженного давления), чему содействует большая разность температур между горячими воздушными массами над землёй и менее тёплыми воздушными массами над океаном. Всё вместе создаёт условие для встречи в этом месте двух мощных воздушных потоков — ветвей муссона: из Аравийского моря и из Бенгальского Залива. В результате в этой области формируется сильный циклон, который аккумулирует влагу из двух источников сразу, который разрешается сильнейшим ливнем в течении 48 часов, знаменуя собой начало муссона в центральной части Индии.

В Северном Пакистане, расположенном у подножья Гималаев, склоны которых окружают эту территорию с севера и запада, в этот момент возникает так называемое западное возмущение (Western Disturbance), которое приводит к образованию мощного антициклона. Антициклон в Северном Пакистане и циклон в Восточных Гатах (ВГ) - разные по температуре и направлению вращения воздушные потоки. Когда они соприкасаются — это приводит к сильным осадкам в ВГ. Возникновение такого циклона в ВГ — необходимое условие для начала муссона, но не достаточное.

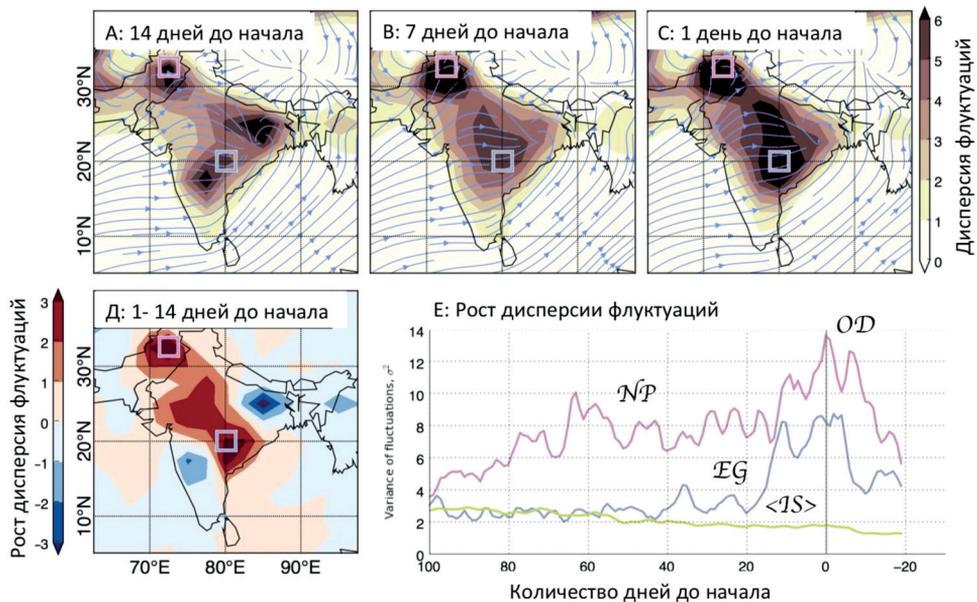


Рис. 2 Идентификация критических элементов муссона. Усиление роста флуктуаций (средние значения) в распределении приповерхностной температуры в районе полуострова Индостан в преддверии муссона по данным ERA-40 за 1971–2002 гг. (А, Б, В). Голубыми линиями обозначено распределение направления ветров. Разность в усилении роста флуктуаций (Г) между третьей (А) и последней неделей (В) перед муссоном. Предмуссонный рост флуктуаций в Восточных Гатах и Северном Пакистане и средние значения флуктуаций по полуострову (Д). Видно, что флуктуации в критических элементах увеличиваются перед муссоном, в то время как среднее значение флуктуаций снижается, не показывая предвестников начала муссона

Важно, что такая расстановка атмосферных неустойчивостей, циклона в ВГ и антициклона в СП, возникает в тот момент, когда внутритропическая зона конвергенции, прибывает на широту 20°N . Именно такая комбинация является необходимым условием для возникновения муссонных дождей в центральной части Индии, которые потом распространяются на северо-запад, в направлении к Северному Пакистану, где сезон дождей очень короткий или не возникает вообще.

Центры циклона и антициклона являются зоной атмосферных возмущений, поэтому флуктуации приповерхностной температуры возникают и растут сильнее всего в этих зонах. Таким образом, атмосферные процессы, происходящие в выявленных областях, играют важную роль в формировании условий возникновения муссона, поэтому эти области получили в нашей работе название «области критического поведения» (Tipping elements of the Indian Summer Monsoon).

Качественное обоснование существования критических элементов необходимо дополнить количественными закономерностями связывающих их в единую систему. Сравнительный анализ временных рядов приповерхностной температуры в критических элементах выявил ещё одно очень важное свойство Индийского муссона — равенство значения температур в Восточных Гатах и Северном Пакистане в момент начала сезона дождей в Восточных Гатах. В пре-муссонный период температура в Восточных Гатах выше, чем в Северном Пакистане. После наступления муссона температура в первой зоне падает, а во второй продолжает расти. Момент начала муссона — это критический порог (по температуре и относительной влажности), после которого возврат в пре-муссон уже невозможен. Эффект усиления флуктуаций исчезает. Подобное явление выравнивания температуры в исследуемых регионах возникает также и в преддверии завершения сезона дождей, потому что расстановка атмосферных неустойчивостей повторяется при возвращении внутритропической зоны конвергенции в область экватора.

Фактически, в этом исследовании впервые удалось объединить в концепции критических элементов две причины возникновения муссона: асимметричный нагрев суши и моря и сезонное перемещение внутритропической зоны конвергенции. Полученные результаты создали основу для создания нового метода прогноза начала и завершения муссона.

4. МЕТОД ПРОГНОЗА НАЧАЛА И ЗАВЕРШЕНИЯ МУССОНА

Сравнительный анализ данных внутри критических областей позволяет определять даты наступления и завершения муссона в центральной части Индии.

На первом этапе необходимо сделать качественное предсказание начала муссона. Для этого сначала необходимо провести оценку линейных трендов временных рядов в двух критических элементах и сравнить значения с трендом средней температуры за предыдущие 14 лет. В результате получим оценку раннего, нормального или позднего наступления муссона. Пример иллюстрирующий метод представлен на данных за 2012 г. на рис. 3 (см. с. 60).

На втором этапе необходимо получить дату начала муссона. Принимая во внимание, что дата начала муссона совпадает с датой, когда значения температуры в Восточных Гатах и в Северном Пакистане становятся равными

(см. рис. 3), для прогнозирования даты начала муссона нужно оценить, когда произойдёт пересечение временных рядов температуры двух критических элементов. Однако в пре-муссонный период температура в Восточных Гатах находится в режиме насыщения, и предсказать, когда она начнёт снижаться и затем упадёт до температуры в Северном Пакистане не возможно. В то же время, температуры в Северном Пакистане растёт в линейном режиме и, может быть аппроксимирована линейным трендом. Используя линейный тренд, мы можем оценить, когда температура в Северном Пакистане достигнет определённого значения, которое для Восточных Гат является критическим порогом для возникновения муссона.

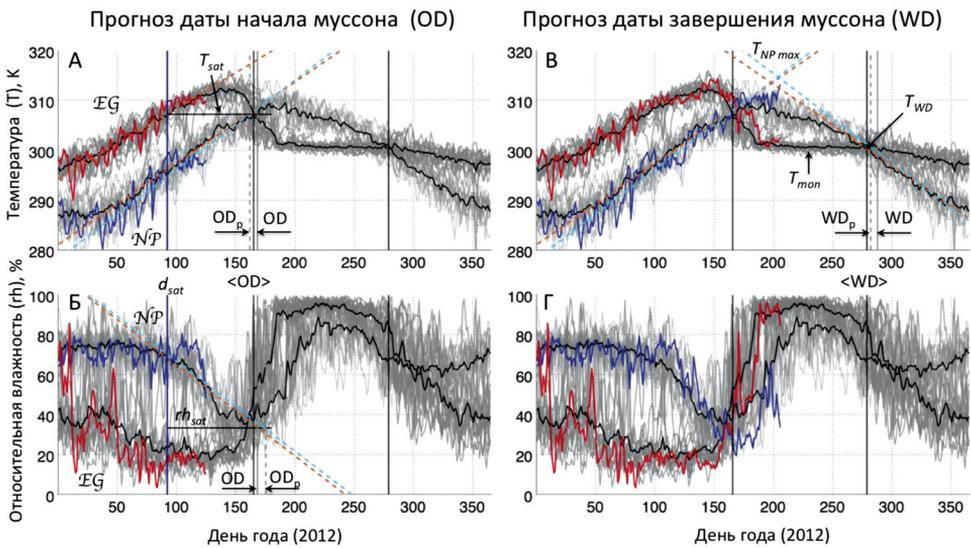


Рис. 3. Прогноз (слева) начала (А, Б) и завершения (справа) (В, Г) муссона в 2012 г. по новой методологии. Вверху — данные по приповерхностной температуре (А, В), внизу — по относительной влажности (Б, Г), данные ре-анализа NCEP/NCAR для точки (20°N, 80°E) в Восточных Гатах (EG) и для точки (32,5°N, 72,5°E) в Северном Пакистане (NP). Серым цветом показаны все данные за предыдущие 14 лет, чёрным — их средние значения. Данные за 2012 г. показаны красным для EG, голубым — для NP. День, когда тренды T и rh 2012 г. в NP (данные показаны голубым) достигают критических порогов T_{sat} и rh_{sat} , соответствует дате фактического начала муссона, завершение муссона (Б, Г) происходит в момент второго пересечения средних значений за 14 лет (данные показаны чёрными линиями), тренды их средних значений — оранжевые пунктирные линии, рассчитанный тренд для завершения муссона показан голубым цветом

Значение критического порога может быть оценено, используя среднее значение критической температуры за период 14 лет в точке пересечения временных рядов в критических элементах. Дата, когда температура в Северном Пакистане достигнет критического значения, является прогнозируемой датой начала муссона.

На третьем этапе прогнозируется дата завершения муссона. Эта оценка основана на свойстве симметрии роста и падения температуры в Северном Пакистане в течение года. Зная тренд роста температуры в пре-муссонный пе-

риод и максимальное годовое значение температуры, можно оценить время, когда температура в Северном Пакистане упадёт до температуры завершения муссона в Восточных Гатах. Значение температуры завершения муссона может быть оценено, используя среднее значение критической температуры за период 14 лет во второй точке пересечения временных рядов в критических элементах.

Долгосрочный прогноз, основанный на предложенной методологии, позволяет предсказать как дату начала муссона за месяц вперёд, так и дату завершения за два месяца вперёд для Центральной части Индии, где такие прогнозы никогда раньше не производились.

Тестирование метода прогноза начала и завершения муссона проводилось по данным ре-анализа NCEP/NCAR за период с 1965 по 2015 г. Проверка метода за 50 лет наблюдений показала положительный результат в 75 % случаях для начала муссона и в 84 % случаях для завершения муссона. В оставшихся 25 %, когда прогноз начала муссона не укладывался в критерий ± 4 дня, сильные флуктуации в пре-муссонный период не позволили получить точный тренд. Однако наличие таких флуктуаций свидетельствует о том, что приближающийся муссон будет слабым, в смысле количества осадков, что также является прогнозом. Такая информация в преддверии муссона является очень важной. Более детальные подробности тестирования представлены в работе [Stobova et al., 2016].

5. РЕЗУЛЬТАТЫ ПРОГНОЗА ПРЯДУЮЩИХ МУССОНОВ

Проведено тестирование метода долгосрочного прогноза начала и завершения летнего Индийского муссона в центральной части Индии в 2016, 2017 г. и 2018 г. Получены самые ранние прогнозы начала и завершения муссона в Индии [<https://www.pik-potsdam.de/services/infodesk/forecasting-indian-monsoon/>]. Результаты тестирования приведены в таблице.

Тестирование показало, что даты прогнозов начала муссона в 2016, 2017 и 2018 г., представленные за 40 дней до начала муссона в центральной части Индии, и даты завершения муссона, представленные за 70 дней, попали в предсказанный диапазон, что подтверждается фактическими значениями зафиксированными постфактум Департаментом метеорологии Индии (IMD).

Годы	Публикация прогноза	Прогноз	Факт	Публикация прогноза	Прогноз	Факт
	Начало муссона			Завершение муссона		
2016	6 мая	13 июня ± 4 дня	17 июня	27 июля	5 октября ± 5 дней	10 октября
2017	8 мая	18 июня ± 4 дня	16–18 июня	30 июля	12 октября ± 5 дней	15–16 октября
2018	7 мая	15 июня ± 4 дня	9–19 июня	30 июля	18 октября ± 5 дней	18–21 октября

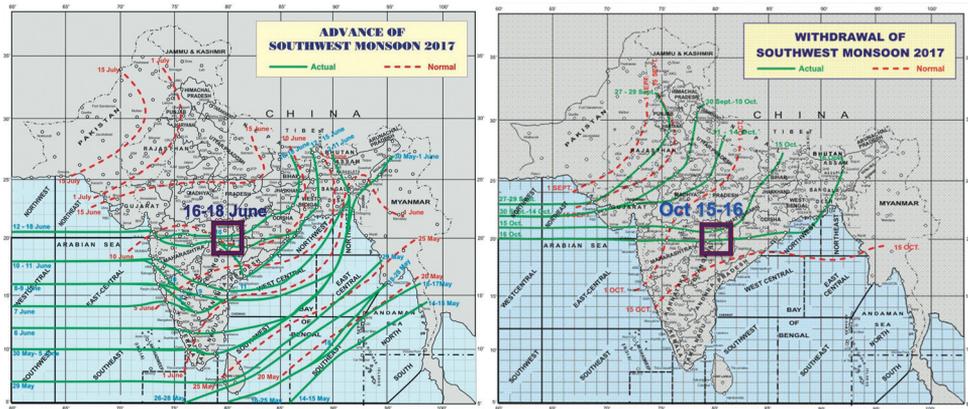


Рис. 4. Карты продвижения (слева) и выхода (справа) летнего Индийского муссона 2017 г. опубликованные Индийским метеорологическим департаментом (IMD) (http://www.imd.gov.in/pages/monsoon_main.php) на момент начала и выхода муссона из Центральной Индии (отмечено квадратом). Изохроны красного цвета показывают средние за последние 50 лет значения, рассматриваемые IMD, как норма. Изохроны зелёного цвета указывают фактические даты начала и завершения муссона в 2017 г. Даты прогнозов начала муссона, 16–18 июня (представленные за 40 дней до начала муссона в этом регионе), и даты завершения муссона, 15–16 октября (представленные за 70 дней до выхода), совпали с фактическими значениями зафиксированными IMD в Центральной Индии

Пример сравнения прогноза за 2017 г. (см. таблицу) и данных зафиксированных Департаментом метеорологии Индии представлен на рис. 4.

Таким образом, самые ранние прогнозы начала и завершения Индийского муссона в 2016, 2017 и 2018 гг. (всего шесть прогнозов) оказались успешными. Принимая также во внимание ретроспективные тесты за период 1964–2015 гг. [Stolbova et al., 2016] можно заключить, что предложенная методология долгосрочного прогноза Индийского муссона достаточно надёжна для предоставления раннего прогноза муссона в центральной части Индии.

ЗАКЛЮЧЕНИЕ

Представлен новый подход к прогнозированию начала и завершения Индийского муссона. Исследования пространственной организации муссона основаны на теории критических переходов. Такой подход базируется на анализе данных наблюдений, который показывает, что приповерхностная температура и относительная влажность демонстрируют наличие критических порогов при переходе к муссону. Этот подход включает в себя два новых метода. Сначала применяется метод определения критических элементов и затем, на основании полученных результатов, применяется метод прогноза начала и завершения муссона.

Анализом флуктуаций в распределении температуры и относительной влажности выявлены области в Восточных Гатах и Северном Пакистане, в которых возникает значительное усиление флуктуаций в преддверии Индийского

муссона. Сравнительный анализ полученных результатов с распределением направления ветров показал, что полученные области совпадают с локализацией центров циклона и антициклона, прилегающих к внутритропической зоны конвергенции, приносящей муссон в центральную часть Индии. Атмосферные процессы, происходящие в выявленных областях, играют важную роль в формировании условий возникновения муссона, поэтому эти области получили в нашей работе название «области критического поведения» (Tipping elements of the Indian Monsoon).

Полученный результат положен в основу нового метода предсказания муссона, в котором сравнительный анализ данных внутри областей критического поведения позволяет определять даты наступления и завершения Индийского муссона. Предложенный метод имеет ряд преимуществ по сравнению с существующими подходами: позволяет осуществлять долгосрочный, на 40 дней вперёд, прогноз начала муссона (вместо 15 дней существующими методами), и завершение муссона за 70 дней (завершение муссона ранее никогда не прогнозировалось, а только декларировалось постфактум), впервые прогноз осуществляется для региона, расположенного в центральной части Индостана.

Предложенная концепция критических элементов не ограничивается приложением к Индийскому муссону, а может быть применена для предсказания внезапных критических переходов в геофизике, технических системах, биофизике и медицине.

ЛИТЕРАТУРА

- [Арнольд, 1990] *Арнольд В. И.* Теория катастроф М.: Наука, 1990. 128 с.
- [Ландау, Лифшиц, 1976] *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. 3-е изд. Ч. 1. М., 1976.
- [Физическая..., 1998] *Физическая энциклопедия* / гл. ред. А. М. Прохоров. М.: Сов. энциклопедия, 1988.
- [Ananthkrishnan, Soman, 1991] *Ananthkrishnan R., Soman M. K.* The Onset of the southwest monsoon in 1990 // *Current Science*. 1991. V. 61. No. 7. P. 447–453. URL: <http://moeseprints.incois.gov.in/3797/1/onset%20of%20southwest.pdf>
- [Gadgil, 1998] *Gadgil S.* The Indian monsoon and its variability // *Annu. Rev. Earth Planet. Sci.* 1998. V. 31. P. 429–467. doi: 10.1146/annurev.earth.31.100901.141251.
- [Kalnay et al., 1996] *Kalnay E. et al.* The NCEP/NCAR 40-year reanalysis project // *Bull. Am. Meteorol. Soc.* 1996. V. 77. P. 437–440.
- [Kéfi et al., 2014] *Kéfi S. et al.* Early warning signals of ecological transitions: Methods for spatial patterns // *PLoS ONE*. 2014. V. 9(3). e92097. doi: 10.1371/journal.pone.0092097.
- [Kravtsov, Surovyatkina, 2003] *Kravtsov Yu. A., Surovyatkina E. D.* Nonlinear Saturation of Prebifurcation Noise Amplification // *Physics Letters A*. 2003. V. 319. No. 3-4. P. 348. URL: <https://www.sciencedirect.com/science/article/pii/S0375960103015640/>
- [Majumdar et al., 2013] *Majumdar A., Ockendon J., Howell P., Surovyatkina E.* Transitions through Critical Temperatures in Nematic Liquid Crystals // *Phys. Rev. E*. 2013. V. 88. 022501. URL: <https://doi.org/10.1103/PhysRevE.88.022501/>
- [Saha, Saha, 1980] *Saha S., Saha K.* A hypothesis on onset, advance and withdrawal of the Indian summer monsoon // *Pure Appl. Geophys.* 1980. V. 118. No. 2. P. 1066–1075. doi:10.1007/BF01593050.
- [Stolbova et al., 2016] *Stolbova V., Surovyatkina E., Bookhagen B., Kurths J.* Tipping elements of the Indian monsoon: Prediction of onset and withdrawal // *Geophys. Res. Lett.* 2016. V. 43. P. 1–9. doi:10.1002/2016GL068392.

- [Surovyatkina, 2004] *Surovyatkina E. D.* Phenomenon of Prebifurcation Rise and Saturation of the Correlation Time // *Physics Letters A*. 2004. V. 329. No. 3. P. 169–172.
- [Surovyatkina et al., 2005] *Surovyatkina E. D., Kravtsov Yu. A., Kurths J.* Fluctuation growth and saturation in nonlinear oscillator on the threshold of bifurcation of spontaneous symmetry breaking // *Phys. Rev. E*. 2005. V. 72. 046125.
- [Uppala et al., 2005] *Uppala S. M.* et al. The ERA-40 re-analysis // *Quarterly J. Royal Meteorological Soc.* 2005. V. 131. No. 612. P. 2961–3012. doi: 10.1256/qj.04.176.

Данные ERA-40 доступны по ссылке:

<http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html>

Данные NCEP/NCAR доступны по ссылке:

<http://apps.ecmwf.int/datasets/data/era40-daily/>

Прогнозы муссонов за период 2017–2018 гг. публиковались на сайте:

<https://www.pik-potsdam.de/services/infodesk/forecasting-indian-monsoon/>

INDIAN MONSOON: TIPPING ELEMENTS APPROACH, FORECAST OF ONSET AND WITHDRAWAL

E. D. Surovyatkina

Space Research Institute of Russian Academy of Sciences (IKI RAN)

The Indian monsoon is an important component of Earth's climate. There are two essential factors for monsoon formation: local, seasonal changes in the atmospheric circulation caused by asymmetric heating of land and sea, and the global meridional movement of the Intertropical Convergence Zone. The important feature of the monsoon is that it starts and ends suddenly. The time of monsoon season's onset and withdrawal varies within a month from year to year. Hence, the monsoon prediction is a considerable scientific challenge with great importance for society. The longterm forecast of monsoon is of vital importance for the population of India, where 70 % of its people directly related to farming. Agriculture, in turn, is dependent on the monsoon.

In our study, we make a step forward in this direction by proposing a new approach forecasting Indian monsoon based on the theory of critical transitions. We developed a new method to detect the origins of critical conditions in spatially distributed systems. These regions are system-forming elements — Tipping Elements of Indian Summer Monsoon. The critical elements localize under the shape of the Intertropical Convergence Zone and the orography of Hindustan. Thus, the concept of critical elements connects two different descriptions of the monsoon phenomenon, based on both global and local factors.

We show that monitoring of the temporal evolution of a system in the origins of the critical condition opens a possibility for prediction of the timing of the upcoming critical transition. We have found the regularities between the Tipping Elements allow us predicting the timing of the forthcoming monsoon onset and withdrawal for 40 and 70 days in advance respectively.

The proposed approach is applicable to different kind of season, which exhibits properties of critical transition. The prediction is based on observational data only when the model cannot accurately anticipate the transition or does not exist yet.

Keywords: Indian monsoon, critical phenomena, long-term forecast

Surovyatkina Elena Dmitrievna — leading researcher, doctor of physical and mathematical sciences, selena@iki.rssi.ru

ОБ ОРБИТАЛЬНЫХ ТОРАХ ВО ВРАЩАЮЩИХСЯ СИСТЕМАХ КООРДИНАТ

В. И. Прохоренко

Институт космических исследований Российской академии наук
(ИКИ РАН), Москва

Понятие орбитальный тор было введено автором около 35 лет назад при выборе высокоапогейных орбит ИСЗ, исходя из задач космических экспериментов, нацеленных на исследование околоземного космического пространства. Именно эти задачи, решения которых основаны на ситуационном анализе, исходя из математических моделей исследуемых областей околоземного космического пространства, способствовали разработке орбитальных торов. В этих задачах ключевую роль играют суточные и годовые орбитальные торы, которые образуются орбитой спутника в результате суточного вращения Земли, а также годового движения Земли вокруг Солнца. Орбитальный тор это пространственная фигура, которую описывает орбита спутника планеты во вращающейся системе координат при начальных значениях орбитальных элементов, заданных в инерциальной (невращающейся) системе координат. Работа посвящена описанию геометрии пространственных орбитальных торов в планетоцентрической вращающейся системе координат, в зависимости от начальных значений орбитальных элементов, заданных в планетоцентрической инерциальной системе координат.

Ключевые слова: Кеплеровы орбитальные элементы, угловые орбитальные элементы Эйлера, инерциальные и вращающиеся планетоцентрические системы координат, орбитальные торы во вращающихся системах координат

ВВЕДЕНИЕ

«Цель расчётов не числа, а понимание»

Ричард Хемминг

В настоящей работе автор ставит перед собой задачу поделиться опытом в области ситуационных исследований, как в процессе выбора орбит спутников, исходя из задач космических проектов в солнечно-земной физике, так и при навигационном сопровождении космических экспериментов. Основываясь на этом опыте, автор предлагает по-новому взглянуть на выбор орбит для будущих космических экспериментов. Стремление достигнуть всеобщего взаимопонимания — вот тот мотив, которым руководствуется автор в своих начинаниях.

Основной вывод, вытекающий из опыта решения рассматриваемых задач, состоит в том, что «особенности эволюции» орбиты спутника и время её существования определяются начальными значениями пяти орбитальных элементов, которые определяют положение орбиты в планетоцентрической инерциальной системе координат, а также датой старта, от которой зависят особенности эволюции орбитальных элементов.

К орбитальным элементам относятся большая полуось и эксцентриситет орбиты (элементы Кеплера), а также три угловых орбитальных элемента Эйлера, задающих положение орбиты в трёхмерном пространстве.

Прохоренко Виктория Ивановна — старший научный сотрудник, кандидат физико-математических наук, vprokhorenko@mail.ru

К угловым орбитальным элементам относятся: наклонение орбиты к плоскости экватора планеты; аргумент широты перицентра, измеряемый в плоскости орбиты спутника от восходящего узла, соответствующего пересечению орбиты с плоскостью экватора планеты; а также прямое восхождение восходящего узла орбиты спутника, измеряемое в плоскости экватора планеты относительно точки «весеннего равноденствия».

Из пяти упомянутых орбитальных элементов, в качестве свободного параметра при выборе орбит для космических экспериментов, можно было рассматривать прямое восхождение восходящего узла орбиты. На остальные орбитальные элементы, как правило, накладывались определённые ограничения.

В качестве инструмента для решения различных задач ситуационного анализа в восьмидесятые годы прошлого века автором были разработаны и использованы орбитальные торы. Геометрии орбитальных торов было посвящён препринт ИКИ АН СССР (Прохоренко, 1983), который теперь стал библиографической редкостью. Предвидя более широкое применение орбитальных торов в астрофизике, автор считает своим долгом привести более полное и более наглядное описание понятия орбитальный тор, основываясь на накопленном опыте и на указанной публикации.

1. ГЕОМЕТРИЯ ОРБИТАЛЬНЫХ ТОРОВ В ПЛАНЕТОЦЕНТРИЧЕСКОЙ ЭКВАТОРИАЛЬНОЙ СИСТЕМЕ КООРДИНАТ, ВРАЩАЮЩЕЙСЯ ВОКРУГ ОСИ OZ

Начнём с начальных значений Кеплеровых орбитальных элементов a_0 , e_0 , описывающих форму орбиты:

- a — большая полуось орбиты;
- e — эксцентриситет орбиты.

Положение орбиты в пространстве описывают Эйлеровы угловые орбитальные элементы в инерциальной (невращающейся) планетоцентрической системе координат $Oxuz$, плоскость Oxu которой совпадает с плоскостью экватора планеты, а ось Ox соответствует линии пересечения плоскости экватора планеты с плоскостью *эклиптики*, (соответствующей плоскости движения планеты вокруг Солнца) и направлена в точку «весеннего равноденствия».

Эйлеровы угловые орбитальные элементы:

- i — наклонение плоскости обиты к плоскости Oxu ;
- ω — аргумент широты точки перицентра, измеряемый в плоскости орбиты от восходящего узла орбиты, лежащего на линии её пересечения с плоскостью Oxu ;
- Ω — прямое восхождение восходящего узла AN (ascending node) орбиты, измеряемое в плоскости Oxu относительно оси Ox . Угол Ω определяет положение линии узлов (AN , DN), соединяющей восходящий и нисходящий узел (descending node) линии пересечения орбиты с плоскостью экватора планеты.

Области возможных начальных значений орбитальных упомянутых орбитальных элементов: $e \in [0, 1]$, $i \in [0^\circ, 90^\circ]$, $\omega \in [0^\circ, 360^\circ]$, $\Omega \in [0^\circ, 360^\circ]$.

Орбитальный тор это пространственная фигура, поверхность которой образует орбита с фиксированными начальными значениями четырёх орбитальных элементов: a_0, e_0, i_0, ω_0 в прямоугольной инерциальной системе координат $Oxyz$ в результате вращения вокруг оси Oz .

В невращающейся (инерциальной) системе координат орбитальный тор образует семейство орбит с фиксированными начальными значениями четырёх орбитальных элементов: a_0, e_0, i_0, ω_0 при свободном значении параметра Ω_0 , покрывающем область его возможных значений от 0 до 360° .

Во вращающейся вокруг оси Oz системе координат орбитальный тор образуется, исходя из орбитального эллипса, сохраняющего (в первом приближении) свою форму и положение в инерциальном пространстве.

Например, «суточный» орбитальный тор образуется в планетоцентрической системе координат $Oxyz$ ($Pmeq$), связанной с экватором планеты, у которой ось Oz ($Pmeq$) совпадает с осью вращения планеты, а плоскость Oxz ($Pmeq$) совпадает с полуденной меридиональной плоскостью. В планетоцентрической солнечно-эклиптической системе координат $OXYZ$ (se), у которой плоскость OXY (se) совпадает с плоскостью эклиптики, а ось OX (se) направлена на Солнце, орбитальный эллипс образует «годовой» орбитальный тор.

Таким образом, суточный орбитальный тор образуется путём вращения орбитального эллипса вокруг оси вращения планеты в планетоцентрической полуденно-экваториальной ($Pmeq$) системе координат. Годовой орбитальный тор — путём вращения орбитального эллипса в планетоцентрической солнечно-эклиптической (se) системе координат вокруг оси, перпендикулярной к плоскости эклиптики. Оба вращения происходят в направлении, обратном вращению системы координат (против часовой стрелки).

Прежде чем переходить к описанию формирования орбитального тора в пространстве, вид которого зависит от первых четырёх указанных выше орбитальных элементов, покажем, как, исходя из массы центрального тела m_1 , формируется орбитальный эллипс и плоскость орбиты тела массы m_2 .

Начнём с фиксированного значения расстояния перицентра r_p , исходя из которого, можно построить планетоцентрическую сферу радиуса r_p . Вектор скорости V , лежащий в плоскости, перпендикулярной к радиус-вектору произвольной точки сферы радиуса r_p , определяет плоскость орбиты тела m_2 , а модуль вектора скорости V_{rp} определяет форму орбиты (её эксцентриситет), исходя из следующего соотношения:

$$\frac{V_{rp}^2}{\mu_{1,2}} = \frac{(1+e)}{r_p},$$

где $\mu_{1,2} = G(m_1 + m_2)$, а G — гравитационная постоянная.

Исходя из значения модуля скорости в точке перицентра V_{rp} , можно определить эксцентриситет орбиты

$$e = \frac{r_p V_{rp}^2}{\mu_{1,2}} - 1.$$

При значении $e = 0$ орбита является круговой, при значении $0 < e < 1$ — эллиптической, а при значении $e = 1$ — параболической.

Исходя из значений r_p и эксцентриситета e , можно определить значения большой полуоси орбиты a , фокального параметра p и скорости V_{rp} в точке перицентра, исходя из следующих выражений:

$$a = \frac{r_p}{(1-e)}, \quad V_{rp} = \pm \sqrt{\frac{\mu_{1,2}(1+e)}{r_p}}, \quad p = r_p(1+e).$$

Напомним, что фокальный параметр p , который через a и e выражается формулой $p = a(1 - e^2)$, — это планетоцентрическое расстояние двух точек орбиты, лежащих на оси Oy , перпендикулярной к линии апсид.

На рис. 1 показаны эллиптические и параболические орбиты, построенные, исходя из фиксированных значений (указанных на рисунках) планетоцентрического расстояния r_p . На этом рисунке и на всех последующих в качестве единицы измерения расстояния используется средний радиус планеты (средний радиус планеты Земля составляет 6371,2 км).

На каждом рисунке показаны две эллиптические орбиты и две параболические орбиты, построенные исходя из указанного на рисунке значения планетоцентрического расстояния перицентра r_p . Каждая из этих двух пар орбит отличается направлением вектора скорости в точке перицентра.

Первая пара орбит (эллиптическая и параболическая) соответствует точке перицентра с координатами $x = r_p, y = 0$. Этой точке перицентра соответствует вектор скорости параллельный оси Oy (и направленный в положительном направлении). Участок орбиты, соответствующий движению от перицентра к апоцентру, окрашен зелёным цветом. А синим цветом, окрашен участок орбиты, соответствующий движению от апоцентра к перицентру. Стрелкой указано направление движения по эллиптической орбите.

Вторая пара орбит соответствует точке перицентра с координатами $x = -r_p, y = 0$. В этом случае точке перицентра соответствует вектор скорости параллельный оси Oy (и направленный в отрицательном направлении). Участки орбит, соответствующие движению от перицентра к апоцентру (и от апоцентра к перицентру), окрашены теми же цветами, что и в первом случае, а стрелкой указано направление движения по орбите. Во всех случаях движение по орбитам происходит в направлении против часовой стрелки.

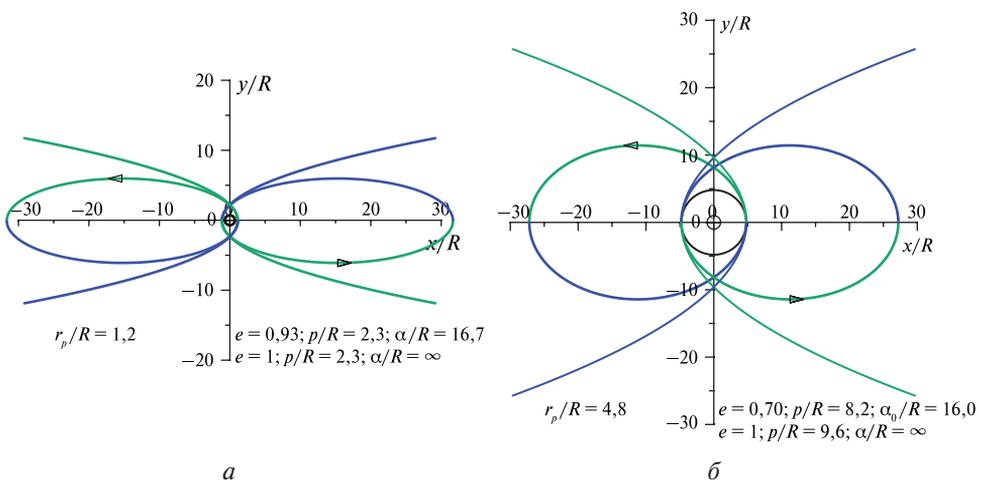


Рис. 1. Геометрия семейства орбит ограниченной задачи двух тел с фиксированным значением расстояния перицентра и различными направлениями скорости в точках перицентра

На рисунках показаны значения эксцентриситета e , фокального параметра p и большой полуоси a , для эллиптической орбиты и для параболической орбиты (соответствующей значению эксцентриситета, равному единице).

Отметим, что на рис. 1а использовано значение r_p (и эксцентриситета e), соответствующее высоко апогейным орбитам спутников серии «Прогноз», которые были запущены в семидесятые годы XX в., а на рис. 1б — экспериментальное значение r_p , которое мы будем использовать в дальнейшем для иллюстрации построения орбитальных торов.

Прежде чем переходить к формированию орбитального тора, исходя из фиксированных начальных значений орбитальных элементов, начнём с описания положения кеплеровского орбитального эллипса в планетоцентрической системе координат, исходя из заданных начальных значений первых четырёх орбитальных элементов.

Планетоцентрическое расстояние произвольной точки орбиты, заданной аргументом широты u , измеряемым относительно восходящего узла AN , принадлежащего линии пересечения плоскости орбиты с плоскостью экватора планеты, определяется следующим соотношением:

$$r(u) = \frac{a_0(1 - e_0^2)}{1 + e_0 \cdot \cos(u - \omega_0)}. \quad (1)$$

Широта соответствующей точки орбиты определяется из соотношения

$$\sin \varphi = \sin u \cdot \sin i_0. \quad (2)$$

Отсюда следует, что в точке перицентра значения аргумента широты u_p , планетоцентрического расстояния r_p и широты φ_p определяются формулами:

$$u_p = \omega_0, \quad r_p = a_0(1 - e_0), \quad \sin \varphi_p = \sin \omega_0 \cdot \sin i_0,$$

а в точке апоцентра значения аргумента широты u_A , планетоцентрического расстояния r_A и широты φ_A определяются формулами:

$$u_A = \omega_0 - 180^\circ, \quad r_A = a_0(1 + e_0), \quad \sin \varphi_A = \sin(\omega_0 - 180^\circ) \sin i_0.$$

На рис. 2 показано формирование орбитального тора для семейства орбит с фиксированным начальным значением большой полуоси $a_0 = 16R$, где $R = 6371,2$ км — средний радиус планеты (Земля), и начальными значениями двух орбитальных элементов (указанными на рис. 2в) при свободном значении долготы восходящего узла Ω_0 в планетоцентрической инерциальной системе координат $Oxyz$.

На рис. 2а (см. с. 70) показано положение орбитального эллипса в плоскости орбиты относительно линии узлов (линии пересечения плоскости орбиты с плоскостью экватора планеты). Ось O_{xN} соответствует восходящему узлу AN . Точки перицентра и апоцентра орбиты обозначены буквами P и A . Соединяющая эти точки линия апсид, показана штрихпунктирной линией. Длина линии апсид равна удвоенному значению большой полуоси $2a_0$. Положение линии апсид относительно линии узлов определяется значением аргумента широты u_A , измеряемого в плоскости орбиты против часовой стрелки от восходящего узла орбиты. Значения аргумента широты апоцентра u_A и перицентра ω_0 отличаются на 180° :

$$u_A = \omega_0 - 180^\circ.$$

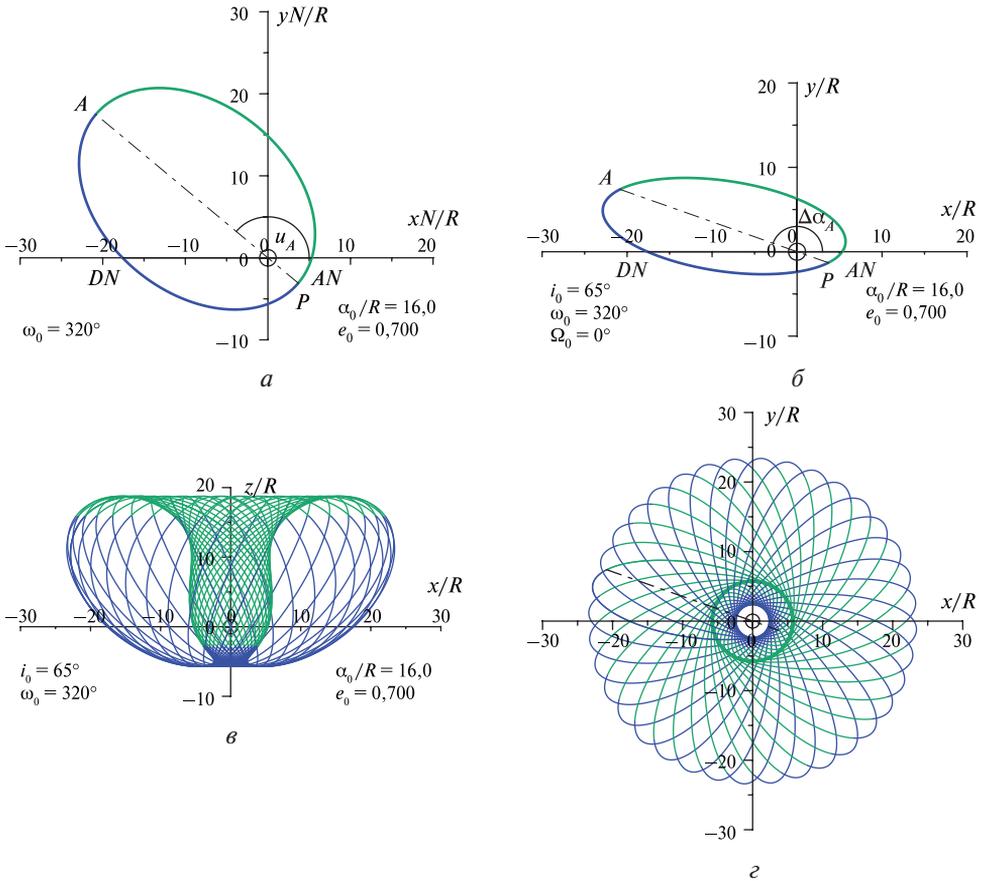


Рис. 2. Орбитальный тор семейства орбит со свободным значением углового элемента Ω_0

Участок обиты, соответствующий движению от перигея к апогею, показан линией зелёного цвета, а участок, соответствующий движению от апогея к перигею, — линией синего цвета.

При прохождении через восходящий узел AN спутник переходит из южного полушария в северное полушарие, а при прохождении через нисходящий узел DN спутник переходит из северного полушария в южное.

На рис. 2б показана проекция орбиты на плоскость экватора при $\Omega_0 = 0^\circ$, а штрихпунктирной линией показана проекция линии апсид. Угловое расстояние проекции линии апсид от восходящего узла AN определяется следующей формулой:

$$\operatorname{tg} \Delta \alpha_A = \operatorname{tg} u_A \cdot \cos i_0.$$

На рис. 2в показано семейство орбит со свободным значением углового элемента Ω_0 , изменяющегося от 0 до 360° с шагом 10° (вид со стороны оси $-Oy$).

На рис. 2г показан вид этого семейства орбит сверху, со стороны оси Oz . Утолщённой линией показана проекция орбиты, соответствующей значению $\Omega_0 = 0^\circ$, а штрихпунктирной линией показана соответствующая проекция линии апсид.

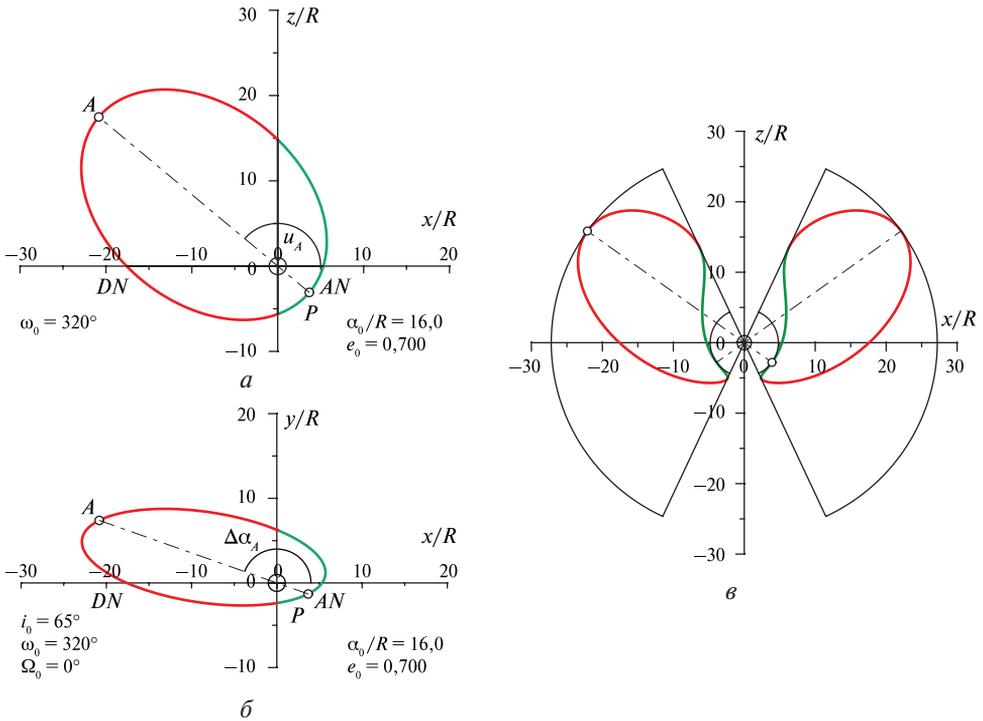


Рис. 3. Формирование орбитального тора, исходя из начальных значений четырёх орбитальных элементов

Теперь перейдём к построению аналогичного орбитального тора во вращающейся системе координат. Для этого воспользуемся рис. 3, на котором будем использовать те же начальные значения первой четвёрки орбитальных элементов, которые использованы при построении рис. 2.

Во вращающейся системе координат плоскость орбиты вращается в противоположном направлении. Положение каждой точки орбиты в её плоскости определяется значением аргумента широты $u \in [0^\circ, 360^\circ]$ и каждая точка орбиты описывает окружность радиуса $r(u) \cdot \cos\varphi(u)$, оставаясь на высоте $r(u) \cdot \sin\varphi(u)$ над плоскостью $z = 0$. Широта точки, заданной аргументом широты u , определяется из соотношения (2). Значения $u \in (-90^\circ, 90^\circ)$ определяют восходящую ветвь орбиты, которой соответствуют значения $\varphi(u) \in (-i_0, +i_0)$, а значения $u \in (90^\circ, 270^\circ)$ определяют нисходящую ветвь орбиты, которой соответствуют значения $\varphi(u) \in (i_0, -i_0)$.

Следующие соотношения

$$\sqrt{x^2 + y^2} = r(u) \cdot \cos(\varphi), \quad z = r(u) \cdot \sin(\varphi), \quad (3)$$

вместе с соотношениями (1) и (2) представляют собой параметрическое уравнение поверхности вращения, которая образуется орбитой в системе координат $Oxyz$.

Вид меридионального сечения этой поверхности, показанной на рис. 3в, даёт основание назвать её орбитальным тором. Меридиональные сечения полученного орбитального тора, соответствующего фиксированному начальному

значению ω_0 , показаны в плоскости Oxz . Восходящая ветвь показана линией зелёного цвета. Нисходящая ветвь — линией красного цвета. В точках, соответствующих граничным значениям широты $\varphi = \pm i_0$, орбитальный тор касается конуса, ось которого совпадает с осью Oz , а угол раствора этого конуса определяется формулой $\beta = 90^\circ - i_0$.

На этом рисунке кружочками обозначены начальные положения точки перицентра и точки апоцентра, а соединяющая их линия апсид, показанная штрихпунктирной линией, «говорит» о том, что эти точки лежат в разных меридиональных полуплоскостях.

Значения широты апоцентра и перицентра при отличных от нуля значениях наклона орбиты имеют разные знаки, и определяются значением аргумента широты перицентра ω_0 , а линия апсид описывает конус, угол раствора которого β_A определяется исходя из широты апоцентра $\beta_A = 90^\circ - \varphi_A$.

В точках, соответствующих минимальному и максимальному расстоянию орбиты от центра планеты, т.е. перицентру $r_p = a_0(1 - e_0)$ и апоцентру $r_A = a_0(1 + e_0)$, орбитальный тор касается соответствующих сфер. Через эти точки проведены окружности касания в двух меридиональных полуплоскостях.

Показанные окружности образуют сферы во вращающейся системе координат и являются границами околопланетного пространства, которые охватывает орбитальный тор при различных начальных значениях аргумента перицентра ω_0 и фиксированных начальных значениях остальных орбитальных элементов из первой четвёрки.

Итак, орбитальный тор с заданными начальными значениями четырёх орбитальных элементов находится внутри сферического кольца, определяемого значениями орбитальных элементов: a_0 , e_0 , и вне конуса, определяемого значением наклона орбиты i_0 . Сравнивая рис. 3в и 2в, мы видим, что меридиональное сечение орбитального тора определяет «силуэт», показанный на рис. 2в.

Орбитальные торы являются инструментом для представления орбиты малого тела в пространстве. Далее мы покажем, как форма орбитального тора зависит от начальных значений угловых орбитальных элементов, и рассмотрим некоторые примеры «экзотических» вариантов орбитальных торов.

2. ГЕОМЕТРИЯ СЕМЕЙСТВА ОРБИТАЛЬНЫХ ТОРОВ СО СВОБОДНЫМ НАЧАЛЬНЫМ ЗНАЧЕНИЕМ АРГУМЕНТА ШИРОТЫ ПЕРИЦЕНТРА

Рассмотрим различные семейства орбитальных торов при фиксированных начальных значениях первых трёх орбитальных элементов a_0 , e_0 , i_0 .

На рис. 4а (см. с. 73) показано семейство орбит, соответствующих начальным значениям $\omega_0 \in (-90^\circ, 90^\circ)$, принадлежащим IV и I четвертям, при $i_0 = 90^\circ$ и $\Omega_0 = 0^\circ$ (в этом случае плоскость орбиты совпадает с плоскостью Oxz).

Положения перицентра и апоцентра начальной орбиты отмечены кружочками голубого цвета. Здесь, как и на всех последующих рисунках, восходящая ветвь орбиты показана зелёным цветом, а нисходящая ветвь — красным цветом. Сечение плоскостью Oxz соответствующего семейства орбитальных торов, при начальном значении наклона орбиты $i_0 = 90^\circ$, показано на рис. 4б.

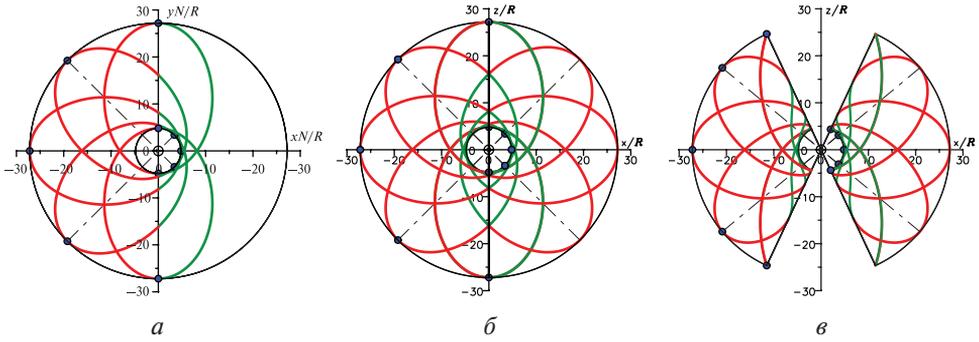


Рис. 4. Два семейства орбитальных торов, соответствующие значениям $\omega_0 \in (-90^\circ, 90^\circ)$ с шагом 45° при различных значениях наклона i_0 : *a* — орбиты в их плоскости; *б* — торы при $i_0 = 90^\circ$; *в* — торы при $i_0 = 65^\circ$

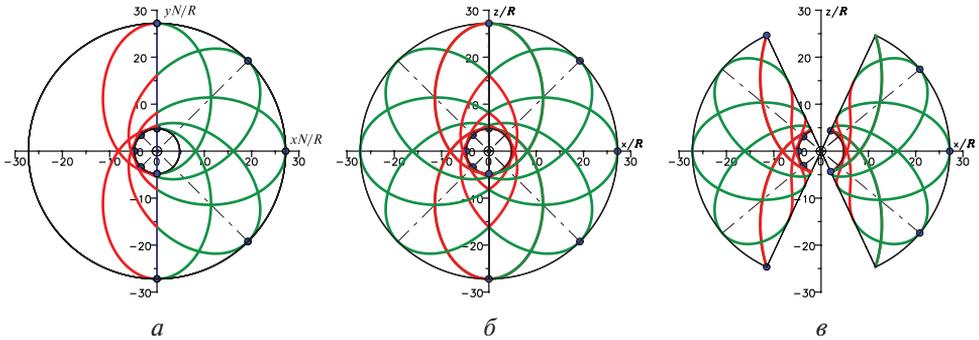


Рис. 5. Два семейства орбитальных торов, соответствующие значениям $\omega_0 \in (90^\circ, 270^\circ)$ с шагом 45° , при различных значениях наклона i_0 : *a* — орбиты в их плоскости; *б* — торы при $i_0 = 90^\circ$; *в* — торы при $i_0 = 65^\circ$

На рис. 4*в* показано сечение плоскостью Oxz того же семейства орбитальных торов, при начальном значении наклона орбиты $i_0 = 65^\circ$.

На рис. 5 представлена аналогичная картина для семейства орбит с начальными значениями аргумента широты перигея ω_0 , принадлежащими II и III четвертям. В этом случае используются начальные значения $\omega_0 \in (90^\circ, 270^\circ)$ с тем же шагом.

На рис. 5*а* это семейство орбит показано при $i_0 = 90^\circ$. На рис. 5*б* показано сечение плоскостью Oxz соответствующего семейства орбитальных торов при значении $i_0 = 90^\circ$. На рис. 5*в* аналогичное семейство орбитальных торов показано при $i_0 = 65^\circ$.

Эти рисунки являются наглядной иллюстрацией следующего факта: при начальных значениях аргумента перигея ω_0 , принадлежащих IV или I четверти (правой полуплоскости), точка перигея находится на восходящей ветви орбиты, а точка апогея на нисходящей ветви орбиты. При начальных значениях аргумента перигея ω_0 , принадлежащих II или III четверти (левой полуплоскости), наоборот, точка перигея находится на нисходящей ветви орбиты, а точка апогея на восходящей ветви. Отметим, что в первом случае (показанном на рис. 4) внутренняя поверхность орбитального тора соответствует восходящей ветви орбиты и окрашена зелёным цветом, а внешняя

поверхность, соответствующая нисходящей ветви орбиты, окрашена красным цветом. Во втором случае (показанном на рис. 5) наоборот, внутренняя поверхность орбитального тора соответствует нисходящей ветви орбиты и окрашена красным цветом, а внешняя поверхность соответствует восходящей ветви орбиты и окрашена зелёным цветом.

На этих рисунках также показаны орбитальные торы, соответствующие начальным значениям аргумента перицентра, лежащим на границах четвертей. Заметим, что при значениях ω_0 , соответствующих переходу перицентра орбиты через плоскость экватора планеты, положение орбиты (в её плоскости) и геометрия орбитальных торов получают зеркальное отражение.

На этих рисунках, среди прочих орбит, показаны орбиты, соответствующие значению $\omega_{04} = 270^\circ$, лежащему на границе четвертей III и IV, а также орбиты, соответствующие значению $\omega_{02} = 90^\circ$, лежащему на границе четвертей I и II. В этих случаях линия аспид перпендикулярна к плоскости экватора планеты. И при начальном значении аргумента перицентра равном ω_{04} , перицентр орбиты r_p находится в «южном» полушарии, а при значении аргумента перицентра равном ω_{02} , перицентр орбиты r_p находится в «северном» полушарии.

На этих рисунках показаны также орбиты, соответствующие значению $\omega_{01} = 0^\circ$, лежащему на границе четвертей I и IV, и орбиты, соответствующие значению $\omega_{03} = 180^\circ$, лежащему на границе четвертей II и III. В этих случаях линия аспид лежит в плоскости экватора планеты, и при значениях аргумента перицентра ω_{01} перицентр r_p находится в области положительных значений x (в правой полуплоскости), а при значениях аргумента перицентра ω_{03} перицентр r_p находится в области отрицательных значений x (в левой полуплоскости).

Легко видеть, что поверхности орбитальных торов, соответствующих значениям аргумента перицентра ω_{01} и ω_{03} , геометрически совпадают, только внешняя и внутренняя поверхности окрашены в разные цвета. Отметим, что эти два варианта орбит совпадают с орбитами, показанными в самом начале, на рис. 1б. Круг замкнулся (мы вернулись к началу описания геометрии орбитальных торов).

3. О ТРАЕКТОРИИ СПУТНИКА НА ОРБИТАЛЬНОМ ТОРЕ

Для того чтобы определить траекторию спутника на торе, необходимо кроме упомянутых выше параметров, определяющих геометрию орбитального тора и положение плоскости орбиты, ввести ещё один параметр, характеризующий положение спутника на орбите. Для этого используем значение момента времени τ прохождения спутника через перицентр орбиты. Тогда (в рамках той или иной модели движения) можно однозначно определить положение спутника в пространстве и во времени. Сферические координаты $r(t)$, $\varphi(t)$, $\lambda(t)$ описывают положение спутника на торе: $r(t)$ и $\varphi(t)$ — в меридиональной полуплоскости, а $\lambda(t)$ — положение этой полуплоскости.

В соответствии с третьим законом Кеплера, период обращения спутника по орбите удовлетворяет соотношению

$$T = \frac{2\pi a^{3/2}}{\sqrt{\mu_{1,2}}}.$$

Ограничиваясь кеплеровским приближением, для любого момента времени можно определить положение спутника в момент времени t . Для этого, исходя из уравнения Кеплера

$$t = \tau + \frac{a^{3/2}}{\sqrt{\mu_{1,2}}}(E - e \cdot \sin E),$$

для любого момента времени можно определить эксцентрическую аномалию $E(t)$, а затем определить угловое положение $u(t)$ спутника на орбите, исходя из следующего соотношения, приведённого в книге [Эльясберг, 1965],

$$u(t) = \omega + \arctan \frac{\sqrt{(1-e^2)} \cdot \sin E(t)}{\cos E(t) - e}.$$

Далее $r(t)$ и $\varphi(t)$ определяются из соотношений (1) и (2), а $\lambda(t)$ — из следующего выражения:

$$\lambda(t) = \arctan(\tan u(t) \cos i_0) + \Omega_0 - \frac{2\pi(t - t_0)}{T_C},$$

где T_C — период вращения системы координат.

Задача о траектории спутника на торе близка к задаче о трассе спутника на поверхности вращающейся планеты. Если отношение между периодами T и T_C представляет собой рациональную дробь: $T/T_C = m/n$, где m и n — рациональные числа, то через n оборотов спутника по его орбите и m оборотов системы координат вокруг оси Oz траектория спутника на торе замыкается, т. е. движение спутника по орбитальному тору является периодическим.

Если же периоды T и T_C являются несоизмеримыми, т. е. отношение между ними представляет собой иррациональное число, то траектория спутника, навивающаяся на орбитальный тор, покрывает его поверхность всюду плотно.

Поскольку период обращения спутника по орбите зависит только от значения большой полуоси a , то и соизмеримость периода T и T_C зависит от значения этого параметра. Заметим, что сами орбитальные торы при различных значениях большой полуоси (и фиксированных начальных значениях остальных орбитальных элементов) геометрически подобны.

Однако вернёмся к орбитальным торах, которые дают возможность представить общую картину движения спутника во вращающейся системе координат, отвлекаясь от его мгновенного положения.

4. ГЕОМЕТРИЯ СЕМЕЙСТВА ОРБИТАЛЬНЫХ ТОРОВ СО СВОБОДНЫМ НАЧАЛЬНЫМ ЗНАЧЕНИЕМ НАКЛОНЕНИЯ ОРБИТЫ

Теперь рассмотрим семейство орбит с фиксированным начальным значением аргумента перицентра ω_0 при свободном начальном значении наклона орбиты $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° .

На рис. 6а (см. с. 76) при фиксированных начальных значениях $i_0 = 90^\circ$ и $\Omega_0 = 0^\circ$ показаны орбиты, соответствующие начальным значениям аргумента широты перицентра ω_0 : $\omega_{04} = 315^\circ$ и $\omega_{03} = 225^\circ$, принадлежащим соответственно IV и III четверти.

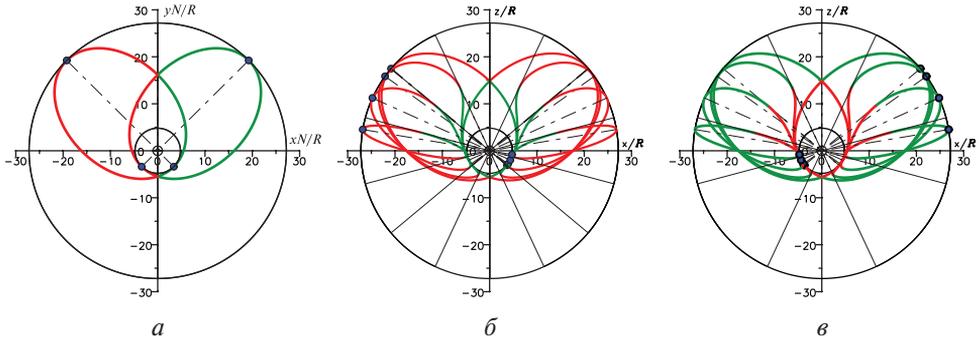


Рис. 6. Два семейства орбитальных торов при начальных значениях $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° : a — орбиты при $\omega_{03} = 225^\circ$, $\omega_{04} = 315^\circ$; b — торы при $\omega_{04} = 315^\circ$; v — торы при $\omega_{03} = 225^\circ$

Семейство орбитальных торов при начальных значениях $i_0 \in (25^\circ, 90^\circ)$ с шагом 20° , соответствующее значению $\omega_{04} = 315^\circ$, показано на рис. $6b$, а на рис. $6v$ — аналогичное семейство орбитальных торов, соответствующее значению $\omega_{03} = 225^\circ$.

На рис. $7a$ при фиксированных начальных значениях $i_0 = 90^\circ$ и $\Omega_0 = 0^\circ$ показаны орбиты, соответствующие начальным значениям аргумента широты перигея ω_0 : $\omega_{01} = 45^\circ$ и $\omega_{02} = 135^\circ$, которые принадлежат соответственно I и II четверти. Семейство орбитальных торов при начальных значениях $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° , соответствующее значению $\omega_{01} = 45^\circ$, показано на рис. $7b$, а на рис. $7v$ — аналогичное семейство орбитальных торов, соответствующее значению $\omega_{02} = 135^\circ$.

На рис. 8 (см. с. 77) показано объединение семейств орбитальных торов при начальных значениях $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° , показанных на рис. 6 и 7. На рис. $8a$ представлены орбиты, соответствующие начальным значениям аргумента широты перигея $\omega_0 \in (45^\circ, 315^\circ)$ с шагом 90° при фиксированных начальных значениях $i_0 = 90^\circ$ и $\Omega_0 = 0^\circ$. На рис $8b$ и v представлены объединения семейств орбитальных торов, показанных на рис. 6 и 7.

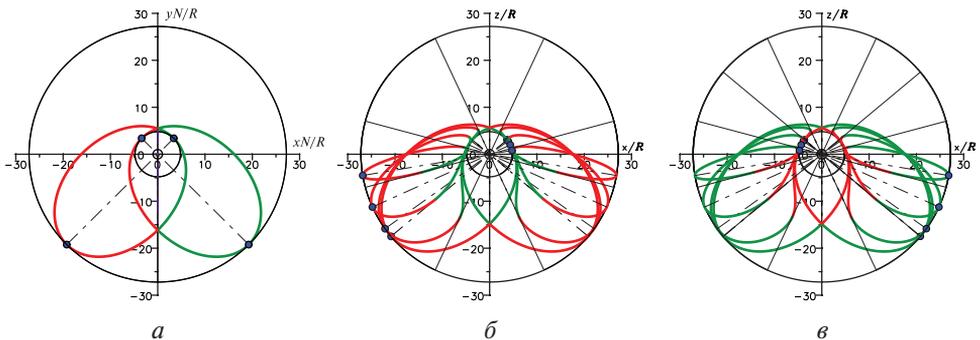


Рис. 7. Два семейства орбитальных торов с начальными значениями $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° : a — орбиты при $\omega_{01} = 45^\circ$, $\omega_{02} = 135^\circ$; b — торы при $\omega_{01} = 45^\circ$, v — торы при $\omega_{02} = 135^\circ$

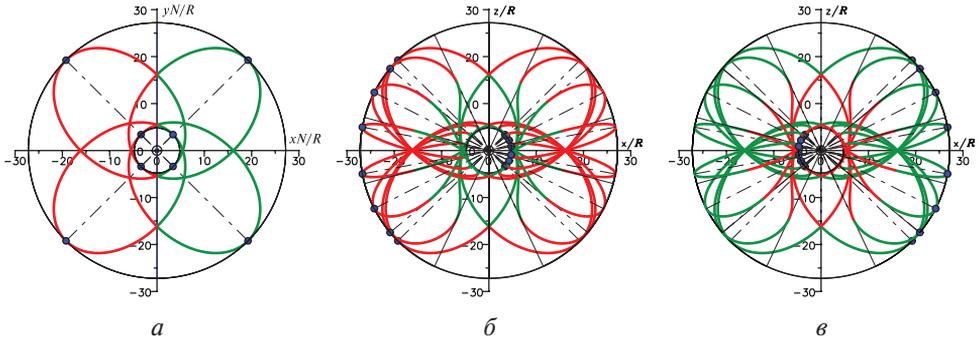


Рис. 8. Два семейства орбитальных торов при значениях $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° : *a* — орбиты при $\omega_0 \in (45^\circ, 315^\circ)$ с шагом 90° ; *б* — торы при $\omega_{04} = 315^\circ$ и $\omega_{01} = 45^\circ$, *в* — торы при $\omega_{02} = 135^\circ$ и $\omega_{03} = 225^\circ$

Сравнивая рис. 6, 7, 8, заметим, что семейства орбитальных торов, показанные на рис. 6 и 7 (и объединённые на рис. 8) отличаются зеркальным отражением относительно плоскости экватора.

Теперь рассмотрим орбиты, соответствующие начальным значениям $\omega_{02} = 90^\circ$ и $\omega_{04} = 270^\circ$. На рис. 9*a* эти орбиты показаны при фиксированных начальных значениях $i_0 = 90^\circ$ и $\Omega_0 = 0^\circ$. Первое из этих значений ω_0 лежит на границе четвертей I и II, а второе — на границе четвертей III и IV в плоскости орбиты.

На рис. 9*б* представлены сечения плоскостью Oxz для двух семейств орбитальных торов, соответствующих фиксированным начальным значениям аргумента перигея, при свободном начальном значении наклонения орбиты $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° .

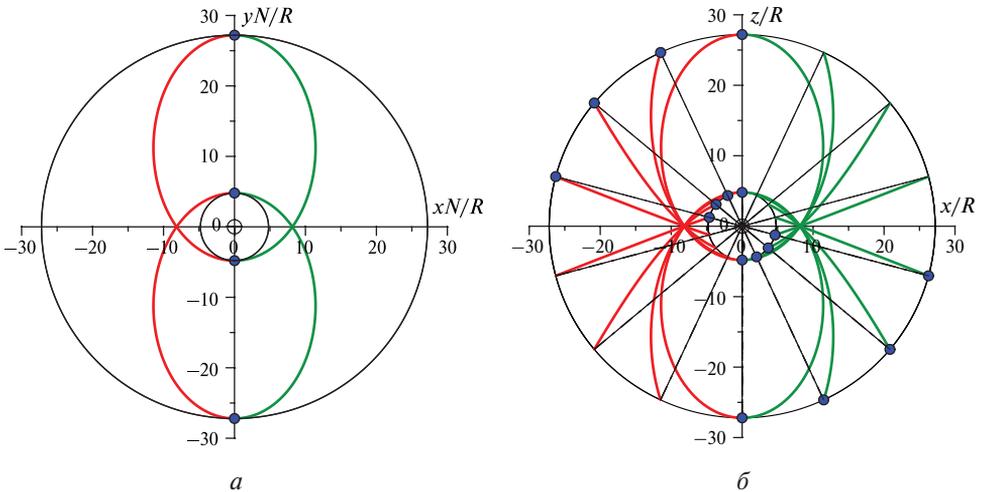


Рис. 9. Два семейства орбитальных торов, соответствующие значениям $\omega_{02} = 90^\circ$ и $\omega_{04} = 270^\circ$ при значениях $i_0 \in (30^\circ, 90^\circ)$ с шагом 20° : *a* — орбиты; *б* — орбитальные торы

В этом случае внешняя и внутренняя поверхности каждого из показанных на рис. 9б орбитальных торов геометрически совпадают, однако движение спутника по этим поверхностям происходит в противоположном направлении.

Теперь мы можем перейти к обсуждению роли орбитальных торов в ситуационном анализе.

5. СИТУАЦИОННЫЙ АНАЛИЗ ПРИ ВЫБОРЕ ОРБИТ В СОЛНЕЧНО-ПЛАНЕТНОЙ ФИЗИКЕ

Рассмотрим задачу ситуационного анализа на примере солнечно-земной физики. Для решения этой задачи используем геоцентрическую инерциальную эллиптическую систему координат $OXYZ$, плоскость OXY которой совпадает с плоскостью эклиптики, а ось OX , направлена в точку весеннего равноденствия.

На рис. 10а показано сечение плоскостью OYZ орбитального тора, полученного в системе координат $Oxyz$ (геоцентрической экваториальной), исходя из фиксированных начальных значений орбитальных элементов (измеренных в экваториальной системе координат). Начальные значения этих орбитальных элементов показаны на рис. 10а.

На рис. 10б показаны проекции на плоскость эклиптики OXY семейства орбит со свободным значением долготы восходящего узла Ω_{eq0} . Проекции линий апсид соответствующих орбит на этом рисунке показаны штрихпунктирными линиями, а значения Ω_{eq0} указаны в соответствующих точках апоцентра. На этом рисунке показана разметка шкалы значений эллиптической долготы α , измеренной относительно точки весеннего равноденствия Υ , а также даты, в которые Солнце проходит через соответствующие значения эллиптической долготы α .

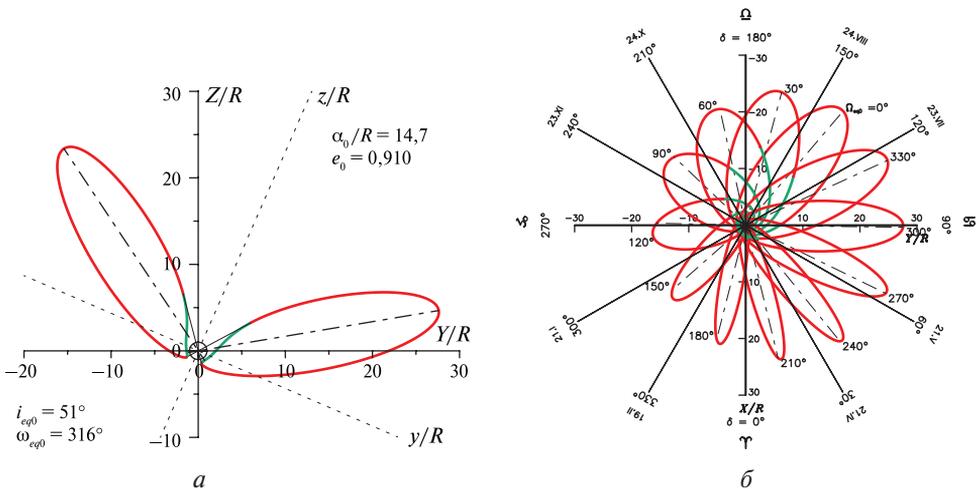


Рис. 10. Положение в инерциальной эллиптической системе координат орбитального тора при фиксированных значениях угловых элементов, измеренных в экваториальной системе координат

Следует учитывать, что разметка плоскости эклиптики по датам зависит от эпохи, которой соответствует система координат $OXYZ$. Будем использовать указанную разметку в первом приближении для понимания сути дела.

Каждая из орбит семейства со свободным значением параметра Ω_{eq0} образует в солнечно-эклиптической системе координат свой индивидуальный годовой орбитальный тор (не похожий на другие).

Все эти торы представлены на рис. 11, каждый тор обозначен значением параметра Ω_{eq0} , показанным в точке апоцентра. Орбитальные торы на этих рисунках показаны на фоне модельных границ ударной волны и магнитопаузы [Fairfield, 1971; Sibeck et al., 1991], определённых в геоцентрической солнечно-эклиптической системе координат.

Для каждого тора штрихпунктирными прямыми линиями показаны радиус-векторы точек апоцентра, а сплошными линиями показаны радиус-векторы точек, соответствующих максимальной эклиптической широте δ_{max} , равной наклонению i_0 каждой орбиты к плоскости эклиптики. Широта δ_{max} достигается в точке перехода спутника от восходящей ветви орбиты, показанной линией зелёного цвета, к нисходящей ветви, показанной линией красного цвета.

Группа орбитальных торов, показанных на рис. 11а, соответствует значениям $\Omega_{eq0} \in (150^\circ, 300^\circ)$. Для большинства орбит этой группы характерен широкий диапазон значений эклиптической широты δ , реализуемый на участках орбиты до апоцентра и после него.

Другая группа орбитальных торов, показанная на рис. 11б, соответствует значениям $\Omega_{eq0} \in (0^\circ, 120^\circ)$ и отличается довольно широким диапазоном значений наклона орбит к плоскости эклиптики, включающим минимальное значение $i_0 = 29^\circ$. Однако, для каждой из орбит этой группы характерен довольно узкий диапазон значений эклиптической широты δ , реализуемых на орбитах.

Подробностям ситуационного анализа, основанного на орбитальных торах, при выборе орбит для космических экспериментов в области магнитосферы, посвящена работа [Прохоренко, 2016].

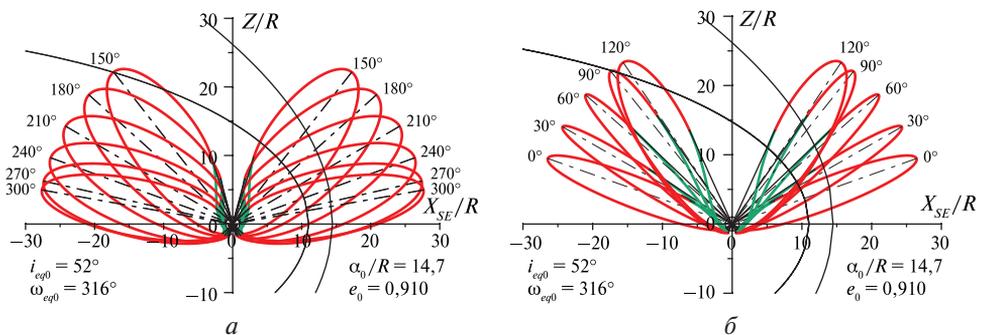


Рис. 11. Два семейства орбитальных торов в солнечно-эклиптической системе координат со свободным значением Ω_{eq0} , при фиксированных значениях четырёх орбитальных элементов, указанных на рисунках

ЗАКЛЮЧЕНИЕ

В заключение нужно сказать, что в настоящей работе даётся представление об инструменте для выбора орбит с учётом геометрии орбитальных торов и о ситуационном анализе при выборе орбит, основанном на орбитальных торах, которые наглядно демонстрируют пространственную картину достижимых орбитой областей в изучаемом пространстве.

Однако следует отметить, что орбитальные элементы, определяющие геометрию орбитального тора, не остаются постоянными, а эволюционируют под влиянием различных возмущений. Имеются различные теории, описывающие долговременную эволюцию орбитальных элементов под влиянием различных факторов, начиная от влияния плотности атмосферы планеты и полярного сжатия планеты, а также влияния внешних гравитационных возмущений, которые могут приводить к соударению спутника с планетой путём уменьшения расстояния перицентра орбиты. Однако учитывая ограниченность объёма публикации, оставим обсуждение этих весьма важных вопросов до следующей публикации. А пока можно сослаться на работы [Лидов, 1961, 1963; Лидов, Ярская, 1974] и на работы автора [Прохоренко, 2015, 2018], в которых можно найти ответы на вопросы о качественной теории возмущений и об особенностях долговременной эволюции орбит и времени баллистического существования спутника.

Заметим, что изменяющиеся во времени кеплеровские орбитальные элементы (и угловые элементы Эйлера) называются оскулирующими орбитальными элементами. Так что, исходя из оскулирующих орбитальных элементов, можно строить оскулирующие (эволюционирующие) орбитальные торы. А для построения оскулирующих орбитальных торов может быть использована настоящая работа.

В заключение хотелось бы выразить благодарность моему учителю профессору П. Е. Эльясбергу, который предложил мне в далёкие восьмидесятые годы прошлого века заняться навигационным сопровождением космических экспериментов на высокоапогейных орбитах КА серии «Прогноз». Опыт работы в этой области естественным образом привёл к ситуационному анализу при выборе орбит КА, исходя из задач проектов и с учётом особенностей эволюции этих орбит под влиянием внешних гравитационных возмущений. Хотелось бы также выразить благодарность Р. Р. Назирову, который предложил написать эту статью, учитывая мой опыт в области ситуационных исследований.

ЛИТЕРАТУРА

- [Лидов, 1961] *Лидов М. Л.* Эволюция орбит искусственных спутников планет под действием гравитационных возмущений внешних тел // Искусственные спутники Земли. 1961, № 8. С. 5.
- [Лидов, 1963] *Лидов М. Л.* О приближенном анализе эволюции орбит искусственных спутников // Проблемы движения искусственных небесных тел. 1963. С. 119–134.
- [Лидов, Ярская, 1974] *Лидов М. Л., Ярская М. В.* Интегрируемые случаи в задаче об эволюции орбиты спутника при совместном влиянии внешнего тела и нецентральности поля планеты // Космич. исслед. 1974. Т. 12. № 2. С. 155–170 (*Lidov M. L., Yarskaya M.* Integrable Cases in the Problem of the Evolution of a Satellite Orbit under the

- Joint Effect of an Outside Body and of the Noncentrality of the Planetary Field // *Cosmic Research*. 1974. V. 12. No. 2. P. 139–152).
- [Прохоренко, 1983] *Прохоренко В. И.* Орбитальные торы в задачах ситуационных исследований: препринт. 1983. Пр-770. М.: ИКИ АН СССР, 1983. 26 с.
- [Прохоренко, 2015] *Прохоренко В. И.* О применении качественных методов теории возмущений при решении практических задач выбора и коррекции орбит спутников планет с учетом вековой и долгопериодической составляющих эволюции под влиянием внешних гравитационных возмущений // *Космическая баллистика — от истоков к будущему: сб. тр.* 2015. С. 130–161.
- [Прохоренко, 2016] *Прохоренко В. И.* О выборе высокоапогейных орбит ИСЗ на основе качественных методов теории возмущений и ситуационного анализа. Ч. I. Ситуационные исследования, основанные на орбитальных торах // *Космич. исслед.* 2016. Т. 54. № 2. С. 143–163.
- [Прохоренко, 2018] *Прохоренко В. И.* О динамических основах теории Лидова – Козаи // *Письма в Астрон. журн.* 2018. Т. 44. № 1. С. 52–70.
- [Эльясберг, 1965] *Эльясберг П. Е.* Введение в теорию полета искусственных спутников Земли. М.: Физматлит, 1965. 540 с.
- [Fairfield, 1971] *Fairfield D. H.* Average and unusual location of the earth's magnetopause and bow shock // *J. Geophys. Res.* 1971. V. 76. P. 6700–6716.
- [Sibeck et al., 1991] *Sibeck D. G., Lopez R. E., Roelof E. C.* Solar wind control of the magnetopause shape, location, and motion // *J. Geophys. Res.* 1991. V. 96. P. 5489–5495.

ON ORBITAL TORI IN ROTATING COORDINATE SYSTEMS

V. I. Prokhorenko

Space Research Institute, Russian Academy of Sciences (IKI RAN)

The concept of the orbital torus was introduced by the author about 35 years ago when choosing high-apogee orbits of satellites, proceeding from the tasks of space experiments aimed at investigating near-Earth space. It is these tasks, whose solutions are based on situational analysis, on the base of mathematical models of the investigated areas of near-Earth space, contributed to the development of orbital tori. In these tasks a key role plays the daily and the annual orbital tori, which are formed by the satellite's orbit as a result of the daily rotation of the Earth, as well as of the annual motion of the Earth around the Sun. The orbital torus is a spatial figure formed by the orbit of the planet's satellite in a rotating coordinate system with the initial values of the orbital elements specified in the inertial (non-rotating) coordinate system. The paper is devoted to the description of the geometry of the spatial orbital tori in the planetocentric rotating coordinate system, depending on the initial values of the orbital elements specified in the planetocentric inertial coordinate system.

Keywords: Keplerian orbital elements, Euler's angular orbital elements, inertial and rotating planetocentric coordinate systems, orbital tori in rotating coordinate systems

Prokhorenko Victoria Ivanovna — senior scientist, PhD, vprokhorenko@mail.ru

О ПЕРЕДАЧЕ ОПТИМИЗАЦИОННОЙ ТЕОРИИ ЛИДЕРСТВА ОТ ПРИКЛАДНОЙ КЛАССИЧЕСКОЙ ТЕОРИИ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ

В. В. Золотарёв¹, Г. В. Овечкин², Р. Р. Назиров¹

¹ Институт космических исследований Российской академии наук (ИКИ РАН)

² Рязанский государственный радиотехнический университет (РГРТУ)

Обсуждаются основные методы реализации алгоритмов многопорогового декодирования (МПД), рассматриваемых как процедуры поиска глобального экстремума функционала с минимальной, т. е. линейной от длины сложностью. Из результатов сравнения возможностей Оптимизационной Теории (ОТ) с эффективностью прочих методов следует, что методы ОТ и МПД вместе с запатентованными нашей научной школой вариантами блочного алгоритма Витерби и новыми парадигмами развития теории помехоустойчивого кодирования полностью заменяют прочие методы декодирования по критериям лёгкости реализации, близости к границе Шеннона и результирующей достоверности.

Ключевые слова: помехоустойчивое кодирование, Оптимизационная Теория, блочный алгоритм Витерби, самоортогональные коды, символьные коды, каскадные коды, многопороговые декодеры, энергетический выигрыш кодирования, канал связи, граница Шеннона

Помехоустойчивое кодирование используется для исправления ошибок, возникающих при передаче информации по каналам связи с помехами [Золотарёв, 2018; Золотарёв, Овечкин, 2004; Зубарев, Овечкин, 2008; см. статью настоящего сборника *Кузнецов Н. А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99]. Для этого при отправлении исходного сообщения в него вносится некоторая избыточность (проверочная информация, полученная по тому или иному алгоритму), которая позволяет на принимающей стороне исправить значительную часть ошибок, возникающих при передаче этого сообщения по каналам связи.

Рассмотрим с общих позиций возможности алгоритмов, созданных для реализации методов достижения решений оптимальных декодеров (ОД) как задачи поиска глобального экстремума функционала (ПГЭФ) и классических методов, которые созданы в рамках алгебраической теории кодирования. Используем для этого данные, представленные в источниках [Золотарёв, 2018; см. статью настоящего сборника *Кузнецов Н. А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99]. При этом ограничимся, в основном, случаями использования моделей двоичного симметричного канала (ДСК) и канала с аддитивным белым гауссовским шумом (АБГШ), которые в течение многих десятилетий считаются главным полигоном, на котором проводится сравнение лучших методов декодирования.

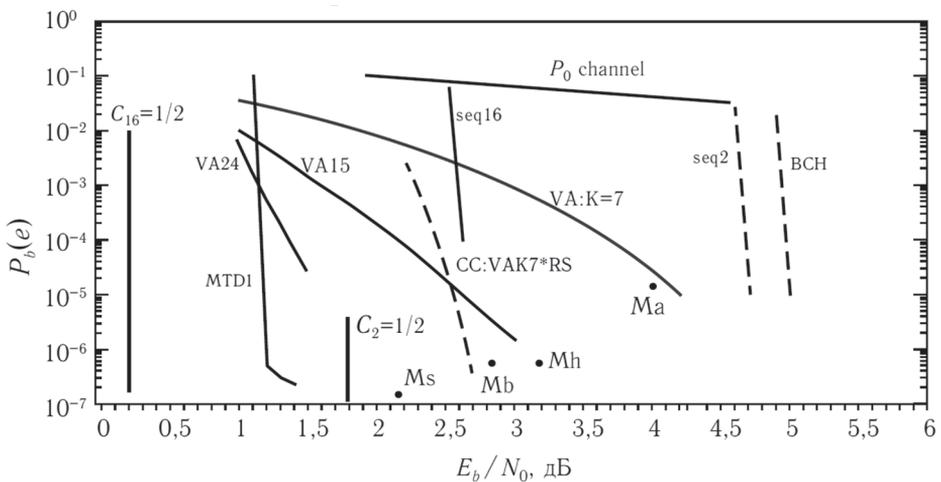
Золотарёв Валерий Владимирович — ведущий научный сотрудник, профессор, доктор технических наук, профессор, zolotasd@yandex.ru

Овечкин Геннадий Владимирович — профессор, доктор технических наук, g_ovechkin@mail.ru

Назиров Равиль Равильевич — заведующий отделом, доктор технических наук, rnazirov@cosmos.ru

Одной из самых впечатляющих революций в технике декодирования для каналов с АБГШ в 70-х гг. прошлого века, на начальном этапе развития теории кодирования, стал алгоритм Витерби (АВ) [Витерби, 1970; Золотарёв, Овечкин, 2004; Кларк, Кейн, 1987; Forney, 1974; Heller, Jacobs, 1971]. В этот период специалисты начали понемногу понимать, что никакие алгебраические методы, например, коды Боуза – Чоудхури – Хоквингема (БЧХ), Рида – Соломона (РС) и другие не смогут решить проблему эффективного простого декодирования при уровне шума, близком к пропускной способности канала. Это было время первого большого разочарования, когда выбранное интересное и важное сначала направление работ по созданию эффективных алгоритмов декодирования оказалось тупиковым. В тот момент и возникла надежда, что именно АВ и связанные с ним методы выведут исследователей на правильное направление поиска. Рассмотрим, каким образом за 50 лет разработок изменились возможности АВ и что вообще сейчас может предложить классическая теория кодирования.

На рисунке представлены вероятности ошибки на бит $P_b(e)$ различных алгоритмов декодирования в двоичном канале с АБГШ и ДСК, которые практически всегда являются одним и тем же физическим каналом. По горизонтальной оси на рисунке отложены отношения битовой энергетике канала к спектральной плотности мощности шума E_b/N_0 . Кривая « P_0 channel» указывает для выбранной кодовой скорости $R = 1/2$ вероятности ошибки на бит в канале. Если в демодуляторе применяется, например, жёсткий модем, определяющий только знак принятых двоичных символов, то в декодер поступают значения «0» и «1» демодулированных битов, пришедших в этом случае из ДСК канала. А если в модеме приёмника перед декодером стоит аналого-цифровой преобразователь, то в этот декодер будут поступать, возможно, решения мягкого модема об очередном переданном бите, квантованные, например, на 16 уровней (4 бита). Это и будет двоичный канал с АБГШ, над проблемой эффективного декодирования в котором уже 25 лет особенно напряжённо работают специалисты всего мира с тех пор, как были открыты турбо коды [Berrou et al., 1993].



Характеристики алгоритмов коррекции ошибок в ДСК и канале с АБГШ при $R = 1/2$

Авторы турбо-декодеров дали нам тогда уже вполне обоснованную надежду на то, что приемлемые по сложности декодеры для уровня шума канала, соответствующего непосредственной близости к границе Шеннона, т.е. когда $R \lesssim C$, создать всё-таки можно. Однако они оказались неоправданно сложными и нетехнологичными.

Кроме того, на рисунке представлены вертикальные границы, отмечающие для ДСК и канала с АБГШ с 16 уровнями квантования уровни энергетики, при которых пропускная способность каналов равна $C = 1/2$, помеченные как C_2 и C_{16} соответственно. График «MTD1» показывает наилучшие на данный момент реальные возможности многопорогового алгоритма (МПД), построенного на основе Оптимизационной Теории (ОТ) при использовании ПГЭФ, который при задержке решения не более 10 Мбит и менее $I = 200$ итераций декодирования может быть спроектирован и создан в виде, который обязательно может иметь теоретически максимально возможное быстродействие при аппаратной реализации [Золотарёв, 2017а; Золотарёв, Овечкин, 2017а]. Способы достижения столь большой производительности МПД декодера были запатентованы [Золотарёв, 1972, 2005, 2009, 2015; Золотарёв, Овечкин, 2005]. Этот конкретный декодер работает, как и другие примеры декодеров на данном рисунке, при $R = 1/2$. А совсем простой МПД с 45 итерациями, отмеченный точкой M_s , лучше каскадной схемы, состоящей из короткого свёрточного кода, декодируемого с помощью АВ, и кода РС (кривая «CC:VAK7*RS») при всего лишь впятеро большей задержке решения по сравнению с ней. Развитие мягких МПД, их приближение к границе Шеннона и работа по снижению размеров задержки их решений будут продолжены. Отметим, что каскадная схема АВ-РС фактически является единственным прикладным достижением классической алгебраической теории, да и то лишь в каскадной схеме со свёрточным кодом. Можно ещё раз упомянуть методы Судана для кодов РС, реально послужившие развитию теории, но и они не имели практического значения.

Далее на рисунке приведены графики для свёрточных кодов длины $K = 7, 15$ и 24 (графики «VA:K=7», «VA15» и «VA24»), где K — длина кодирующего регистра, для которых при их декодировании использовался алгоритм Витерби. Как следует из вида этих графиков, самый первый из декодеров АВ с $K = 7$, созданный около 50 лет назад, гораздо слабее последующих. Но и АВ для $K = 15$ также был впервые сделан ещё в прошлом тысячелетии для проекта НАСА «Кассини», причём для небольшой кодовой скорости. Таким образом, АВ для $K = 15$ тоже является вполне реальным устройством. Разумеется, АВ для $K = 24$ пока недоступен для реализации просто из-за экспоненциально растущей с K сложностью декодера. Здесь же показаны характеристики главной каскадной схемы АВ с $K = 7$ и кодом РС [Золотарёв, Овечкин, 2004; Кларк, Кейн, 1987]. Это основные успешные и немногие реальные сейчас методы кодирования для гауссовских каналов. Характеристики конкретных низкоплотностных (LDPC) и турбо декодеров, которые сейчас относятся к реализуемым системам, детально обсуждались в работах [Золотарёв, 2017а; Золотарёв, Овечкин, 2017а; Золотарёв, 2018]. Там же перечислены их основные недостатки. Самым главным свойством этих алгоритмов, которое ограничивает их возможности, является то, что эти методы не измеряют расстояния своих решений до принятого из канала связи вектора. Таким образом, хотя эти декодеры весьма высокой степени сложности и относятся к итеративным процедурам, они не являются оптимизационными процедурами, а это не позволяет отнести их к перспективным методам кодирования. Этот их главный недостаток оказалось возможным

точно сформулировать только сейчас. Но и остальные ограничения, которые им свойственны, не позволяют сделать серьёзные ставки на их успешное развитие [см. статью настоящего сборника *Кузнецов Н.А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99]. И поэтому их реальные возможности уже значительно отстали от достижений МПД алгоритмов, особенно при большом уровне шума.

На рисунке указаны также границы эффективности «seq2» и «seq16» для последовательных алгоритмов, которые так и не смогли преодолеть уровня вычислительных скоростей R_1 гауссовского и ДСК каналов, которые более чем на 2 дБ выше уровня энергетике канала при $R \lesssim C$. Таким образом, последовательные алгоритмы также уже очень давно не участвуют в конкурсе эффективно процедур декодирования для большого уровня шума канала.

Подчеркнём далее в связи с этим, что сейчас только все модификации МПД, а также свёрточные и блочные АВ точно измеряют расстояние своих решений до принятого сообщения. Мы объединяем все эти алгоритмы в группу декодеров прямого контроля метрики (ДПКМ). Скорее всего, только на них и надо ориентироваться при разработке методов, которые позволят быстрее других достичь ещё более близких к границе Шеннона рабочих значений уровня шума при декодировании, чем это уже сделано сейчас. Разумеется, движение рабочей области алгоритмов МПД к этой границе будет продолжено, как это и было всё последнее время. А пока для гауссовских каналов, ДСК, символьных кодов и стирающих каналов наиболее высокие характеристики при очень умеренном уровне сложности обеспечивают только алгоритмы на базе ОТ и МПД, дивергентных принципов, простых приёмов каскадирования и реализации методов ПГЭФ [Золотарёв, 2018; см. статью настоящего сборника *Кузнецов Н.А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99].

Как ещё одно полезное направление развития укажем также на то, что каскадирование АВ при $K \sim 15$ с символьными кодами, как видно из графика для АВ, возможно, могло бы обеспечить для этой схемы хорошие характеристики при $E_b/N_0 \sim 1$ дБ. Этот вопрос стоит детального рассмотрения, чтобы ещё раз оценить возможности, предоставленные теорией и современной элементной базой. Использование с этим АВ двоичных кодов тоже нужно изучать.

Рассмотрим возможности других алгоритмов в канале ДСК. Неконкурентность последовательных процедур, сыгравших в своё время определённую положительную роль в теории кодирования, мы уже отметили выше и увидели по граничным кривым для них на рисунке. А возможности кодов БЧХ оказались крайне слабыми из-за того, что при росте длины кодов отношение кодового расстояния к длине кода d/n для этих кодовых структур быстро падает. И кроме того, в отличие от кодов с мажоритарным декодированием, декодеры БЧХ всегда ошибаются, если число ошибок в принятом блоке превышает $d/2$. Эти два серьёзнейших недостатка вместе со сложностью их декодирования, не доведённой до линейной от длины кода, приводят к тому, что граница для кодов БЧХ по энергетике лежит на рисунке даже дальше, чем для последовательных алгоритмов для ДСК. На это давно известное свойство кодов БЧХ указывалось в работе [Золотарёв, Овечкин, 2004]. А поскольку алгебраическая теория кодирования не смогла решить задачу эффективного декодирования в гауссовских каналах, у неё не осталось никаких перспектив в решении реальных прикладных проблем, т. е. в создании эффективных декодеров для работы вблизи границы Шеннона [Золотарёв, 2018].

Далее точкой M_a на рисунке помечен результат 30-летней давности для крайне простого МПД декодирования при $E_b/N_0 = 4$ дБ для свёрточного кода с $d = 11$ и $I = 14$ итерациями декодирования [Золотарёв, 1986]. Тем самым наглядно видно, что уже в те годы были полностью перекрыты все возможные достижения для кодов БЧХ и всех прочих методов алгебраической теории. Это однозначно определялось достижением МПД алгоритмом и в этом случае весьма малой энергетики канала типа ДСК уровня оптимального (т.е. переборного!) декодирования при всего лишь $I = 14$ итерациях декодирования для довольно хорошего в плане размножения ошибок кода. Крайне странно, что наши теоретики совершенно не оценили тогда этот результат, который ясно указывал на завершение «классического» этапа теории кодирования.

При снижении энергетики ДСК при $R = 1/2$ до ~ 3 дБ достижение высокоуровня достоверности декодирования тоже не составляет проблемы, если значительный рост задержки решения в блоковом или свёрточном вариантах декодирования до ~ 1 Мбит признаётся допустимым. Эта достижимая для многих видов МПД декодеров и кодов точка отмечена рисунке как M_h . При отличии рабочей энергетики МПД от шенноновского уровня для $R = 1/2$ в ДСК на ~ 1 дБ сложность и задержка МПД будут, как следует из стандартных оценок и экспериментальных данных, примерно в 1,5 раза меньше, чем у довольно не простого уже МТД1, что вполне естественно для столь большого уровня шума. Значит, и в ДСК у ОТ, МПД и методов ПГЭФ конкурентов вообще нет.

Таким образом, из комплексного обзора сведений об алгоритмах декодирования получается, что в гауссовских каналах пока что нет реально других перспективных алгоритмов, кроме МПД и АВ. Мы также обсудили, почему турбо и LDPC коды выпали из конкурса перспективных алгоритмов. История последних 15 лет изучения декодеров этого типа, когда рост их эффективности прекратился, однозначно подтверждает наш вывод о слабых характеристиках всех алгоритмов, не относящихся к группе прямого контроля метрики ДПКМ. В самом деле, достаточно сложно ожидать от алгоритма, который не измеряет расстояние своих решений до принятого вектора, успешного хоть в какой-то мере декодирования, тем более при максимально допустимом теорией уровне шума. В этих условиях методы для LDPC кодов могут вообще «не увидеть» достижение решения ОД, так как требуемые для этого замеры расстояния декодеры LDPC вообще не производят.

Столь же отрицательное и аргументированное отношение нашей научной школы к полярным кодам было нами изложено в источниках [Золотарёв, 2017а, 2018; Золотарёв, Овечкин, 2016; www.mtdbest.ru] с использованием необходимых в этих случаях ссылок. Все классы этих и многих других методов кодирования последнее время безо всяких причин объявляются эффективными и перспективными при вообще полном отсутствии за прошедшие годы хотя бы одного-трёх примеров достижения каких-то разумных и проверяемых конкретных характеристик эффективности и сложности. Мы полагаем, что полная текущая профанация теории и прикладных идей всей теории помехоустойчивого кодирования требует крайних мер в оценках характеристик публикуемых алгоритмов. Например, полезно всегда публиковать только те алгоритмы декодирования для классических моделей каналов связи, которые следует обязательно сопровождать заслуживающими доверия характеристиками эффективности и быстродействия, например, на языке C++. Это введёт в нормальное русло обучение студентов и специалистов, а также повысит уровень диссертаций по прикладным вопросам теории кодирования. Сделать это крайне важно, поскольку сейчас

идёт просто настоящий поток монографий и защит столь «великих» докторских диссертаций, в которых нет вообще никаких графиков характеристик эффективности и производительности «новых» алгоритмов, а также каких-либо численных оценок их возможностей. Эти «опусы» переполнены десятками фрагментов псевдопрограмм, цена которых в миллионы раз меньше стоимости бумаги, на которой они напечатаны. Теория кодирования может быстро вернуть себе престиж самой важной отрасли теории информации, если российские специалисты, наконец, обратятся к разработкам, технологиям и прикладным исследованиям на базе современных теорий глобальной оптимизации функционалов [Золотарёв, 2017а, 2018; Золотарёв, Овечкин, 2004, 2017а; см. статью настоящего сборника *Кузнецов Н.А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99].

Отметим кратко другие эффективные алгоритмы декодирования. В публикациях [Золотарёв, 2017а; Золотарёв, Овечкин, 2017а] мы неоднократно отмечали, что символьные коды навели, наконец, в случайных недвоичных каналах полный порядок после 50-летнего застоя, случившегося после важнейшего для теории кодирования шестидесятих годов открытия кодов Рида – Соломона (РС). Но коды РС реальны только в своих коротких версиях. Использовать длинные коды РС нельзя, да и не надо, т.к. они тоже малоэффективны, а их декодеры неоправданно сложны. Ну, а больше ничего и нет. А символьные коды могут быть любой длины и даже при большом уровне шума декодируются оптимально, как это сделал бы недвоичный АВ, создать который для каких-либо не очень коротких недвоичных кодов, скорее всего, совершенно нереально. Уровень шума, при котором символьный МПД (QMПД) декодирует символьный код оптимально, в разы больше по вероятности ошибки канала, чем это могут позволить себе декодеры для кодов РС. Обширные данные по символьным МПД широко опубликован уже очень давно [Золотарёв, 2018; Золотарёв, Назиров, 2008; Золотарёв, Овечкин, 2004; Золотарёв и др., 2010; см. статью настоящего сборника *Кузнецов Н.А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99; www.mtdbest.ru]. Важно при этом и то, что для недвоичных кодов не существует сколько-нибудь эффективных алгоритмов Витерби, так как их сложность при больших q слишком велика [Золотарёв, 2017а; Золотарёв, Овечкин, 2017а]. Таким образом, уже созданные системы с кодами РС должны работать, пока они нужны, а все новые задачи для недвоичных каналов следует решать с использованием исключительно символьных кодов, т. к. конкурентов у них теперь уже просто вообще не может быть.

В качестве примера высокой эффективности QМПД при очень малой избыточности, когда $R = 19/20$, читатели могут переписать на свой компьютер демо-программу для МПД (QMПД) блочного символьного кода по гиперссылке [http://www.mtdbest.ru/program/qmtd_demo_r.zip] со страницы «Обучение» сайта www.mtdbest.ru. Эта демо-программа, как и все другие программные средства на наших порталах, сопровождается инструкцией, которая позволяет настроить ряд параметров декодера. Очень важно, что на самых обычных ПК производительность QМПД, как и алгоритмов для двоичных кодов достигает десятков мегабит в секунду, что отлично иллюстрирует огромное преимущество методов, созданных на базе ОТ, над другими методами декодирования.

Отметим, что причина высокой скорости работы декодеров этого типа одна: единственный активный блок МПД — пороговый элемент, простейшее устройство. В случае аппаратной реализации в соответствии с теми же уже

запатентованными решениями для ДСК и АБГШ каналов [Золотарёв, 1972, 2005, 2009, 2015; Золотарёв, Овечкин, 2005] символьный декодер также может обеспечить теоретически максимально возможную производительность для любой элементной базы, которая будет просто совпадать со скоростью продвижения данных по регистрам сдвига выбранной микросхемы или ПЛИС.

Таким образом, при наличии до недавнего времени работоспособной системы кодирования на базе только недвоичных коротких кодов РС (а в силу этого и малоэффективных) алгоритмы QМПД оказываются сейчас единственными реальными высокопроизводительными декодерами для недвоичных каналов со случайными ошибками. Они характеризуются высочайшей производительностью и способностью находить решения ОД даже в условиях очень большого уровня шума. Это означает, что ОТ полностью решила вопросы простого и высокодостоверного декодирования в недвоичных каналах и на сегодняшний день, и на перспективу.

Наконец, очень кратко отметим, что и в стирающих каналах запатентованные алгоритмы, идеологически относящиеся к МПД методам, на больших скоростях декодируют данные при $R = 1/2$ и при вероятности приёма из канала стёртых символов $\sim 0,49$ снижают долю невосстановленных алгоритмом символов до уровня менее 10^{-6} [Золотарёв, 2018; Золотарёв, 2017б; см. статью настоящего сборника Кузнецов Н. А. и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99]. А недавно запатентованный нами АВ [Золотарёв, Овечкин, 2017б] для блоковых кодов (БАВ) со сложностью $\sim 2^K$, а не 2^{2^K} , как умели буквально до недавнего времени наши теоретики, полностью исключает из конкурсов для гауссовских каналов вообще все алгебраические алгоритмы.

Таким образом, представленные результаты для всех основных типов каналов чётко свидетельствуют о безусловном и очень большом по всем параметрам эффективности и сложности преимуществе технологий и идеологии ОТ при решении задач теории кодирования, относящихся к исправлению, восстановлению, контролю и хранению цифровых данных в современных цифровых системах. ОТ успешно приняла эстафету во всех прикладных вопросах от классической алгебраической теории кодирования и выходит в новое бескрайнее интеллектуальное пространство оптимизационных алгоритмов, с линейной от длины кодов сложностью решающих все проблемы достижения оптимальной по максимуму правдоподобия достоверности цифрового контента нашей информационной цивилизации.

Мы полагаем, что число наших последователей и сторонников методов ПГЭФ, как и раньше, будет расти, а новые сферы исследований и разработок с использованием парадигм нашей новой «квантовой теории» в области помехоустойчивого кодирования, т. е. технологий ОТ и МПД, новых видов АВ-декодирования, дивергентных и каскадных схем будет быстро расширяться [Золотарёв, 2017а, 2018; Золотарёв, Овечкин, 2017а; Zolotarev et al., 2015; см. статью настоящего сборника Кузнецов Н. А. и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99].

Большинство ссылок, относящихся к этой статье, как и многие другие материалы по ОТ и МПД алгоритмам, в том числе множество демо-программ, удобных для предварительного знакомства и исследований МПД, можно найти на наших сетевых порталах [www.mtdbest.ru; www.mtdbest.iki.rssi.ru] и в нашей новой монографии [Золотарёв, 2018]. Большое число ссылок на конкретные высокоскоростные демо-программы и исследовательские платформы по ОТ

и МПД также представлены в источниках [www.mtdbest.ru, см. статью настоящего сборника *Кузнецов Н. А.* и др. Многопороговые алгоритмы на базе оптимизационной теории вблизи границы Шеннона, с. 99].

ЛИТЕРАТУРА

- [Витерби, 1970] *Витерби А. Дж.* Границы ошибок для свёрточных кодов и асимптотически оптимальный алгоритм декодирования // Некоторые вопросы теории кодирования. М.: Мир, 1970. С. 142–165.
- [Золотарёв, 1972] *Золотарёв В. В.* Устройство для декодирования линейных свёрточных кодов. А. С. СССР № 492878 от 1972.
- [Золотарёв, 1986] *Золотарёв В. В.* Многопороговое декодирование // Проблемы передачи информации. 1986. Т. 22. Вып. 1. С. 104–109.
- [Золотарёв, 2005] *Золотарёв В. В.* Высокоскоростное устройство многопорогового декодирования линейных кодов: патент на полезную модель. № 44216 от 27.02.2005.
- [Золотарёв, 2009] *Золотарёв В. В.* Способ декодирования помехоустойчивого кода: патент на изобретение. РФ № 2377722 от 27.12.2009.
- [Золотарёв, 2015] *Золотарёв В. В.* Способ декодирования помехоустойчивого кода: патент на изобретение. РФ № 2557454 от 25.06.2015.
- [Золотарёв, 2017а] *Золотарёв В. В.* О новом этапе развития оптимизационной теории // Цифровая обработка сигналов. 2017. № 1. С. 33–41.
- [Золотарёв, 2017б] *Золотарёв В. В.* Способ обнаружения и исправления стираний при приёме дискретной информации: патент на изобретение. РФ № 2611235 от 21.02.2017.
- [Золотарёв, 2018] *Золотарёв В. В.* Теория кодирования как задача поиска глобального экстремума // под науч. ред. акад. РАН Н. А. Кузнецова. 2-е изд., испр. М.: Горячая линия – Телеком, 2018, 228 с.
- [Золотарёв, Назиров, 2008] *Золотарёв В. В., Назиров Р. Р.* Сверхнадёжное исправление ошибок на основе МПД алгоритмов для баз данных систем ДЗЗ // Современные проблемы дистанционного зондирования Земли из космоса. 2008. Вып. 5. Т. 1. С. 267–272.
- [Золотарёв, Овечкин, 2004] *Золотарёв В. В., Овечкин Г. В.* Помехоустойчивое кодирование. Методы и алгоритмы: Справочник. М.: Горячая линия – Телеком, 2004. 126 с.
- [Золотарёв, Овечкин, 2005] *Золотарёв В. В., Овечкин Г. В.* Устройство многопорогового декодирования линейных кодов для гауссовских каналов: патент на полезную модель. № 44215 от 27.02.2005.
- [Золотарёв, Овечкин, 2016] *Золотарёв В. В., Овечкин Г. В.* О сопоставлении новых методов помехоустойчивого кодирования // Доклады 18-й Международной конф. «Цифровая обработка сигналов и её применение». Москва. 2016. Т. 1. С. 59–64.
- [Золотарёв, Овечкин, 2017а] *Золотарёв В. В., Овечкин Г. В.* Эффективные многопороговые методы декодирования самоортогональных кодов // Вестник РГРТУ. 2017. № 60. С. 113–122.
- [Золотарёв, Овечкин, 2017б] *Золотарёв В. В., Овечкин П. В.* Способ кодирования и декодирования блочного кода с использованием алгоритма Витерби: патент на изобретение. РФ № 2608872 от 25.01.2017.
- [Золотарёв и др., 2010] *Золотарёв В. В., Овечкин Г. В., Назиров Р. Р., Овечкин П. В., Чулков И. В.* Эффективное недвоичное многопороговое декодирование помехоустойчивых кодов для систем дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 2. С. 269–274.
- [Зубарев, Овечкин, 2008] *Зубарев Ю. Б., Овечкин Г. В.* Помехоустойчивое кодирование в цифровых системах передачи данных // Электросвязь. 2008. № 12. С. 2–11.

- [Кларк, Кейн, 1987] *Кларк Дж., Кейн Дж.* Кодирование с исправлением ошибок в системах цифровой связи. М.: Радио и связь, 1987.
- [Berrou et al., 1993] *Berrou C., Glavieux A., Thitimajshima P.* Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes // Proc. Intern. Conf. Commun. Geneva, Switzerland. 1993. P. 1064–1070.
- [Forney, 1974] *Forney G. D.* Convolutional codes. II. Maximum-likelihood decoding // Information and control. 1974. V. 25. No. 3.
- [Heller, Jacobs, 1971] *Heller J. A., Jacobs J. M.* Viterbi decoding for satellite and space communication // IEEE Trans. Comm. Technology. Pt. 2. 1971. V. COM-19. No. 5.
- [Zolotarev et al., 2015] *Zolotarev V. V., Zubarev Y. B., Ovechkin G. V.* Optimization Coding Theory and Multithreshold Algorithms // ITU Publications. 2015. URL: <http://www.itu.int/pub/S-GEN-OCTMA-2015>.

OPTIMIZATION THEORY: THE RECEPTION OF THE BATON OF LEADERSHIP FROM THE APPLIED CLASSIC THEORY OF ERROR-CORRECTING CODING

*V. V. Zolotarev*¹, *G. V. Ovechkin*², *R. R. Nazirov*¹

¹ Space Research Institute of Russian Academy of Sciences (IKI RAN)

² Ryazan State Radio Engineering Institute (RSREU)

We discuss the main implementation methods of multithreshold decoding (MTD) algorithms, considered as the procedure of searching the global extremum of functional with minimal, i.e. linear complexity with the code length growth. Due to results of the Optimization Theory (OT) possibilities comparison with the other methods, it follows that the methods of OT and MTD, together with the options of block Viterbi algorithm (BVA) patented by our scientific school and new paradigms of the OT development of noise-resistant coding, completely replace other methods of decoding by the criteria of simplicity implementation, proximity to the Shannon bound and desired veracity.

Keywords: error-correction coding, optimization theory, block Viterbi algorithm, self-orthogonal codes, symbolic codes, concatenated codes, multithreshold decoders, the code gain, the communication channel, the Shannon bound

Zolotarev Valery Vladimirovich — leading researcher, professor, doctor of technical sciences, zolotasd@yandex.ru

Ovechkin Gennady Vladimirovich — professor doctor of technical sciences, g_ovechkin@mail.ru

Nazirov Ravil Ravilyevich — head of department, doctor of technical sciences

ПРОГРАММНЫЙ КОМПЛЕКС «ГЕОДИАЛОГ»: МОДЕЛЬ ЭКОЛОГИЧЕСКОГО УПРАВЛЕНИЯ

Б. М. Балтер¹, Д. Б. Балтер¹, В. В. Егоров¹,

М. В. Стальная¹, М. В. Фаминская²

¹ Институт космических исследований Российской академии наук (ИКИ РАН)

² Российский государственный социальный университет (РГСУ)

Приводится описание программного комплекса (ПК) «Геодиалог», предназначенного для моделирования процесса мониторинга и управления экологическим состоянием экосистем в замкнутом контуре автоматического регулирования. Изложена идея информационного цикла, имеющего место в этом контуре, описаны функции и организация ПК «Геодиалог», а также основные области его практического применения.

Ключевые слова: мониторинг, управление, информация, энтропия, цикл, экосистема, программный комплекс, динамика, состояние

ВВЕДЕНИЕ

Проблемы экологии приобретают всё более актуальный характер и становятся главной заботой мирового сообщества [Балтер, Егоров, 1994; Балтер, Фаминская, 2014; Балтер и др., 1994, 2007]. Очевидно, что мониторинг состояния окружающей среды с адекватным временным и пространственным охватом возможен только с помощью аэрокосмических систем дистанционного зондирования. В настоящее время такие системы достигли соответствующего уровня разработки. Однако их потенциал используется недостаточно эффективно. Это обусловлено несколькими причинами.

Во-первых, выходная информация таких систем не вполне соответствует практическим задачам управления состоянием окружающей среды. Поэтому, она теряет часть своей значимости или, говоря языком термодинамики, просто рассеивается. Существует также обратная сторона этой проблемы: системы аэрокосмического мониторинга не ориентированы на то, чтобы принимать во внимание эффекты экологического управления, основанные на получаемой ими информации, т. е. обладают недостатком, связанным с отсутствием канала обратной связи необходимого для оптимизации использования собираемой информации. Следовательно, сенсоры мониторинга и области практического применения этих данных функционально разъединены и не образуют полной системы мониторинга и управления с обратной связью.

Балтер Борис Михайлович — научный сотрудник, balter@mail.ru

Балтер Дмитрий Борисович — инженер-программист 2-й категории, db.balter@gmail.com

Егоров Виктор Валентинович — старший научный сотрудник, кандидат технических наук, victor_egorov@mail.ru

Стальная Мария Валентиновна — младший научный сотрудник, stalnayam@mail.ru

Фаминская Марина Вадимовна — доцент, кандидат физико-математических наук, faminskaya@mail.ru

Во-вторых, процесс разработки сенсоров дистанционного зондирования обычно не сопровождается моделированием возможного практического применения получаемых ими данных. Поэтому параметры сенсоров: бортовых радаров, радиометров, сканеров, спектрометров, выбираются в значительной степени на базе субъективных решений. В результате информация, получаемая такими сенсорами, оказывается часто бесполезной для практики. Оптимизация систем аэрокосмического мониторинга и более эффективное использование получаемых ими данных для целей экологического мониторинга вызывают необходимость применения новых методов информационного дизайна и анализа. Это явилось причиной разработки программного комплекса (ПК) «Геодиалог».

ГЛАВНАЯ ИДЕЯ ПК «ГЕОДИАЛОГ»

Временная эволюция экосистемы, дистанционный мониторинг, целевое управление состоянием окружающей среды, построение моделей и их подгонка — все эти аспекты могут быть хорошо описаны в теории, но они завязаны в один сложный узел в почти любой практической задаче. Это выглядело бы многообещающим, будучи внедрённым в единый пакет знаний математических экологов, специалистов в области оптимального управления, дистанционного зондирования, создателей сложных моделей, программистов и обработчиков информации. Пользователи могли бы успешно использовать эту базу знаний, даже являясь специалистом хотя бы только в одной из этих областей, поскольку всё это увязано в программную структуру. Она в «Геодиалоге» отражается на экране, на котором показаны средства управления ею, что открывает новые возможности для пользователя.

Главная возможность — это создание замкнутой системы управления, включающей в себя подсистемы мониторинга и управления с обратной связью (рис. 1).

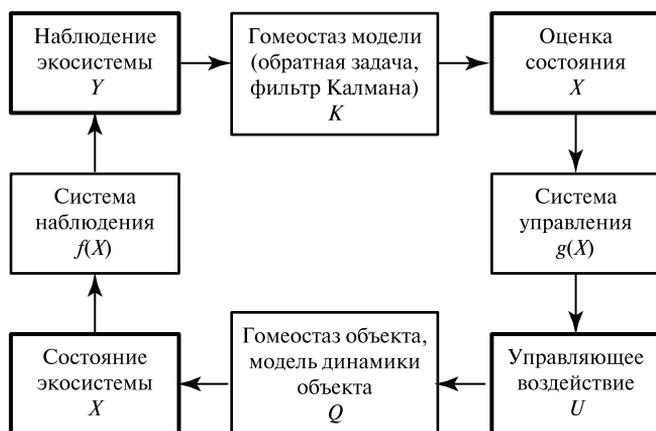


Рис. 1. Информационный цикл мониторинга и управления; K — коэффициент передачи фильтра Калмана; Q — аналог коэффициента передачи экосистемы

Обе подсистемы удовлетворяют теориям оптимальности: одна из них предназначена для фильтрации принимаемого сигнала, а другая для управления состоянием экосистемы и приведения её к заданному целевому состоянию. Обе подсистемы имеют дело с информацией и критерий их оптимальности может быть сформулирован посредством связанной с информацией понятием энтропии. Управление основано на информации, поставляемой мониторингом, а с другой стороны, посредством управления могут открываться новые возможности для мониторинга. Таким образом, формируется замкнутый информационный цикл. Возникает вопрос применительно к мониторингу и управлению: не можем ли мы, используя возможности одного из них, повышать эффективность другого? В таком случае, естественно использовать энтропию полного информационного цикла вместо двух парциальных энтропий в качестве критерия оптимизации. При этом невозможно приводить состояние экосистемы к цели (оптимуму) любыми способами. Это возможно только, если сложность системы, измеряемая негэнтропией, при этом не убывала, так как это привело бы к снижению оптимизируемого критерия качества управления. Этот критерий адаптирует наши усилия по управлению к естественному гомеостазу системы. Поэтому далее становится неочевидным кто кем управляет, и кто за кем наблюдает: технологическая подсистема за природной или наоборот. Обе выполняют одни и те же функции, как в диалоге. Отсюда и возникло название программного комплекса — «Геодиалог».

Другая возможность открывается в результате создания динамических сценариев мониторинга, которые адаптируются к изменению состояния экосистемы таким образом, чтобы оптимизировать информационную энтропию. Это открывает путь к проектированию адаптивных космических сенсоров. «Геодиалог» предназначен управлять этой комплексной проблемой оптимизации таким же способом, как он управляет мониторингом и динамикой состояния экосистемы по отдельности. Такие «смешанные сценарии» могут использоваться также для подгонки моделей в ситуациях, в которых происходит дрейф состояния экосистемы и изменение условий мониторинга. Эти ситуации реально имеют место на практике. Тогда, как и в предыдущем случае, критерий оптимальности должен зависеть от параметров, описывающих как состояние экосистемы, так и условия мониторинга.

Мы надеемся, что «Геодиалог» будет способствовать выполнению симулированных экспериментов в области мониторинга и управления, которые ранее были технически невозможны.

ФУНКЦИИ КОМПЛЕКСА «ГЕОДИАЛОГ»

«Геодиалог» — многоцелевой модульный программный комплекс, предназначенный для моделирования динамики экосистем и других комплексных двумерных структур типа изображений или одномерных типа векторов, процесса их мониторинга средствами дистанционного зондирования, управления их параметрами пользователем или средствами автоматического управления. Комплекс моделирует как динамику экосистемы, так и управление ею:

- подгоняет модель нелинейного дифференциального уравнения динамики к данным;
- приводит состояние экосистемы к желаемой цели;

- воспроизводит отклик, наблюдений, получаемых в ходе мониторинга, к изменению состояния экосистемы;
- строит карты состояния экосистемы по изображениям, полученным средствами дистанционного зондирования, и информации о функции отклика мониторинга;
- оптимизирует информативность процесса мониторинга;
- использует принцип обратной связи для поддержания симулированной экосистемы в гомеостазе в соответствии с выбранным критерием качества управления;
- выполняет все эти функции, объединённые в информационный цикл мониторинга/управления;
- посредством одновременного использования определённой пользователем цели и поддержания гомеостаза экосистемы как критерия, реализует новый диалоговый режим взаимодействия между технологической системой «наблюдатель/контроллер» и экосистемой.

ОРГАНИЗАЦИЯ ПРОГРАММНОГО КОМПЛЕКСА

«Геодиалог» имеет четыре программных блока (рис. 2). Каждое симулированное изображение соответствует одному из прямоугольников, выделенных жирными линиями на рис. 1. Они выполняют функции (по часовой стрелке, начиная с левого нижнего угла):

- определения состояния экосистемы и её динамики;
- отображения спектрального отклика;
- обработки сигнала;
- выбора сценария управления.

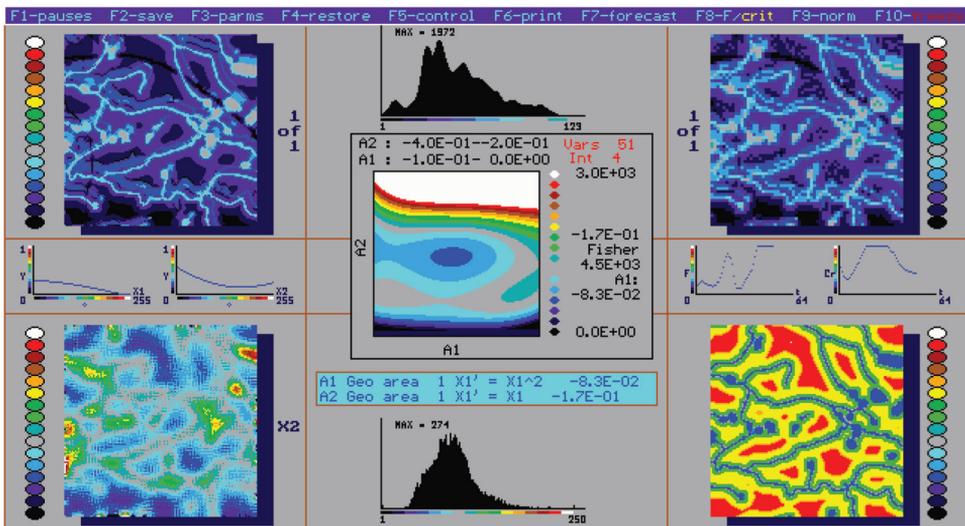


Рис. 2. Главный экран комплекса «Геодиалог», демонстрирующий цикл мониторинга-управления

Каждый блок может быть открыт, после чего пользователь может определить соответствующие параметры и учесть их влияние на «передаточную функцию» этого блока без учёта воздействия других блоков. Таким образом, например, чтобы установить сценарий динамики экосистемы, пользователю нет необходимости прибегать к использованию блоков мониторинга и управления.

В этом отдельном блоковом режиме можно загрузить из базы данных параметры состояния экосистемы и сенсора, функции связи параметров состояния со спектром, сценарий управления и др. В том же режиме сохраняются новые назначаемые параметры. После этого каждый блок поставляет своё обновлённое содержание в общий «форум», предусмотренный для их взаимодействия и отображаемый на экране информационного цикла. В нём каждый блок представлен окном изображения в одном из четырёх углов экрана (см. рис. 2).

Изображение состояния экосистемы (внизу слева) показывает величины в цветовом коде: чёрный – фиолетовый для малых значений, красный – белый для больших значений. Отдельно отображаются поля переменных состояния X_1 и X_2 . Переменные X_1 и X_2 взаимодействуют друг с другом в динамических уравнениях. Пользователь может иметь несколько объектов на этих изображениях, и параметры динамики переменных X_1 и X_2 имеют различные значения для каждого объекта. Гистограммы значений переменных X_1 и X_2 показаны в прямоугольнике внизу экрана. Они используются в алгоритмах управления.

Окно с изображением отражательной способности объектов экосистемы (вверху слева) показывает, как выглядит экосистема в выбранном узком спектральном канале при её освещении белым светом единичной интенсивности. При этом яркость в этом канале можно отобразить оттенками серого, а на экране к ним применяется цветовое кодирование, как и к блоку состояния экосистемы. Обычно задаётся несколько спектральных каналов сенсора, параметры которых выбираются в ходе моделирования цикла мониторинга и управления. Карта значений отражательной способности вычисляется по карте значений переменных состояния с помощью функции связи состояние – спектр (слева, посередине). Последняя устанавливается заранее. Она различна для разных объектов.

Блок сенсора использует карту отражательной способности, добавляет условия освещения, затем пропускает сигнал через детектор и другие элементы модели сенсора. Доступны пять типов сенсоров: многоспектральный сканер, спектрометр, локаторы бокового обзора с реальной и синтезированной апертурой, микроволновый и инфракрасный радиометры.

Результирующее изображение показано вверху справа. Это то, что видно на изображении, полученном в соответствующем спектральном канале сканера, на экране радара и т.п. Гистограммы выходного сигнала каждого канала показаны вверху посередине. Они используются в алгоритмах фильтрации. Выходом процесса фильтрации сигнала сенсора являются оценки реального состояния экосистемы. Карты оценённого состояния показываются также в виде цветокодированного изображения слева вверху.

Эти карты поступают на блок управления. Он производит оценки будущего состояния экосистемы и вычисляет сигнал оптимального управления для приведения экосистемы к целевому состоянию, заданному соответствующей картой (показана внизу справа). В зависимости от выбранного режима, это изображение демонстрирует карту предсказываемых состояний, карту целевого состояния, карту управляющего воздействия или карту информации Фишера, содержащейся в выходном сигнале сенсора относительно средних значений

переменных состояния для каждого объекта. Вклады каждой переменной и каждого объекта в общий критерий оптимальности учитываются с весами, определяемыми пользователем. Обходом одного цикла завершается модельный шаг и устанавливается пауза, когда пользователь может менять установки в стратегии управления. После этого блок экосистемы принимает новые значения сигнала управления и цикл выполняет следующий шаг. Производится отображение истории значений критерия, полученных на каждом шаге (справа, посередине), чтобы отобразить общую тенденцию.

В паузе пользователь имеет доступ к меню рабочего цикла, отображаемого в верхнем ряду. Оно предоставляет несколько технических возможностей, таких как сохранение/восстановление результатов, изменение нормализаторов, которые определяют клин цветовых кодов и т.п. Два из них являются идеологически важными: установка/изменение контролируемых параметров и установка критерия оптимальности.

Несколько параметров каждого из четырёх основных блоков (из нескольких десятков, в зависимости от числа объектов или каналов сенсора) могут выбираться в качестве оптимизируемых в процессе управления. Пользователь может использовать только два оптимизируемых параметра на каждой итерации цикла, поскольку выполнять процедуру оптимизации одновременно с большим числом параметров нерационально. Эти параметры могут изменяться пользователем или оптимизироваться автоматически. Остальные параметры, не устанавливаемые как управляющие, могут меняться только вручную. Например, чтобы оптимизировать размещение спектральных каналов, пользователь может выбрать центры двух из них в качестве управляемых и одновременно установить их ширину вручную. Для оптимизации состояния экосистемы можно ввести два коэффициента состояния в дифференциальные уравнения динамики экосистемы в качестве двух управляемых параметров и изменять значения остальных коэффициентов вручную.

Блок оптимизации функционально является пятым основным блоком. Он имеет своё собственное окно, высвечиваемое по запросу в центре экрана. Однако алгоритмически он является частью блока фильтрации/управления, и делает в интерактивном режиме то, что блок фильтрации/управления делает автоматически. Центральное окно показывает значения информации Фишера или отклонение от целевой функции в виде цветокодированного критерия по двум управляющим параметрам. Окно показывает значения критерия, полученное по всем итерациям цикла, после того, как последняя назначенная процедура управления выполнена. Пользователь может вручную выбрать следующую точку (т.е. управляющие параметры), чтобы получить ожидаемое улучшение значения критерия. Блок оптимизации производит также интерполяцию значений критерия для всех точек или для их подмножества и предлагает перспективные значения параметров управления в виде красных или белых зон.

ОБЛАСТИ ПРИМЕНЕНИЯ ПК «ГЕОДИАЛОГ»

Существует несколько перспективных областей использования ПК «Геодиалог».

Дистанционное зондирование. Комплекс «Геодиалог» представляет собой средство тестирования существующих и поиска новых конструкций сенсоров дистанционного зондирования, процедур обработки информации, режимов наблюдения и т.п. Он выполняет эти процедуры с чётко определённым и вы-

числяемым критерием в виде величины информационного потока (только проблемно-ориентированной информации) и/или показателя управляемости экосистемы. В этой области «Геодиалог» — автоматизированное средство проектирования конструкции сенсоров и выбора параметров космических систем.

В области фундаментальной науки. Комплекс «Геодиалог» может применяться в следующих областях:

- для экологов: изучать математические модели популяционной динамики, гомеостаза экосистем, критических явлений, самоорганизации, вызванной диффузионными процессами, исследовать перспективы управления, оценивать точность, масштаб и периодичность процедур мониторинга;
- для географов: исследовать модели, воспроизводящие природные пространственные и временные структуры;
- для специалистов оптимального управления: тестировать алгоритмы управления для нелинейных технических систем, имеющих прикладное значение и чёткий критерий качества;
- для специалистов в области нелинейной математической физики: визуализировать сложные феноменологические процессы, имеющие чёткое математическое описание, и иметь возможность управлять ими.

Управление природными ресурсами. Комплекс «Геодиалог» может быть инструментом экспертизы крупномасштабных проектов, имеющих экологическую направленность. Это предполагает использование нового подхода к взаимодействию с экосистемами, который допускает воздействие на них только на основе данных мониторинга и такое, которое не уменьшает критерия общего информационного потока. Это обеспечивает формирование канала обратной связи для экосистемы для предотвращения любого действия, приводящего к нарушению её гомеостаза. Так построенная имитационная модель сценария ресурсного управления, удовлетворяющего этим ограничениям, может быть применена на практике с достаточной вероятностью не причинить необратимый вред экосистеме.

В области образования. Комплекс «Геодиалог» может быть демонстрационным экспериментальным комплексом для студентов, обучающихся в области математической экологии, дистанционного зондирования, оптимального управления. Он даёт обширные возможности визуализации и позволяет пользователю изменять практически любой параметр с непосредственным отображением его влияния на критерий качества.

ЗАКЛЮЧЕНИЕ

Программный комплекс «Геодиалог» представляет собой модель замкнутого контура мониторинга и управления окружающей средой и экологическими процессами, протекающими в экосистемах.

Актуальность разработки такого комплекса была вызвана двумя причинами: 1) выходная информация глобальных систем мониторинга не вполне соответствует практическим задачам управления состоянием окружающей среды;

2) процесс разработки сенсоров дистанционного зондирования обычно не сопровождается моделированием возможного практического применения получаемых ими данных.

Комплекс «Геодиалог» — это новый подход к решению научных и прикладных задач в таких областях, как экология, география, в управлении природными ресурсами и образовании.

ЛИТЕРАТУРА

- [Балтер, Егоров, 1994] Балтер Б. М., Егоров В. В. Природные и информационные циклы в экологическом моделировании и дистанционном зондировании. Концепция и принципы выбора объектов мониторинга // Исслед. Земли из космоса. 1994. № 6. С. 19–25.
- [Балтер, Фаминская, 2014] Балтер Б. М., Фаминская М. В. Природно-технологические и социально-информационные циклы в ноосферных подходах В. И. Вернадского, Н. Н. Моисеева и П. Тейяра де Шардена // Международный науч. ин-т “Educatio”. 2014. № 5. С. 152–154.
- [Балтер и др., 1994] Балтер Б. М., Егоров В. В., Каурова И. Э., Козин М. Б. Изоморфизм природных циклов мониторинга: применение в компьютерном моделировании // Исслед. Земли из космоса. 1994. № 1. С. 29–35.
- [Балтер и др., 2007] Балтер Б. М., Балтер Д. Б., Егоров В. В., Стальная М. В. Программный комплекс «Геодиалог» для моделирования съемки и обработки данных аэрокосмических РСА // Современные проблемы дистанционного зондирования Земли из космоса. 2007. Т. 4. № 1. С. 13–19.

“GEODIALOG”: SOFTWARE PACKAGE FOR MODELING ENVIRONMENTAL CONTROL

B. M. Balter¹, D. B. Balter¹, V. V. Egorov¹, M. V. Stalnaya¹, M. V. Faminskaya²

¹ Space Research Institute of Russian Academy of Sciences

² Russian State Social University, Moscow

We describe the software package “Geodialog”, which models the monitoring and control of the state of environment and ecological systems. The model is based on a closed cycle of the type considered in the theory of automatic control. We describe the basic features of the information flow in this cycle, the architecture of the software and the application domains envisioned for it.

Keywords: monitoring, control, information, entropy, cycle, ecosystem, software package, dynamics, state

Balter Boris Mikhaylovich — researcher, balter@mail.ru

Balter Dmitriy Borisovich — engineer-programmer 2nd cat, db.balter@gmail.com

Egorov Victor Valentinovich — senior researcher, PhD, victor_egorov@mail.ru

Stalnaya Maria Valentinovna — junior researcher, stalnayam@mail.ru

Faminskaya Marina Vadinovna — associate professor, PhD, faminskaya@mail.ru

МНОГОПороГОВЫЕ АлГОРИТМЫ НА БАЗЕ ОПТИМИЗАЦИОННОЙ ТЕОРИИ ВБЛИЗИ ГРАНИЦЫ ШЕННОНА

Н. А. Кузнецов¹, В. В. Золотарёв², Г. В. Овечкин³,

Р. Р. Назиров², С. В. Аверин⁴

¹ Институт радиотехники и электроники им. В. А. Котельникова
Российской академии наук (ИРЭ им. В. А. Котельникова РАН)

² Институт космических исследований Российской академии наук
(ИКИ РАН)

³ Рязанский государственный радиотехнический университет (РГРТУ)

⁴ ООО «Объединённые радиоэлектронные технологии»

Анализируются характеристики алгоритмов декодирования, построенных на основе оптимизационной теории (ОТ) помехоустойчивого кодирования. Обсуждаются многопороговые декодеры для двоичных гауссовских, недвоичных, стирающих и других каналов. Указано на большие возможности блочного алгоритма Витерби. Подчёркиваются перспективы дивергентного кодирования. Отмечается группа алгоритмов прямого контроля метрики. Обсуждается ценность основных парадигм ОТ. Определяются основные пути развития теории кодирования на ближайшие годы и на перспективу.

Ключевые слова: телекоммуникационные системы, системы хранения данных, помехоустойчивое кодирование, блочный алгоритм Витерби, компьютерное моделирование, самоортогональные коды, символьные коды, каскадные коды, многопороговые декодеры, энергетический выигрыш кодирования, дивергентное кодирование

ВВЕДЕНИЕ

В 2015 г. исполнилось 25 лет со дня защиты диссертации [Золотарёв, 1990], где для очень простых по сегодняшним меркам кодов были доказаны многие основные результаты, которые позднее были систематизированы и представлены в полном объёме в оптимизационной теории помехоустойчивого кодирования [Золотарёв, 2005, 2006а, б, 2015а; Золотарёв, Овечкин, 2005, 2014, 2015а, 2017; Золотарёв, Чулков, 2013; Золотарёв и др., 2009а, б, 2010, 2012, 2014; Кузнецов и др., 2010; Zolotarev et al., 2015]. Она развивалась на основе идей мажоритарного декодирования [Месси, 1966]. Оптимизационная теория (ОТ) создала новые технологии итеративной коррекции ошибок декодирования, базовые методы которой были запатентованы ещё в 1972 г. [Золотарёв, 1972]. Они от-

Кузнецов Николай Александрович — академик РАН, доктор технических наук, профессор, kuznetsov@cplire.ru

Золотарёв Валерий Владимирович — ведущий научный сотрудник, профессор, доктор технических наук, профессор, zolotasd@yandex.ru

Овечкин Геннадий Владимирович — профессор, доктор технических наук, g_ovechkin@mail.ru

Назиров Равиль Равильевич — заведующий отделом, доктор технических наук, rnazirov@cosmos.ru

Аверин Сергей Владимирович — заместитель директора, доктор технических наук, ser-averin@yandex.ru

крыли новую эру эффективных итеративных алгоритмов декодирования, которые могут с линейной от длины кода сложностью сходиться к решениями оптимального декодера (ОД), для чего раньше применялись только переборные методы. Сейчас именно ОТ и только её парадигмы определяют практически все улучшения в характеристиках декодирования для классических моделей каналов.

В настоящее время все основные этапы создания, исследования и проектирования многопороговых декодеров (МПД) проводятся на основе специальных мощных оптимизационных процедур, эффективность и сложность которых быстро растут. При этом сложность самих методов МПД, разработанных на основе ОТ, остаётся минимальной, по-прежнему растущей всего лишь линейно с длиной кода. Но эффективность декодирования на базе МПД оказывается совпадающей с возможностями переборных, т. е. наилучших возможных методов даже для больших уровней искажений в канале связи. Отметим в связи с этим, что в современных сложных научных изысканиях невозможно переоценить значение теорий оптимизации как таковых. Уже весьма давно, например, считается общеизвестной истиной, что роль оптимизационных теорий в математике столь же велика, как и роль математики вообще в науках.

Благодаря применению ОТ с увеличением числа итераций коррекции ошибок и появлением кодов с всё меньшей степенью подверженности размножению ошибок (РО) возможности МПД существенно растут при сохранении весьма небольшой сложности самого алгоритма. К настоящему времени характеристики многопороговых декодеров по сравнению с другими методами при всех практически интересных для техники связи параметрах систем уже стали существенно лучше [Золотарёв, 2006а, б; Золотарёв, Овечкин, 2014; Золотарёв и др., 2012; Кузнецов и др., 2010; Zolotarev et al., 2015]. Ниже сначала рассмотрены возможности МПД, соответствующих главным кодовым кластерам (типичным наборам параметров кодов и моделей каналов) [Гринченко и др., 2006, 2015; Золотарёв, 1972, 1983, 2009; Золотарёв, Овечкин, 2015б; Золотарёв, Чулков, 2013; Золотарёв и др., 2009а, б, 2010; Овечкин, Као, 2014; Овечкин и др., 2013; Самойленко и др., 1981; Zolotarev, Averin, 2007; Zolotarev et al., 2009, 2014а, б; www.mtdbest.ru; www.mtdbest.iki.rssi.ru], а затем состоится обсуждение представленных результатов.

ГАУССОВСКИЕ КАНАЛЫ

Рассмотрим характеристики основных алгоритмов декодирования в гауссовском канале при кодовой скорости $R = 1/2$, представленные на рис. 1 (см. с. 101). На нём показаны в традиционном виде зависимости вероятности ошибки на бит $P_b(e)$ различных алгоритмов декодирования как функции от уровня битовой энергии канала E_b/N_0 в децибелах. Вертикаль $C = 1/2$ отмечает уровень шума, при котором пропускная способность гауссовского канала равна кодовой скорости $C = R = 1/2$. Пунктир P_0 отмечает вероятность ошибки в рассматриваемом канале при $R = 1/2$. Граница АТ указывает на предельные возможности турбо кодов, которые, однако, до сих пор не могут быть воплощены в характеристики аппаратуры из-за сложности алгоритмов этого класса. Кривая VA:K7 отражает эффективность повсеместно применяемого алгоритма Витерби (АВ) для свёрточных кодов с длиной кода $K = 7$, а зави-

симось $CC:VA^*RS$ соответствует каскадной схеме на основе АВ и кода Рида – Соломона (РС) [Золотарёв, Овечкин, 2004; Кларк, Кейн, 1987]. Кривая LDPC приведена для min-sum-декодера кода с низкой плотностью проверок на чётность (LDPC) стандарта DVB-S2 длиной 64 800 бит, реализованного в НИИ Радио [Овечкин, Чикин, 2009]. График TR представляет реальные возможности декодера для турбо-кода длиной 3060 бит стандарта CDMA2000.

МПД-алгоритм в свёрточном варианте реализации в двоичном гауссовском канале и 4-битовом квантовании сигнала в демодуляторе показан на рис. 1 на графике MTD1. В настоящее время он практически столь же оптимально, как и переборные алгоритмы, декодирует длинные коды при очень низкой энергетике гауссовского канала $E_b/N_0 = 1,2$ дБ, когда до его пропускной способности C оказывается всего лишь около 1 дБ [Золотарёв, 2009; Золотарёв и др., 2012; Zolotarev et al., 2015; Золотарёв, 2018]. Таким образом, МПД1 работает при таком уровне шума, когда мощность передатчика только на $\sim 26\%$, т.е. лишь на четверть превышает её уровень при $C = 1/2$. Для работы декодера требуется не более $I = 192$ итераций. Величина задержки декодирования для построенного кода с малым РО и с кодовым расстоянием $d = 21$ составляет менее 8 Мбит кодовых символов. Объём собранной статистики для всех точек этого графика превышает $2,3 \cdot 10^9$ бит.

Далее, при незначительном снижении уровня шума до $E_b/N_0 = 1,5$ дБ для МПД уже нужны лишь $I = 90$ итераций и задержка решения свёрточного декодера около 1 Мбит, как это показывает график MTD2. А при $E_b/N_0 = 1,8$ дБ обычный МПД декодер с 40 итерациями (кривая MTD3) оказывается лучше несравненно более сложной последовательной каскадной схемы алгоритма Витерби с декодером кода Рида – Соломона, при всего лишь втрое большей задержке. Не в последнюю очередь надо помнить, что эта каскадная схема имеет ещё и на $\sim 0,6$ дБ (на 12,5 %) меньшую кодовую скорость R , чем у рассматриваемых МПД алгоритмов, что ещё более наглядно показывает преимущества алгоритмов, созданных в соответствии с ОТ.

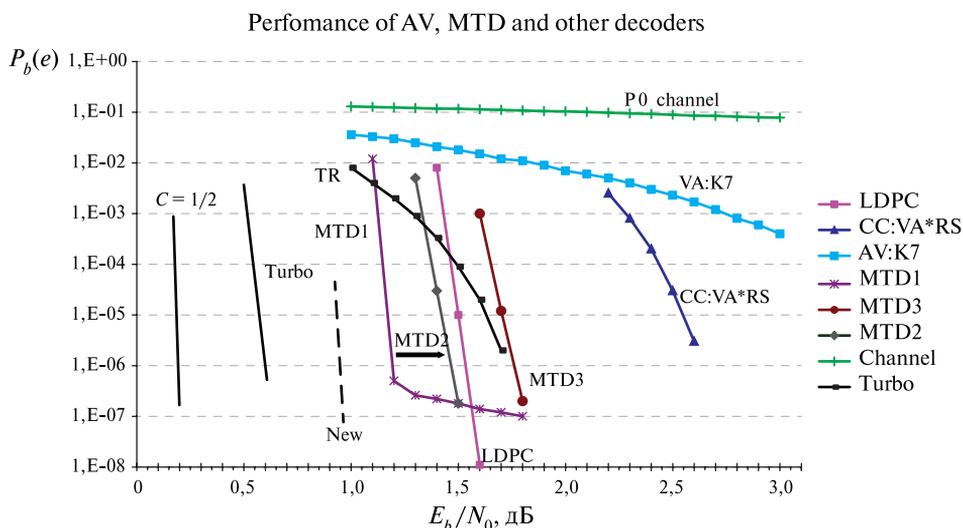


Рис. 1. Характеристики основных алгоритмов декодирования в канале с аддитивным белым гауссовским шумом (АБГШ) при кодовой скорости $R = 1/2$

Детальное сравнение МПД декодеров и других основных алгоритмов для гауссовских каналов, которые относительно успешно развиваются исследователями кодирования у нас и в зарубежье, показывает, что к характеристикам МПД при $E_b/N_0 \sim 1,2$ дБ за последнее десятилетие даже не приблизились никакие алгоритмы типа LDPC, турбо или какие-либо ещё методы с сопоставимой сложностью декодирования. Реальные их характеристики за последнее десятилетие так и не преодолели условную границу энергетики $E_b/N_0 \sim 1,5$ дБ даже с использованием каскадных схем, что при такой близости к границе Шеннона является огромной разницей по сравнению с первым графиком для МПД. В работе [Назаров, Щеглов, 2017] указано, что у этих кодов сложна «процедура формирования кодовых слов, определяемая отсутствием упорядоченных структур порождающих и проверочных матриц». Мы же со своей стороны неоднократно указывали, что алгоритмы для кодов LDPC, являясь итеративными, тем не менее не измеряют расстояние своих решений до принятого сообщения и поэтому не относятся к процедурам глобального поиска. Поэтому их относительно приемлемая эффективность является следствием больших вычислительных затрат этих алгоритмов, что не обеспечивает, однако, их хороших характеристик при $R \sim 1/2$ в области энергетики гауссовского канала менее 1,5 дБ. В настоящее время нет никаких указаний на то, что эта граница будет преодолена при разумном уровне сложности декодирования кодов LDPC. Отметим здесь только, что и структура любых схем кодирования на базе свёрточных МПД декодеров при аппаратной реализации всегда гораздо проще, чем у других методов как на передающей, так и, особенно, на приёмной стороне.

Далее, укажем, что МПД алгоритмы могут быть реализованы аппаратно так, что они становятся как бы однократной решающей схемой мгновенного действия. А это приводит к тому, что, как и все алгоритмы, МПД декодеры создают задержку решения, но совершенно не снижают скорости работы любого устройства, в составе которого они работают. В итоге получаем, что представленные выше реализации МПД алгоритмов обеспечивает в аппаратуре любую произвольно высокую производительность, что принципиально невозможно для других методов. Это происходит согласно работам [Золотарёв, 2005, 2009; Золотарёв и др., 2014] так, словно кодовая последовательность просто поступает в декодер и без какой-либо обработки с той же предельно высокой скоростью обычного сдвига данных в чипе, с какой она была введена в декодер, выходит из него, но уже почти без ошибок в информационной части принятого сообщения. Кроме того, и структура связей между ячейками в аппаратном МПД декодере много проще, чем у прочих алгоритмов. Он более чем на 99 % состоит из простейшей памяти на длинных регистрах сдвига, что дополнительно облегчает его создание и отладку [Золотарёв, Овечкин, 2004]. Такое свойство свёрточного МПД естественно назвать максимальной аппаратной теоретической производительностью. Поэтому даже просто сопоставимого с алгоритмами МПД быстрого действия другие алгоритмы коррекции достичь не могут в принципе. Но при этом МПД декодеры, как было показано выше, ещё и работают при таком уровне шума, при котором неработоспособны прочие методы. А так как эта область уже очень близка к пропускной способности канала типа АБГШ, то разница с другими методами, составляющая несколько десятых децибел, оказывается принципиальной и на самом деле очень значимой, непреодолимой, определяющей и перспективы развития алгоритмов МПД в будущем. Наконец, отметим, что и абсолютная разница в числе итераций в пользу МПД там, где прочие алгоритмы всё-таки ещё работают, также весь-

ма значительна. Причём, сложность каждой итерации у декодеров МПД, в которых лишь суммируются небольшие целые числа, тоже существенно меньше. А в случае программной реализации методов МПД за счёт хранения вычисленных сумм на пороговом элементе (ПЭ) каждая итерация вообще превращается в простую проверку значений этих сумм, на что обычно требуется не более 1-2 операций сравнения [Золотарёв, 2006а, б; Золотарёв, 2006; Zolotarev et al., 2015]. Совокупность перечисленных преимуществ алгоритмов МПД в двоичных каналах с АБГШ перед прочими методами столь значима, что МПД декодеры как продукт ОТ к настоящему моменту стали абсолютными лидерами в мировом конкурсе прикладных достижений в одном из главных кластеров разработок систем кодирования как по совокупности параметров близости к границе Шеннона, сложности, быстродействия и помехоустойчивости, так и по всем им в отдельности.

Указанное соотношение по эффективности и сложности реализации между различными декодерами полностью подтверждается и результатами работы демо-программ для различных декодеров, представленных на наших уникальных сетевых порталах [www.mtdbest.ru; www.mtdbest.iki.rssi.ru], которые полностью соответствуют всем рассмотренным выше соотношениям свойств между алгоритмами. В частности, например, МПД алгоритм в программной демоверсии при весьма высоком уровне шума исправляет все ошибки до оптимального (переборного!) уровня всего за 10 итераций на ПК с процессором Core-i7 на тактовой частоте ~ 3 ГГц на скорости более 16 Мбит/с! (см. www.mtdbest.ru, страница «Обучение», блоки (demo_quick.zip) и инструкцию). И пока что не просматривается никаких путей для каких-либо других алгоритмов коррекции ошибок, чтобы хотя бы приблизиться по своим характеристикам простоты реализации, быстродействия и энергетической эффективности к возможностям МПД.

В качестве важнейшей калибровки скорости работы алгоритмов МПД в двоичных симметричных и гауссовских каналах при большом уровне шума вблизи границы Шеннона можно использовать эту демо-программу сверхскоростного МПД декодирования, которая указывает очень жёсткие ориентиры для числа выполняемых этим алгоритмом операций, а также для быстродействия программных версий декодеров этого класса при двоичной передаче.

Для этого на нашем сетевом портале www.mtdbest.ru_co страницы «Обучение» возьмите инструкцию по применению этого быстрого МПД, его exe-модуль и файл входных данных (demo_quick.zip). Далее по простой инструкции запустите его exe-модуль и по выходному файлу проверьте, что демо-модуль работает на скорости 8–22 Мбит/с в зависимости от процессора IBM PC. Демо-программа написана на языке C++ и при кодовом расстоянии $d \sim 10$ для $I \sim 10$ итерации выполняет $\sim dI = 100$ операций сложения и сравнения для каждого декодируемого символа. Это и есть масштабная калибровка сложности декодирования. Производительность МПД легко пересчитывается на любой алгоритм МПД с другими d , I и уровнями шума. Несомненно, сопоставлять возможности МПД с другими методами нужно только если такие методы также написаны на языке C++, создающем лишь небольшие накладные расходы в демо-программе, реализующей алгоритм МПД. Понятно, что предложенный нами подход легко учитывает и масштабирует производительность процессора, если авторы новых конкурирующих алгоритмов будут запускать на своих компьютерах и эту программу. Написанные на языке Матлаб и прочих медленно работающих учебных средствах алгоритмы декодирования не позволяют

корректно сравнивать какие-либо декодеры между собой. Такие исследования не должны публиковаться из-за невозможности выполнения корректного сравнения методов в этих несопоставимых ситуациях.

Наша научная школа полагает, что престижные журналы по телекоммуникациям могли бы активно и очень эффективно поспособствовать росту качества разработок декодеров, которое, к сожалению, сейчас находится на прискорбно низком уровне. Необходимо лишь публиковать все алгоритмы декодирования только в соответствии с нашими предложениями на языке C++ и с калибровкой по скорости декодирования.

Отметим, наконец, что представленные на рис. 1 МПД алгоритмы за последнее время значительно улучшили свои энергетические характеристики. Есть ли сейчас у нас реальная перспектива успешной работы МПД при энергетике гауссовского канала менее 1,0 дБ, как это показывает пунктир “New”? Да, но здесь получать новые результаты очень сложно, так как в этой области уровень абсолютной упругости границы Шеннона нарастает уже чрезвычайно быстро. Таким образом, дополнительное улучшение возможностей МПД реально, но вычислительные затраты уже могут быть весьма значительными. Реальность возможных малозатратных улучшений характеристики МПД в этой области может определиться уже в ближайшем будущем.

СИМВОЛЬНЫЕ КОДЫ

В высшей степени ценным для ОТ, теории кодирования и различных прикладных цифровых систем передачи, обработки, хранения и восстановления данных стало открытие в 1984 г. и к настоящему времени уже полное исследование нашей научной школой символьных кодов [Золотарёв, 1984, 1985, 1990, 2005, 2006а, б; Золотарёв, Овечкин, 2004, 2015а; Золотарёв и др., 2010; Зубарев и др., 2016а; Кузнецов и др., 2010; Овечкин, 2009, 2011; Zolotarev, Averin, 2007; Zolotarev et al., 2015], реализация символьного многопорогового декодирования (QMПД) которых также чрезвычайно проста, как и в случае двоичных их аналогов.

Строго говоря, впервые эти недвоичные коды с мажоритарным декодированием рассмотрел Дж. Мессе (J. Massey), который доказал теоремы 1–4 для них в работе [Мессе, 1966]. Но затем он очень негативно оценил возможности таких кодов в разделах 1.2, 6.2, 6.5, 6.6 и 8.2 этой же книги и больше уже никогда не занимался этой темой. При этом нам до сих пор неизвестны другие сколько-нибудь содержательные работы по мажоритарному декодированию недвоичных кодов, а тем более публикации по итеративным алгоритмам для них.

Рассмотрим возможности недвоичных кодов. На рис. 2 (см. с. 105) представлены характеристики декодеров кодов Рида – Соломона и QМПД при кодовой скорости $R = 1/2$. По горизонтальной оси отложены вероятности ошибки на символ при различных размерах алфавита q , $q = 2^8 = 256$ и $q = 2^{16} = 65\,536$. По вертикальной оси откладываются вероятности ошибки декодера на символ $P_d(s)$ для любых q . График P_0 показывает вероятность ошибки в симметричном недвоичном канале q СК. Кривая $RS2^8$ даёт представление о возможностях кода РС с (n, k, d) параметрами $(255, 128, 128)$, в котором размер символа соответствует одному байту. Далее пунктир $RS-Su2^8$

соответствует нижней оценке возможностей довольно сложного декодера для этого же кода, предложенного Суданом [Sudan, 1997]. График RS2¹⁶ относится к коду РС длины 65 535, который ещё долго будет считаться очень сложным в реализации.

К настоящему времени для символьных кодов получены все основные оценки характеристик декодирования и детально проработана их общая теория как для блоковых, так и для свёрточных кодов. Они могут полностью заменить коды РС во всех возможных приложениях, выигрывая у них и у других потенциальных конкурентов все конкурсы по достоверности и быстродействию [Золотарёв, 1984, 1985; Золотарёв, Овечкин, 2004; Золотарёв и др., 2010]. Причиной этого преимущества являются возможность выбора любой длины этих символьных кодов, которая совершенно не зависит от алфавита выбранного кода q . Но, самое главное (!), символьные коды, как и двоичные, также обеспечивают фактически оптимальное декодирование (эквивалентное переборному!) при использовании простейших мажоритарных методов даже при достаточно большом уровне шума. Это оказалось возможным благодаря очень точной и удачной модификации недвоичного ПЭ по сравнению с двоичным [Золотарёв, 1984, 1985, 20066, 2009, 2015а; Золотарёв, Овечкин, 2004, 2014, 2015а; Кузнецов и др., 2010; Золотарёв и др., 2010, 2012, 2016а; Zolotarev et al., 2015; Zolotarev, Averin, 2007; www.mtdbest.ru; www.mtdbest.iki.rssi.ru]. При этом сложность символьного МПД (QMПД), как и в двоичном случае, оказывается теоретически минимальной, линейной от длины кода.

Полезность символьных кодов становится особенно важной и впечатляюще значимой, если вспомнить, что для недвоичных кодов фактически вообще невозможно создать сколько-нибудь работоспособный алгоритм Витерби, сложность которого в большинстве случаев будет порядка $\sim q^K$, где q — размер алфавита, K — длина кода в символах. Например, для кода с $R = 1/2$ уже при $q \gtrsim 32$ и $K \gtrsim 7$ число отслеживаемых в таком АВ путей превысит миллиард, а характеристики будут слабее, чем у кода Рида — Соломона.

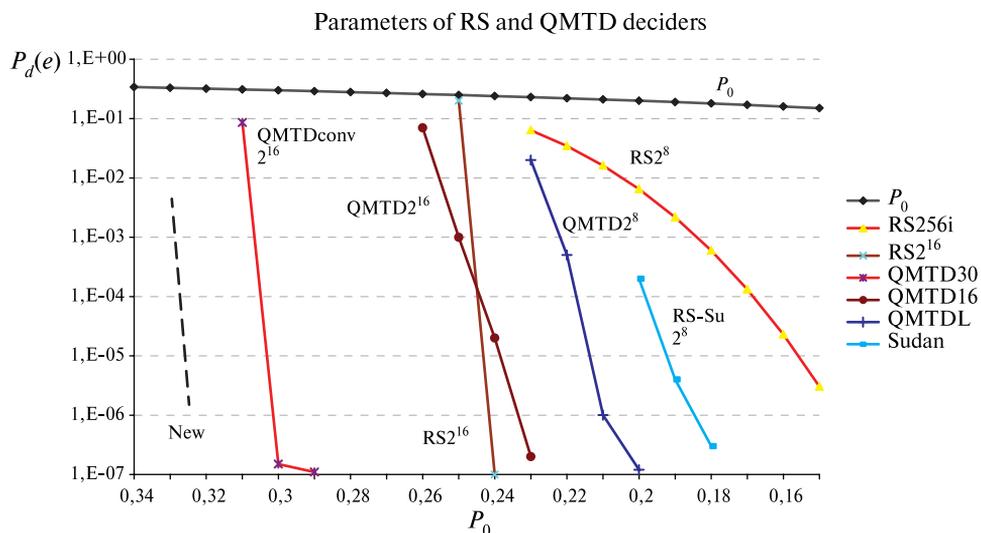


Рис. 2. Характеристики основных алгоритмов декодирования в q -ичном симметричном канале при кодовой скорости $R = 1/2$

Вспомним теперь, что за 50 лет развития недвоичных кодов РС, которые надолго и вполне заслуженно имели широкое поле приложений, ничего реально эффективнее их для недвоичных алфавитов придумано не было. Но реально применяемые коды РС для каскадных схем, оптических дисков и прочих систем относятся к коротким кодам и поэтому малоэффективны. Методы Судана [Зубарев и др., 2016а; Sudan, 1997], позволяющие несколько снизить вероятность ошибки в кодах РС по сравнению со стандартными алгоритмами, также не решили проблему малой эффективности этих кодов даже за счёт весьма значительного усложнения процесса коррекции ошибок.

С другой стороны, большое число публикаций, диссертаций, в том числе и докторских, а также демо-программы для символьных МПД свидетельствуют, что длинные символьные коды с простейшей реализацией и программным быстродействием до десятков мегабит в секунду эффективно работают при вероятностях ошибок, кратно превосходящих уровень допустимых вероятностей для кодов РС [Золотарёв, 1990, 2006а, б, 2009; Кузнецов и др., 2010; Овечкин, 2009, 2011; Zolotarev et al., 2015; www.mtdbest.ru; www.mtdbest.iki.rssi.ru].

Можно указать также на недавнее достижение символьными свёрточными МПД с кодовой скоростью $R = 1/2$ для $q = 2^{16}$ особо высокой помехоустойчивости даже в q СК канале с вероятностью ошибки $P_0 = 0,3$, как показано на рис. 2 в виде графика QMTDconv 2^{16} [Зубарев и др., 2016а, б]. Этот уровень шума недостижим при $R = 1/2$ даже для кодов РС в поле $GF(2^{16})$, реализовать декодер для которых исключительно сложно, а из-за невысоких характеристик декодирования и бесполезно.

В качестве важнейшего на особый текущий момент прикладного результата можно указать на недавнее достижение свёрточными q МПД для кодов с кодовой скоростью $R = 1/2$ при $q = 2^{16}$ возможности субоптимального декодирования в q СК при вероятности ошибки $P_0 = 0,325$, что проиллюстрировано на рис. 2 кривой “New”. Общее число итераций свёрточного QМПД декодера не превышало $I = 185$. Этот результат недостижим для декодеров других практически реализуемых кодов с $R = 1/2$.

Можно также указать в связи с этим на то, что скорости работы символьных МПД действительно очень велики, и благодаря крайне простой идее их реализации на самом деле могут достигать очень высоких значений [Золотарёв, 2005; Золотарёв, Овечкин, 2015а; Золотарёв и др., 2010; Овечкин, 2009, 2011; Zolotarev, Averb, 2007]. График QMTD 2^8 для символьного кода длины 8000 байт при $q = 256$ эффективнее короткого кода РС с декодером Судана [Sudan, 1997]. Сложность программной версии QMTD 2^8 (для того же процессора) определяется скоростью декодирования, равной ~ 300 кс/с (кс — кило-символы), которая, конечно, огромна. Менее чем за 1 ч набирается статистика на более чем миллиард символов, что может составить до $3 \cdot 10^{10}$ бит. Читателям доступна также демо-программа QMTD для малоизбыточного символьного кода с $R = 0,95$ с простой инструкцией по настройке параметров и использованию [www.mtdbest.ru; www.mtdbest.iki.rssi.ru]. Она работает на том же процессоре Core-i7 при очень высоком уровне шума для этого значения R со скоростью порядка 40 Мбит/с. Для перезаписи этой демо-программы на свой ПК нужно с нашего портала www.mtdbest.ru со страницы «Обучение» переписать упакованные инструкцию и программный модуль с рабочими файлами (qmtd_demo.zip). Запуск по инструкции покажет на любом ПК производительность символьного МПД декодера более десятка Мбит/с, причём результаты будут соответствовать оптимальному (переборному!) декодированию (ОД), которое для

недвоичных кодов при $q \gg 1$ никогда не может быть реализовано в полномасштабном виде иначе, чем на основе QМПД. Данная демо-программа, как и модуль для двоичных кодов, может использоваться как калибровочная для точной оценки сложности символьных и иных декодеров в недвоичных каналах.

Каскадные схемы с использованием QМПД были полностью исследованы в работах [Золотарёв, 2006б; Овечкин, 2009, 2011; Zolotarev, Averin, 2007; Zolotarev et al., 2015; www.mtdbest.ru; www.mtdbest.iki.rssi.ru].

Символьные МПД на много лет вперёд полностью решили все проблемы цифрового мира по передаче и, особенно, хранению и восстановлению цифровых данных в любых информационных системах в диапазоне кодовых скоростей $R = 0,5-0,98$ при любом реальном уровне надёжности используемых информационных байтово-ориентированных массивов.

Как и двоичные МПД, символьные алгоритмы декодирования могут быть реализованы на теоретически максимальных скоростях при их аппаратной реализации [Золотарёв, 2005, 2009, 2018; Золотарёв и др., 2014].

СТИРАЮЩИЕ КАНАЛЫ

Обратимся ещё к одной важной области теории кодирования, которая создаёт и изучает декодеры для каналов со стираниями. Однако до применения к ней методов ОТ и алгоритмов МПД результаты для этих каналов были крайне скромны у всех методов [Гринченко и др., 2006; Золотарёв, 1983, 1990, 2006а; Месси, 1966; Zolotarev et al., 2015; Зубарев и др., 2016а]. Основной причиной этого было, конечно, гораздо большее внимание специалистов именно к каналам с ошибками. Собственно, именно поэтому исследования декодеров для стирающих каналов были как бы в тени, а их характеристики весьма далеки от допускаемых теорией пределов эффективной работы этих алгоритмов.

Ограничимся только кратким рассмотрением тех характеристик МПД, которые показаны на рис. 3. Из последних результатов в этой области следует, что для восстановления стираний методами МПД [Золотарёв, 1983, 1990, 2006а, б, 2017; Zolotarev et al., 2015; Золотарёв и др., 2017б] при $R = 1/2$ необходим свёрточный код с минимальным кодовым расстоянием $d \geq 21$ и особенно малым группированием своих неправильных решений при декодировании, т. е. с минимальной подверженностью эффекту РО. Соответствующий МПД алгоритм с таким кодом, представленный графиком MTD05, успешно восстанавливает поток стёртых (любых!) символов с вероятностью независимых стираний $P_{es} = 0,48$ в канале до уровня вероятности оставшихся невосстановленных символов $P_s(s) \sim 3 \cdot 10^{-7}$. Таким образом, в канале с пропускной способностью $C = 0,52$ МПД алгоритм даже без использования каскадирования успешно работает при отношении $R/C \sim 0,961$, что является абсолютно уникальным достижением для процедур восстановления стираний. При этом скорость работы соответствующей МПД демо-программы при $I = 90$ итерациях коррекции была равна 95 Кс/с. Другие быстрые методы восстановления стираний при $R = 1/2$ для таких уровней искажений на входе декодера неизвестны. График MTD45 соответствует свёрточному коду с $R = 4/5$, с которым МПД при $P_{es} \leq 0,19$ успешно работал при $R/C \approx 0,988$ [Зубарев и др., 2016а]. Такой результат также крайне труден для повторения какими-либо другими алгоритмами коррекции стираний.

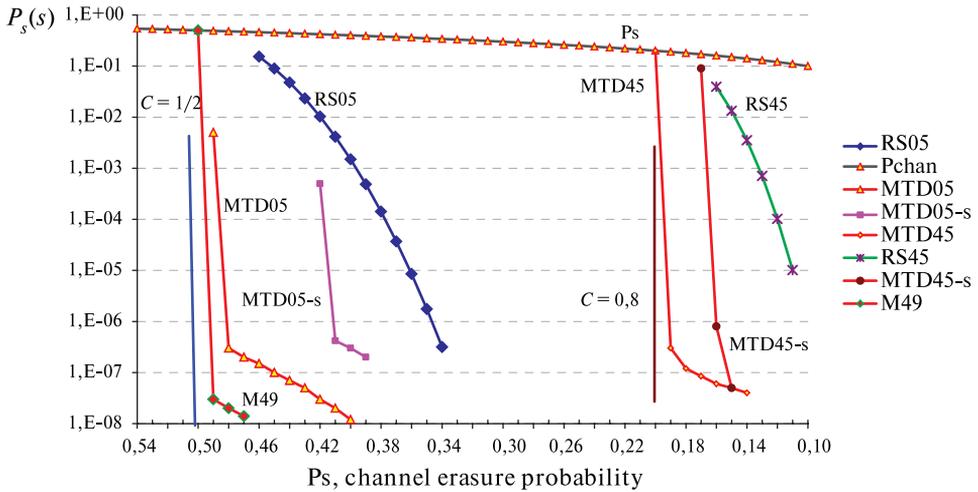


Рис. 3. Характеристики МПД и кодов РС в каналах со стираниями

Графики MTD05-s и MTD45-s для гораздо более простых алгоритмов этого типа также соответствуют той области искажений символов в канале, в которой декодеры для кодов РС совершенно неработоспособны [Золотарёв и др., 2017б; Зубарев и др., 2016а].

Весь материал этого раздела основан на данных исследований нескольких последних лет. Однако буквально в последний перед публикацией самых новых данных по МПД алгоритмам интерес к исследованиями декодеров, исправляющих стирания, заметно возрос. Хорошей иллюстрацией этого является [Золотарёв, 2018], где обсуждались различные алгоритмы такого типа, предлагавшиеся многими авторами. Нелишне подчеркнуть, что среди всех рассмотренных там методов МПД декодеры были самыми простыми и быстрыми, а их характеристики эффективности были в числе лучших. Но за последнее время нашей научной школой были проведены новые обширные исследования именно среди кодов, предназначенных для коррекции стираний. Результат таких исследований представлен на рис. 3 графиком M49. Он соответствует моделированию работы МПД декодера, восстанавливающего стирания с новым кодом при $R = 1/2$ и $d = 25$.

Как следует из вида графика, МПД алгоритм эффективно работает в стирающем канале уже при его вероятности стираний $p_s = 0,49$. Это соответствует отношению $R/C = 50/51 \sim 0,98$ и, видимо, достижение вероятности ошибки (невосстановления значения символа) порядка $3 \cdot 10^{-8}$, что демонстрирует обсуждаемый СтМПД, является для прочих методов неразрешимой задачей. Предложенная схема СтМПД относится к обычным уже описанным версиям некаскадных декодеров, активно использующим принцип дивергентности, который применялся при построении моделировавшегося кода и при создании декодера несколько раз [Золотарёв, 2015б; Золотарёв, Овечкин, 2015б]. Число использовавшихся итераций коррекции было при $p_s = 0,49$ не более $I = 198$ раз, а полная задержка решения декодера составила менее 2 млн символов. Поскольку проведённому эксперименту придавалось особенно большое значение, дополнительно было исследовано и поведение МПД алгоритма в довольно небольшой области вероятностей канала при $p_s > 0,49$. Код и алгоритм

показали хорошую устойчивость работы в этой буферной области, что подтвердило правильность выбора направления завершающих исследований по этой тематике. Для более точного измерения характеристик декодирования в области $ps > 0,49$ объём эксперимента был выбран заведомо бóльшим, чем этого требовали статистические критерии, и составил более $6 \cdot 10^{10}$ информационных символов в каждой из нижних точек графика. Полученные вероятности соответствовали потенциальным возможностям кода с $d = 25$ при оптимальном декодировании.

Производительность декодера на ПК с процессором Core-i7 также в демо-программе на языке C++ близка к 60 кс/с, что для столь эффективного восстановления принятых символов при $R \approx C$ следует считать отличным результатом, закрывающим, наконец, гонку создания кодов и алгоритмов теперь уже и для стирающих каналов. Полезно отметить, что именно простота и высокая скорость декодирования стирающего МПД позволили быстро набрать указанный выше огромный объём статистики на обычном ноутбуке.

Демо-платформу «Программа моделирования работы многопорогового декодера в канале связи со стираниями...», которая представляет собой демо-программу моделирования на языке C++ работы МПД алгоритма, предназначенного для восстановления стёртых при передаче данных символов, можно переписать с портала www.mtdbest.ru со страницы «Обучение». После перезаписи на свой ПК этой программы и инструкции по её использованию можно с использованием предложенных там или иных кодов заняться исследованиями работы МПД в таком стирающем символами канале. Такие исследования могут стать основой для разработок более сложных кодовых систем, например, исправляющих и ошибки, и стирания.

Основной вывод по этому разделу состоит в том, что и в стирающих каналах алгоритмы МПД тоже не имеют равных по простоте реализации и эффективности декодирования вблизи границы Шеннона, т.е. при $R \lesssim C$.

Напомним ещё раз, что на портале www.mtdbest.ru как раз рядом с указанной выше демо-платформой находятся аналогичные расширенные программные ресурсы, т.е. такие же демо-платформы для моделирования работы МПД в гауссовском: «Программа моделирования работы многопорогового декодера» и недвоичном канале: «Программа моделирования работы недвоичного многопорогового декодера», где вы можете посмотреть возможности как кодов, которые предлагаем мы, так и тех кодов, которые захотите выбрать вы сами. Там также можно менять как разные параметры для МПД алгоритмов, так и уровни шума каналов, при которых вы хотите их проверить.

СПЕЦИАЛЬНЫЕ ПРИЛОЖЕНИЯ ОПТИМИЗАЦИОННОЙ ТЕОРИИ

Отметим кратко другие результаты ОТ, которые оказались важнейшими для теории кодирования и её приложений.

Во-первых, это масштабные исследования различных высокоскоростных алгоритмов МПД при кодовой скорости $R \sim 0,8$ и выше. Все свойства МПД и их соотношение с возможностями других алгоритмов, сохраняются и в этой области кодовых скоростей [Золотарёв, 1984, 1985, 1990, 2006а, б, 2009; Золотарёв, Овечкин, 2004; Золотарёв, Чулков, 2013; Овечкин, 2009, 2011; Zolotarev, Averin, 2007; Zolotarev et al., 2015].



Рис. 4. Действующий макет МПД декодера на ПЛИС Altera для гауссовского канала на информационную скорость 1 Гбит/с, созданный в ИКИ РАН

Например, в работах [Золотарёв, 2006б; Золотарёв, Овечкин, 2014; Zolotarev et al., 2015; Zolotarev et al., 2014a, b] рассмотрены методы высокоскоростного МПД декодирования при $R = 4/5$ для оптических каналов при гауссовской модели канала, полученные нашей научной школой и зарубежными авторами, которых мы консультировали. Результаты российских исследователей были существенно лучше по энергетической эффективности и простоте реализации МПД алгоритмов. С учётом теоретически максимально возможной скорости декодирования МПД наши методы имеют существенные преимущества.

Применение методов МПД декодирования оказывается особенно полезным при передаче и хранении очень больших массивов данных, в том числе и на борту спутников систем дистанционного зондирования Земли (ДЗЗ). Вероятность нарастающих со временем искажений в таких системах памяти может быть весьма высокой, а необходимая скорость передачи этих данных с орбиты часто также очень велика, что приводит к необходимости использования там только МПД алгоритмов. Рабочий макет простого МПД-декодера на программируемой логической интегральной схеме (ПЛИС) Altera для спутникового канала на скорость более 1 Гбит/с, созданный в ИКИ РАН, представлен на рис. 4. Подобный ему чип с МПД декодером был ранее разработан в НИИ Радио. Такие характеристики высокоскоростного декодирования для других алгоритмов, видимо, будут ещё очень долго недостижимы.

Далее можно указать на совсем небольшую сложность реализации МПД алгоритмов в канале типа двоичных симметричных каналов (ДСК), с которых и начались много лет назад исследования МПД [Золотарёв, 1986]. В качестве наглядного примера высокой эффективности МПД в этой области укажем на важнейшую фундаментальную и одновременно сложнейшую техническую задачу декодирования для флеш-памяти. Специфика этой проблемы состоит в том, что вероятность ошибки на бит для таких систем должна быть не хуже

$P_b(e) \sim 10^{-15}$, а в скором времени требования к достоверности решений соответствующих декодеров ещё более усилятся. В исходных массивах такой памяти вероятность ненадёжного хранения символов может быть на уровне $p_e \sim 10^{-3}$ и даже существенно большей. Найденные специалистами нашей научной школы технические решения на основе МПД алгоритмов позволили довольно просто решить эту проблему [Золотарёв, 2005; Зубарев и др., 2016а; Овечкин и др., 2013; Zolotarev et al., 2015]. При этом использовался код с МПД декодированием при $R = 3/4$, что гарантированно позволяет обеспечить требуемые уровни достоверности даже при вероятности ненадёжного хранения $p_e \sim 10^{-2}$. Столь же простые и одновременно эффективные результаты по сверхдостоверному декодированию в условиях применения малонадёжной памяти у других разработчиков нам неизвестны.

Дальнейшие несложные эффективные методы улучшения надёжности подобных систем могут быть сделаны с помощью методов каскадного кодирования, которые особенно просты для МПД декодеров [Золотарёв, 1990, 2006а, б; Золотарёв, Овечкин, 2004; www.mtdbest.ru; www.mtdbest.iki.rssi.ru].

СОСТОЯНИЕ КЛАССИЧЕСКОЙ ТЕОРИИ КОДИРОВАНИЯ

Ограничимся кратким обсуждением тех наиболее важных теоретических, математических и технологических ресурсов, которые образуют пространство, совокупность знаний, понятий и методов решения задач, называемых обычно парадигмами той или иной отрасли науки, в нашем случае — теории кодирования.

Отметим сначала главные парадигмы (результаты, методы и технологии) классической теории кодирования, длительное время связываемой с алгебраической теорией. Конечно же, она имела гораздо более широкое поле развития. Однако даже после появления алгоритма Витерби основные направления развития классической теории включали как важнейший аспект развития разработку алгебраических методов декодирования (в дискретных полях). Но хотя классы новых кодов множились, а декодеры понемногу становились более простыми, долгое время исправление ошибок за пределами половины кодового расстояния было для алгебраических кодов неразрешимой задачей. Кроме того, относительное расстояние d/n этих кодов уменьшается с ростом их длины n . Таким образом, исправление ошибок становилось всё более слабым. Поэтому наилучшим научным результатом, широко применяемым в технике декодирования, уже 40 лет является каскадная схема АВ с кодами РС, известная уже 40 лет. И это всё!

Хорошими примерами эвристического подхода, выросшего из классической теории, стали в последние десятилетия турбо и низкоплотностные (LDPC) коды, которые позволили немного приблизиться к пропускной способности канала. Однако с началом нового тысячелетия это движение фактически остановилось, хотя возможности элементной базы и теории увеличивались. Дело в том, что эти методы излишне сложны и, например, многие вопросы их применения для сверхточных кодов и больших скоростей всё ещё не решены. Кроме того, все эти декодеры не учитывают расстояние между полученным вектором и текущими решениями. Есть веские основания считать, что именно эти обстоятельства ограничивают возможности подобных

декодеров. Поэтому поиски новых методов в рамках алгебраической теории продолжают.

Одним из промежуточных итогов этого поиска стали полярные коды, к которым сразу обратилось большое число специалистов по кодам [Arıkan, 2009]. Но до сих пор реальные результаты разработки таких кодов представлены очень невнятно, а иногда и весьма странно. Анализ этого направления по имеющимся работам был нами сделан в работе [Золотарёв, Овечкин, 2016]. Более содержательные данные по этим кодам, если они появятся, позволят сделать более точные прогнозы по их перспективам. А сейчас весьма слабые результаты для полярных кодов компенсируются обращением к декодированию списками и созданию крайне сложных каскадных схем для этих кодов, что пока полностью исключает их из группы перспективных алгоритмов.

Не лишне напомнить, что аналогичные подходы «переключения интереса» несколько ранее вынуждены были использовать специалисты по алгебраическим методам и последовательным алгоритмам, возможности которых даже теоретически были очень далеки от границы Шеннона. Однако значительно улучшить эти методы не удалось. Весьма интересное направление 1980-х гг. [Блох, Зяблов, 1976] тоже не получило тогда хорошего импульса для своего развития. Возможно, это было связано с тем, что это направление было в большей степени сориентировано на алгебраические коды со слабыми характеристиками. Возможно, последним единственным успешным результатом алгебраической теории следует считать алгоритмы Судана [Sudan, 1997], которые немного улучшили эффективность декодирования кодов РС и послужили развитию дискретной математики. К сожалению, эти методы тоже оказались весьма сложными, а их реальная эффективность декодирования выросла незначительно. Этими подходами и исчерпывается пока реальный прогресс классической теории, хотя предложенный здесь комментарий, конечно, должен быть гораздо более детальным.

ОСНОВНЫЕ РЕСУРСЫ ОПТИМИЗАЦИОННОЙ ТЕОРИИ

Определим важнейшие парадигмы оптимизационной теории, которые позволяют сейчас постоянно получать принципиально новые результаты практически для всех групп кодовых кластеров, некоторые основные типы которых мы рассматривали выше.

Разработка оптимизационной теории началась около 45 лет назад с появлением метода особого итеративного декодирования на основе мажоритарных процедур [Золотарёв, 1972; Самойленко и др., 1981] и Основной Теоремы многопорогового декодирования (ОТМТД), философскую целеполагающую роль которой невозможно переоценить [Золотарёв, 1990, 2006а, б; Золотарёв, Овечкин, 2014; Кузнецов и др., 2010; Zolotarev et al., 2015]. Этот метод сразу позволил в двоичных симметричных каналах при использовании как блоковых, так и свёрточных кодов исправлять число ошибок в несколько раз большее, чем половина кодового расстояния [Золотарёв, 2006а, б; Zolotarev et al., 2015]. При этом, что было абсолютно неожиданным для специалистов, ОТМПД гарантировала при каждом изменении декодируемых символов строгое приближение к оптимальному решению, которое обычно достигается только переборными

методами, известным элегантным примером которых является АВ для свёрточных кодов. Дальнейшим развитием ОТ стало распространение возможностей МПД на все прочие каналы с независимыми искажениями.

Вторым фундаментальным результатом ОТ стала всесторонняя разработка теории эффекта размножения ошибок (РО), без которой значимость ОТМПД была бы весьма ограниченной. А совместное синергетическое воздействие этих двух фундаментальных теоретических результатов на эффективность МПД оказалась столь огромным, что при линейной от длины кода сложности этим алгоритмом сразу оказалось возможным исправлять на фактически оптимальном уровне эффективности почти все ошибки в блоковых и свёрточных кодах длины $n = 10^4 - 10^5$ битов и более. Напомним, что реальная длина кода для переборного АВ даже в будущем, видимо, не будет более 30 битов, чем и определяются все преимущества МПД, который остался алгоритмом с теоретически минимальной сложностью и наилучшими возможными результатами декодирования.

Дальнейший прогресс ОТ поддерживался созданием нескольких поколений высокопроизводительного программного обеспечения для разработки и исследований различных типов МПД и средств поиска кодов с малым уровнем РО, что позволило постепенно приближать уровень практически оптимального (эквивалентного переборному!) декодирования на основе МПД к границе Шеннона.

Следующими очень мощными методами улучшения характеристик МПД для всех каналов стали открытые нами ещё в 1986 г. параллельные методы каскадирования, технологии улучшенного способа назначения порогов, весов проверок и подбора кодов, а также методы ускоренного вычисления сумм на пороговых элементах [Золотарёв, 2006а, б; Zolotarev et al., 2015]. Особое влияние на идеологию ОТ оказало выделение в особую группу декодеров с прямым контролем метрик (ДПКМ) методов МПД и АВ, которые точно измеряют расстояние до принятого сообщения, что и позволяет сейчас только этим алгоритмам успешно работать при больших уровнях шума. Другие методы с такими важнейшими свойствами не известны.

Для сохранения простоты реализации алгоритмов МПД мы также пересмотрели принципы каскадирования, известные нам, и нашли очень простые решения, которые без усложнения этих декодеров многократно повышают эффективность каскадных схем [Золотарёв, 1990, 2006а, б; Золотарёв, Овечкин, 2014; Овечкин, 2009, 2011; Zolotarev et al., 2015]. Но эта тема заслуживает отдельных публикаций.

Принципиальное значение, решающим образом меняющее соотношение свойств эффективности и сложности между различными алгоритмами, имеет открытие нашей научной школой блоковых АВ (БАВ). Их сложность, как и в случае свёрточных кодов по порядку числа операций и памяти равна 2^K , K — длина кода, тогда как до недавнего времени наши теоретики полагали, что сложность таких методов имеет порядок, более близкий к 2^{2K} [Кудряшов, 2016]. БАВ также относится к группе ДПКМ, что очень сильно уменьшает потенциальные перспективы полярных кодов, методов Чейза и других алгоритмов для блоковых кодов как в гауссовских каналах, так и в ДСК [Золотарёв, Овечкин, 2015в, 2017; Золотарёв, Назиров, 2013; Золотарёв и др., 2017а]. Очевидно, что большое число проблем эффективного декодирования блоковых кодов теперь будут очень успешно решать простые каскадные схемы с использованием БАВ.

Ещё больший простор в разработке методов декодирования на основе ОТ открылся после точной формулировки принципов дивергентного декодирования [Золотарёв, 2015б, в; Золотарёв, Овечкин, 2015б; Зубарев и др., 2016б; www.mtdbest.ru; www.mtdbest.iki.rssi.ru]. Этот подход чётко определил отдельный класс простых некаскадных методов увеличения кодового расстояния используемых кодов, который дополнительно приблизил рабочие характеристики алгоритмов ОТ к пропускной способности канала С. Они использовались и в МПД алгоритмах, представленных на рис. 1–3, что особенно сильно улучшило работу МПД при экстремально высоких уровнях шума. Важно, что при этом можно для некоторых кодов одновременно использовать несколько различных методов коррекции ошибок.

Оптимизационная теория успешно осваивается мировым научно-техническим сообществом. Наши крупнейшие в России сетевые порталы по ОТ и МПД алгоритмам www.mtdbest.ru (РГРТУ) и www.mtdbest.iki.rssi.ru (ИКИ РАН) посещают за год, как показано на рис. 5 до 100 тысяч читателей из более 80 стран мира и переписывают около 20 Гб информации по теории и технологиям нашей научной школы [Зубарев и др., 2016б; www.mtdbest.ru; www.mtdbest.iki.rssi.ru], включая множество демо-программ по всем известным лучшим алгоритмам декодирования. Их можно переписать на свои ПК и сразу непосредственно заняться исследованиями по технике декодирования согласно прилагаемым к ним инструкциям. Такое масштабное внимание мировой науки к разработкам российской научной школы является уникальным и свидетельствует о серьёзных достижениях нашей теории кодирования.

Наши порталы по ОТ и МПД

www.mtdbest.ru

www.mtdbest.iki.rssi.ru

**За 2016 год - более 105 тыс. читателей
на наших порталах из 94 стран мира**

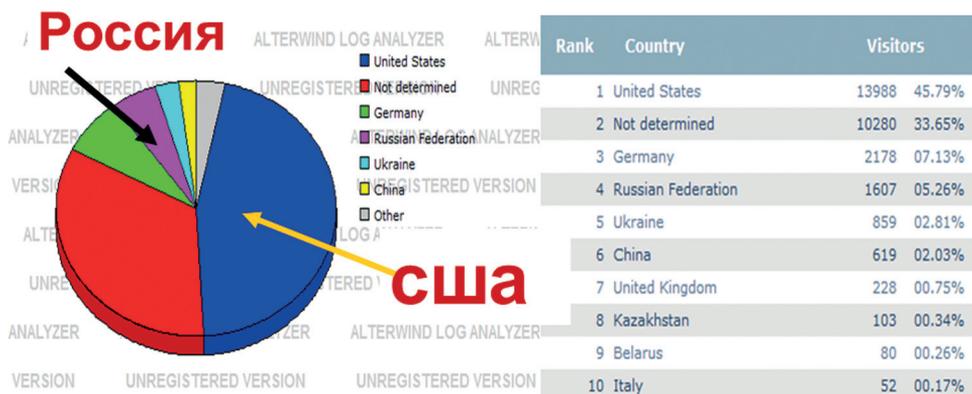


Рис. 5. Диаграмма посещаемости сайтов в течение года научно-методических порталов ИКИ РАН и РГРТУ по тематике ОТ и МПД

Рассмотренные методы, свойства и возможности кодов и алгоритмов в рамках ОТ, как видно из их описания, все без исключения позволяющие улучшить характеристики декодирования простейшими способами, образуют очень мощное перспективное интеллектуальное пространство развития методов декодирования. Эти новые парадигмы ОТ уже позволили создать методы для разных кодовых кластеров, которые эффективно работают непосредственно вблизи границы Шеннона. Развитие всех этих методов быстро переносится на новые кодовые структуры.

ЗАКЛЮЧЕНИЕ

Нашей научной школой представлены результаты успешного решения главной проблемы цифрового мира - достижение простыми методами требуемой достоверности цифровых потоков при уровнях шума каналов основных классических типов, близких к пропускной способности каналов связи.. Сложность процедур поиска глобального экстремума растёт с увеличением длины используемых кодов в теоретически минимальном темпе, т.е. всего лишь линейно. Это значит, что в год 70-летнего юбилея публикации великой статьи Шеннона [Shannon, 1948] мы предъявили наилучшее технологичное решение проблемы высокодостоверной передачи и хранения цифровых потоков в условиях сильно шумящих каналов и ненадёжной памяти. Как в физике достижение скорости света материальными телами невозможно, так и совсем небольшое расстояние (никогда не достижение!) рабочей области алгоритмов МПД до пропускной способности C рассматривавшихся выше каналов, что следует из графиков на рис. 1–3, является особенно наглядным свидетельством абсолютной успешности и состоятельности созданной нами ОТ. А огромное преимущество наших методов по эффективности и одновременно по сложности реализации перед всеми прочими алгоритмами коррекции ошибок позволяет утверждать, что поставленная 70 лет назад Шенноном проблема сегодня решена, причём, на вполне приемлемом технологическом уровне, что подтверждает и большое число патентов по алгоритмам ОТ и МПД, лишь частично представленных в этой статье [Золотарёв, 1972, 2005, 2009, 2015а, 2017; Золотарёв, Овечкин, 2005, 2015а, 2017; Золотарёв и др., 2014].

Важность, успешность и перспективность ОТ можно сравнить только с появлением квантовой механики в самом начале прошлого века. Но тогда физика как наука не столь много значила в жизни людей. Тем не менее, результаты Планка, Шрёдингера и целой плеяды других великих физиков были высоко оценены научным сообществом, и они заслуженно стали Нобелевскими лауреатами.

Достижение, фактически в одиночку, аналогичных по масштабу и значимости для цифрового мира результатов нашей научной школой, конечно, имеет **гораздо** большее значение, так как мы решили важнейшую и чрезвычайно сложную проблему простого достижения высокой достоверности цифровых потоков в период взрывного развития нашей информационной цифровой цивилизации, своевременно создав все возможности для применения наших простых и понятных методов помехоустойчивого кодирования в современной цифровой технике и науке будущего.

Таким образом, изложенные в статье данные свидетельствуют об успешном решении главной научной и технологической проблемы всей нашей информационной цифровой цивилизации – создании обширных классов простых методов достижения произвольно высокой достоверности передачи, хранения и восстановления цифровой информации в непосредственной близости от границы Шеннона. Решение этой сложнейшей технологической задачи позволяет переключить усилия инженеров и учёных на решение новых важных проблем цифровой информатики нашего мира.

БЛАГОДАРНОСТИ

Исследования оптимизационной теории в течение прошедшего полувекового периода её эволюции с момента первых публикаций нашей научной школы поддерживали: МФТИ, ИППИ РАН, концерт «Созвездие», Совет по кибернетике АН СССР, ОНИТ РАН, РГРТУ, НИИ Радио, МНИТИ, ИКИ РАН. Методы многопороговых декодеров тестировали ООО «ОРТ», НПО им. С. А. Лавочкина, а также ряд организаций и предприятий отрасли связи. Ценность финансовой поддержки разработки методов многопороговых декодеров со стороны РАН, РГРТУ, РФФИ и других источников также была исключительно велика (проекты РФФИ № 14-07-00859, 14-07-00824, 08-07-00078 и 05-07-00024). Мы полагаем, что число наших последователей и сторонников, как и раньше, будет расти, а новые сферы исследований и разработок с использованием парадигм нашей новой «квантовой теории» в области помехоустойчивого кодирования, т.е. оптимизационной теории, технологий многопороговых декодеров и новых видов алгоритма Витерби, а также новых каскадных схем будет быстро расширяться.

Большинство ссылок, относящихся к этой статье, как и многие другие материалы по оптимизационной теории и МПД-алгоритмам, в том числе множество демо-программ, можно найти на наших сетевых порталах [www.mtdbest.ru; www.mtdbest.iki.rssi.ru] и в нашей новой монографии [Золотарёв, 2018].

ЛИТЕРАТУРА

- [Блох, Зяблов, 1976] Блох Э. Л., Зяблов В. В. Обобщённые каскадные коды. М.: Связь, 1976.
- [Гринченко и др., 2006] Гринченко Н. Н., Золотарёв В. В., Овечкин Г. В., Овечкин П. В. Применение многопорогового декодера в каналах со стираниями // Труды НТО-РЭС им. А. С. Попова. 2006. С. 338–340.
- [Гринченко и др., 2015] Гринченко Н. Н., Као В. Т., Овечкин Г. В. Повышение эффективности многопорогового декодера // 17-я Международная конф. «Цифровая обработка сигналов и ее применение — DSPA 2015». 2015.
- [Золотарёв, 1972] Золотарёв В. В. Устройство для декодирования линейных свёрточных кодов. А. С. СССР № 492878 от 1972.
- [Золотарёв, 1983] Золотарёв В. В. Многопороговое декодирование в стирающих каналах // Вопросы радиоэлектроники. Сер. ЭВТ. 1983. Вып. 10. С. 67–70.
- [Золотарёв, 1984] Золотарёв В. В. Многопороговое декодирование в недвоичных каналах // Вопросы радиоэлектроники. Сер. ЭВТ. 1984. Вып. 12. С. 73–76.

- [Золотарёв, 1985] *Золотарёв В. В.* Алгоритмы коррекции символьных данных в вычислительных сетях // Вопросы кибернетики: сб. ст. 1985. Вып. 105. С. 54–62.
- [Золотарёв, 1986] *Золотарёв В. В.* Многopороговое декодирование // Проблемы передачи информации. 1986. Т. 22. Вып. 1. С. 104–109.
- [Золотарёв, 1990] *Золотарёв В. В.* Субоптимальные алгоритмы многopорогового декодирования: дис. ... д-ра техн. наук. М., 1990. 278 с.
- [Золотарёв, 2005] *Золотарёв В. В.* Высокоскоростное устройство многopорогового декодирования линейных кодов: патент на полезную модель. № 44216 от 27.02.2005.
- [Золотарёв, 2006а] *Золотарёв В. В.* Теория и алгоритмы многopорогового декодирования. М.: Радио и связь; Горячая линия – Телеком, 2006 (2-е изд. 2014).
- [Золотарёв, 2006б] *Золотарёв В. В., Зубарев Ю. Б., Овечкин Г. В.* Многopороговые декодеры и оптимизационная теория кодирования / под ред. акад. РАН В. К. Левина. М.: Горячая линия – Телеком, 2012. 238 с.
- [Золотарёв, 2009] *Золотарёв В. В.* Способ декодирования помехоустойчивого кода: патент на изобретение. РФ № 2377722 от 27.12.2009.
- [Золотарёв, 2015а] *Золотарёв В. В.* Способ декодирования помехоустойчивого кода: патент на изобретение. РФ № 2557454 от 25.06.2015.
- [Золотарёв, 2015б] *Золотарёв В. В., Овечкин Г. В., Ташатов Н. Н.* Применение принципа дивергенции при декодировании свёрточных кодов // 3-я Международная научно-практич. конф. «Информационная безопасность в свете Стратегии Казахстана-2050». 2015. С. 158–164.
- [Золотарёв, 2015в] *Золотарёв В. В.* Применение дивергентного кодирования в каналах спутниковой связи и ДЗЗ // 13-я Всерос. открытая конф. «Современные проблемы дистанционного зондирования Земли из космоса». 2015. С. 29.
- [Золотарёв, 2017] *Золотарёв В. В.* Способ обнаружения и исправления стираний при приёме дискретной информации: патент на изобретение. РФ № 2611235 от 21.02.2017.
- [Золотарёв, 2018] *Золотарёв В. В.* Теория кодирования как задача поиска глобального экстремума // под науч. ред. акад. РАН Н. А. Кузнецова. 2-е изд., испр. М.: Горячая линия – Телеком, 2018, 228 с.
- [Золотарёв, Назиров, 2013] *Золотарёв В. В., Назиров Р. Р.* Блоковая модификация алгоритма Витерби // 11-я Всерос. открытая конф. «Современные проблемы дистанционного зондирования Земли из космоса». 2013. С. 136.
- [Золотарёв, Овечкин, 2004] *Золотарёв В. В., Овечкин Г. В.* Помехоустойчивое кодирование. Методы и алгоритмы: Справочник. М.: Горячая линия – Телеком, 2004. 126 с.
- [Золотарёв, Овечкин, 2005] *Золотарёв В. В., Овечкин Г. В.* Устройство многopорогового декодирования линейных кодов для гауссовских каналов: патент на полезную модель. № 44215 от 27.02.2005.
- [Золотарёв, Овечкин, 2014] *Золотарёв В. В., Овечкин Г. В.* Применение многopороговых методов декодирования помехоустойчивых кодов в высокоскоростных системах передачи данных // Электросвязь. 2014. № 12. С. 10–14.
- [Золотарёв, Овечкин, 2015а] *Золотарёв В. В., Овечкин Г. В.* Способ работы символьного порогового элемента в символьном мажоритарном декодере: патент на изобретение. РФ № 2573741 от 22.12.2015.
- [Золотарёв, Овечкин, 2015б] *Золотарёв В. В., Овечкин Г. В.* Дивергентное кодирование свёрточных кодов // 18-я Международная научно-техн. конф. «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций». Рязань. 2015. С. 27–32.
- [Золотарёв, Овечкин, 2015в] *Золотарёв В. В., Овечкин П. В.* Характеристики декодирования блоковых кодов по алгоритму Витерби для систем ДЗЗ // 13-я Всерос. открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса». 2015. С. 163.

- [Золотарёв, Овечкин, 2016] *Золотарёв В. В., Овечкин Г. В.* О сопоставлении новых методов помехоустойчивого кодирования // Доклады 18-й Международной конф. «Цифровая обработка сигналов и её применение». Москва. 2016. Т. 1. С. 59–64.
- [Золотарёв, Овечкин, 2017] *Золотарёв В. В., Овечкин П. В.* Способ кодирования и декодирования блокового кода с использованием алгоритма Витерби: патент на изобретение. РФ № 2608872 от 25.01.2017.
- [Золотарёв, Чулков, 2013] *Золотарёв В. В., Чулков И. В.* Малоизбыточное кодирование для высокоскоростных каналов: тез. докл. // 11-я Всерос. открытая конф. «Современные проблемы дистанционного зондирования Земли из космоса». 2013. С. 31.
- [Золотарёв и др., 2009а] *Золотарёв В. В., Назиров Р. Р., Никифоров А. В., Чулков И. В.* Новые возможности многопорогового декодирования по высокодостоверной передаче данных ДЗЗ // Современные проблемы дистанционного зондирования Земли из космоса. 2009. Вып. 6. Т. 1. С. 167–173.
- [Золотарёв и др., 2009б] *Золотарёв В. В., Назиров Р. Р., Чулков И. В., Овечкин Г. В.* Алгоритмы МПД // Российский космос. 2009. № 1. С. 60–63.
- [Золотарёв и др., 2010] *Золотарёв В. В., Овечкин Г. В., Назиров Р. Р., Овечкин П. В., Чулков И. В.* Эффективное недвоичное многопороговое декодирование помехоустойчивых кодов для систем дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса. 2010. Т. 7. № 2. С. 269–274.
- [Золотарёв и др., 2014] *Золотарёв В. В., Овечкин Г. В., Сатыбалдина Д. Ж., Таишатов Н. Н., Адамова А. Д.* Способ мягкого многопорогового декодирования помехоустойчивого кода: удостоверение автора (патент Республики Казахстан) № 93989 от 15.10.2014.
- [Зубарев и др., 2016а] *Зубарев Ю. Б., Золотарёв В. В., Овечкин Г. В., Овечкин П. В.* Оптимизационная теория кодирования: итоги 25 лет развития // 18-я Международная конф. «Цифровая обработка сигналов и её применение — DSPA 2016». 2016. Т. 1. С. 6–12.
- [Зубарев и др., 2016б] *Зубарев Ю. Б., Золотарёв В. В., Овечкин Г. В., Овечкин П. В.* Итоги 25-летнего развития оптимизационной теории кодирования // Научноёмкие технологии. 2016. Т. 17. С. 26–32.
- [Золотарёв и др., 2017а] *Золотарёв В. В., Овечкин Г. В., Овечкин П. В.* Блоковая модификация алгоритма Витерби // Вестник РГРТУ. Рязань. 2017. № 59. С. 30–35.
- [Золотарёв и др., 2017б] *Золотарёв В. В., Овечкин Г. В., Баринев И. В.* Применение самоортогональных кодов в каскадных схемах кодирования для каналов связи со стираниями // 19-я Международная конф. «Цифровая обработка сигналов и её применение — DSPA 2017». 2017. Т. 1. С. 75–79.
- [Кларк, Кейн, 1987] *Кларк Дж., Кейн Дж.* Кодирование с исправлением ошибок в системах цифровой связи. М.: Радио и связь, 1987.
- [Кудряшов, 2016] *Кудряшов Б. Д.* Основы теории кодирования: учеб. пособие. СПб.: БХВ-Петербург, 2016. 393 с.
- [Кузнецов и др., 2010] *Кузнецов Н. А., Золотарёв В. В., Овечкин Г. В., Овечкин П. В.* Недвоичные многопороговые декодеры и другие методы коррекции ошибок в символьной информации // Радиотехника. 2010. № 6. Вып. 141. С. 4–9.
- [Месси, 1966] *Месси Дж.* Пороговое декодирование / пер. с англ. Ю. Л. Сагаловича; под ред. Э. Л. Блоха. М.: Мир, 1966. 208 с.
- [Назаров, Щеглов, 2017] *Назаров Л. Е., Щеглов М. А.* Характеристики полных и укороченных помехоустойчивых низкоплотностных кодов на основе конечных геометрий. // Успехи современной радиоэлектроники. 2017. № 6. С. 38–48.
- [Овечкин, 2009] *Овечкин П. В.* Разработка алгоритмов повышения эффективности недвоичных многопороговых декодеров в системах передачи и хранения больших объемов информации: дис. ... канд. техн. наук. Рязань: РГРТУ, 2009. 131 с.
- [Овечкин, 2011] *Овечкин Г. В.* Теория каскадного декодирования линейных кодов для цифровых радиоканалов на основе многопороговых алгоритмов: дис. ... д-ра техн. наук. Рязань: РГРТУ, 2011. 301 с.

- [Овечкин, Као, 2014] *Овечкин Г. В., Као В. Т.* Многопороговые декодеры для гауссовских каналов // 19-я Всероссийская научно-техн. конф. «Новые информационные технологии в научных исследованиях и в образовании». Рязань, РГРТУ. 2014. С. 121–122.
- [Овечкин, Чикин, 2009] *Овечкин Г. В., Чикин А. В.* Помехоустойчивость приемника спутниковых сигналов DVBS2 // 11-я Международная конф. и выставка «Цифровая обработка сигналов и ее применение — DSPA 2009». 2009. С. 578–580.
- [Овечкин и др., 2013] *Овечкин Г. В., Золотарев В. В., Федюков В. С.* Повышение достоверности хранения цифровых данных на флеш-памяти // 6-я Международная научно-техн. конф. «Космонавтика. Радиоэлектроника. Геоинформатика». Рязань. 2013. С. 201–203.
- [Самойленко и др., 1981] *Самойленко С. И., Давыдов А. А., Золотарев В. В., Третьякова Е. Л.* Вычислительные сети. М.: Наука, 1981. 278 с.
- [Arikan, 2009] *Arikan E.* Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels // IEEE Trans. Information Theory. 2009. V. 55. No. 7. P. 3051–3073.
- [Shannon, 1948] *Shannon C.* A Mathematical Theory of Communication // Bell System Technical J.. 1948. V. 27. July P. 379–423, October P. 623–656. (= Шеннон К. Э. Математическая теория связи / пер. В. Ф. Писаренко // Работы по теории информации и кибернетике. М.: Иностранная литература, 1963. С. 242–332.).
- [Sudan, 1997] *Sudan M.* Decoding of Reed Solomon codes beyond the error-correction bound // J. Complexity. 1997. V. 13. P. 180–193.
- [Zolotarev, Averin, 2007] *Zolotarev V. V., Averin S. V.* Non-Binary Multithreshold Decoders with Almost Optimal Performance // 9th Intern. Symp. Communication Theory and Applications (ISCTA'07). UK, Ambleside, 2007.
- [Zolotarev et al., 2009] *Zolotarev V. V., Ovechkin G. V., Averin S. V.* Algorithm of Multithreshold Decoding for Self-Orthogonal Codes over Gaussian Channels // 11th Intern. Symp. Communication Theory and Applications (ISCTA'09. UK, Ambleside, 2009. 6 p.
- [Zolotarev et al., 2014a] *Zolotarev V., Ovechkin G., Satybalдина D., Tashatov N., Adamova A., Mishin V.* Effective multithreshold decoder for optical and other data transmission systems // Latest trends on Communications: Proc. 18th Intern. Conf. Communications (part of CSCC'14). Santorini Island, Greece. 2014. P. 152–156.
- [Zolotarev et al., 2014b] *Zolotarev V., Ovechkin G., Satybalдина D., Tashatov N., Adamova A., Mishin V.* Efficiency multithreshold decoders for self-orthogonal block codes for optical channels // Intern. J. Circuits, Systems and Signal Processing. 2014. V. 8. P. 487–495.
- [Zolotarev et al., 2015] *Zolotarev V. V., Zubarev Y. B., Ovechkin G. V.* Optimization Coding Theory and Multithreshold Algorithms // ITU Publications. 2015. URL: <http://www.itu.int/pub/S-GEN-OCTMA-2015>.

CHARACTERISTICS OF OPTIMIZATION THEORY ALGORITHMS NEAR THE SHANNON BOUNDARY

N. A. Kuznetsov¹, V. V. Zolotarev², G. V. Ovechkin³, R. R. Nazirov², S. V. Averin⁴

¹ Kotel'nikov institute of radio engineering and electronics of Russian Academy of Sciences

² Space Research Institute of Russian Academy of Sciences (IKI RAN)

³ Ryazan State Radio Engineering Institute (RSREU)

⁴ LLC "United Radio-Electronic Technologies"

The characteristics of decoding algorithms based on the Optimization theory (OT) of noise-resistant coding are analyzed. Multithreshold decoders are discussed for Gaussian, binary, non-binary,

erase and other channels. Opportunities of the block Viterbi algorithm are discussed. The prospects for divergent coding are highlighted. A group of direct metric control algorithms is noted. The value of the main paradigms of the OT is discussed. The main ways of development of coding theory for the coming years and for the future are determined.

Keywords: telecommunication systems, data storage systems, error-correction coding, block Viterbi algorithm, computer simulation, self-orthogonal codes, symbolic codes, threaded codes, multithreshold decoders, the coding gain, divergent coding

Kuznetsov Nikolay Alexandrovich — academician of RAS, doctor of technical sciences, professor, kuznetsov@cplire.ru

Zolotarev Valery Vladimirovich — leading researcher, professor, doctor of technical sciences, zolotasd@yandex.ru

Ovechkin Gennady Vladimirovich — professor doctor of technical sciences, g_ovechkin@mail.ru

Nazirov Ravil Ravilyevich — head of department, doctor of technical sciences

Averin Sergey Vladimirovich — Deputy Director, doctor of technical sciences, ser-averin@yandex.ru

АППАРАТУРА ДЛЯ ИССЛЕДОВАНИЯ ПОЛОЖЕНИЯ ЛИНИИ МОРСКОГО ГОРИЗОНТА В ВИДИМОЙ ОБЛАСТИ СПЕКТРА С ИСПОЛЬЗОВАНИЕМ ПОЛЯРИЗАЦИОННЫХ СВЕТОФИЛЬТРОВ

И. А. Маслов¹, В. А. Гришин^{1,2}, И. В. Николенко³

¹ Институт космических исследований Российской академии наук (ИКИ РАН)

² Московский государственный технический университет им. Н. Э. Баумана (МГТУ)

³ Институт астрономии Российской академии наук (ИНАСАН)

В настоящее время наблюдается значительный интерес к исследованиям методов навигации, обладающих устойчивостью и надёжностью в условиях возможного радиоэлектронного противодействия. Одним из возможных источников навигационной информации, обладающим очень высокой устойчивостью к противодействию, является классический метод навигации, основанный на использовании видимого положения Солнца, Луны и звёзд. При этом для навигации используется информация о точном времени, а также информация о направлении местной вертикали, которую можно получить из наблюдаемой линии горизонта. Однако вариации температуры, давления и влажности атмосферы порождают вариации степени её рефракции, что вызывает ошибки определения ориентации местной вертикали. Для летательных аппаратов указанные ошибки ещё более возрастают из-за процессов рассеивания и ослабления оптического излучения на атмосферных аэрозолях. Все это приводит к тому, что основным источником ошибок навигации являются как раз ошибки определения местной вертикали. Некоторое уменьшение влияния рассеянного атмосферой излучения может быть достигнуто при использовании поляризационных светофильтров. Для оценки перспективности использования таких светофильтров для выделения горизонта, а также для оценки стабильности атмосферной рефракции, была разработана и запущена в эксплуатацию описываемая ниже аппаратура. Отличительной особенностью аппаратуры является её автономная работа, обычно не требующая внешнего управления, с автоматическим запуском сразу после подачи на неё напряжения питания.

Ключевые слова: навигация, мониторинг, линия горизонта, море, атмосфера, автоматизация, удалённое управление

Маслов Игорь Алексеевич — заведующий сектором, imaslov@iki.rssi.ru

Гришин Владимир Александрович — старший научный сотрудник, кандидат технических наук, доцент, vgrishin@iki.rssi.ru

Николенко Игорь Владимирович — младший научный сотрудник, niki63@yandex.ru

ВВЕДЕНИЕ

Поскольку широко используемые в настоящее время спутниковые навигационные системы обладают принципиальным недостатком — очень малой устойчивостью к средствам радиоэлектронного противодействия (что обусловлено очень малой мощностью навигационного сигнала на поверхности Земли), то ведутся активные работы по расширению числа источников навигационной информации, в частности, для беспилотных летательных аппаратов (БПЛА). Например, это могут быть навигационные системы технического зрения, работающие по форме наблюдаемой линии горизонта (при полёте над материками или в непосредственной близости к ним), либо по звёздному небу или по положению Луны и Солнца (при полете над морями и океанами). В последнем случае наблюдение линии горизонта используется для оценки ориентации местной вертикали, необходимой для расчёта координат. При этом основным источником ошибок навигации являются как раз ошибки оценки ориентации местной вертикали. Рассмотрим использование датчиков горизонта.

Датчики горизонта находят широкое применение на борту искусственных спутников Земли и используются для угловой ориентации аппаратов [STD 15, 2004; STD 16, 2001; IRES, 2015]. Большинство этих датчиков работает в инфракрасном диапазоне и использует механическое сканирование. Более продвинутые сенсоры используют матричные детекторы [Meller et al., 2000; Scheidegger et al., 2007].

Датчики горизонта используются также на беспилотных летательных аппаратах (БПЛА). Однако в данном случае такие датчики используются, главным образом, не для навигации, а для угловой стабилизации летательных аппаратов [Cornall, Egan, 2004; Ettinger et al., 2002; Horiuchi, 2009; Shabayek et al., 2012; Taylor et al., 2003]. Точность таких датчиков невелика, поэтому они неприменимы для навигации.

Нельзя не отметить классические методы навигации на море, когда производятся угловые измерения положения Солнца, Луны и звёзд относительно линии горизонта и по специальным таблицам, а также по показаниям хронометра, производится определение широты и долготы судна.

Однако наблюдение линии горизонта с борта БПЛА имеет свои особенности. Дело в том, что расстояние до горизонта существенно зависит от высоты полёта. Чем больше высота полёта, тем, естественно, это расстояние больше. Поскольку прозрачность атмосферы ограничена, то начиная с некоторой высоты, мы перестаём видеть реальный горизонт, а видим лишь линию максимального градиента яркости в вертикальном направлении, положение которой определяется уже только процессами распространения и рассеивания света в атмосфере [Grishin, Maslov, 2018].

Следует отметить, что даже когда высота полёта невелика, и мы можем наблюдать реальный горизонт, то процессы рефракции в атмосфере могут очень сильно влиять на оценку местной вертикали a , следовательно, и оценку текущих координат. Например, в работах [Sampson et al., 2003; Schaefer, Liller, 1990; Young, 2004] представлены результаты высокоточного определения степени рефракции в атмосфере. Рефракция определялась путём фиксации времени восхода и захода Солнца, по которым определялась вариация величины рефракции. Измерения показали, диапазон разброса величины рефракции в атмосфере весьма велик — порядка $0,234$ – $1,678^\circ$ [Schaefer, Liller, 1990]

и $0,402-2,081^\circ$ [Sampson et al., 2003]. В других измерениях, угловое положение линии горизонта варьировалось в интервале от $-0,246$ до $0,326^\circ$ [Young, 2004]. Несложно определить, что ошибка оценки направления местной вертикали всего в одну угловую минуту даёт ошибку по координатам приблизительно в одну морскую милю.

СОСТОЯНИЕ ПРОБЛЕМЫ

Получить более надёжные и представительные в статистическом смысле результаты о возможности наблюдения линии горизонта и колебаниях его положения возможно при проведении исследований, для проведения которых требуется разработка специализированных приборов. В настоящее время известно довольно большое число приборов, предназначенных для исследования состояния атмосферы. Некоторые из них выпускаются серийно (например, TSI-440 и TSI-880 производства Yankee Environmental Systems). Однако пожалуй наиболее серьёзной разработкой являются Whole sky imagers, описанные в работах [Shields et al., 1992, 2003, 2013] и большом количестве других публикаций. Приборы разрабатывались в Marine Physical Laboratory of the Scripps Institution of Oceanography University of California, San Diego. Приборы обладают большим динамическим диапазоном, функционируют в нескольких спектральных диапазонах. Данные приборы предназначены для длительной круглосуточной работы в автономном режиме и в режиме телеуправления с автоматическим сбором информации с периодом от 1 до 10 мин. Приборы имеют сетевой доступ. Самые первые версии приборов были разработаны в 1950-х гг. С тех пор были разработаны несколько поколений приборов, последние версии которых используются и в настоящее время. Информация, полученная с таких и подобных приборов, используется для мониторинга (определению и классификации) облаков, изучению свойств атмосферных аэрозолей, измерения ночной прозрачности атмосферы и других задач. К сожалению, для разработки, изготовления и эксплуатации подобных приборов необходимо соответствующее финансирование.

ЦЕЛИ РАЗРАБОТКИ АППАРАТУРЫ

Проведённые исследования [Grishin, Maslov, 2018] показали заметный суточный дрейф линии горизонта. Очевидно, что некоторое уменьшение влияния рассеянного атмосферой излучения может быть достигнуто при использовании поляризационных светофильтров, подавляющих часть рассеянного атмосферой солнечного излучения. Для оценки перспективности использования таких светофильтров для выделения горизонта, а также для оценки стабильности атмосферной рефракции и накопления соответствующей статистики была разработана и запущена в эксплуатацию описываемая ниже аппаратура. Отличительной особенностью аппаратуры является её автономная работа, не требующая внешнего управления, с запуском сразу после подачи на неё напряжения питания.

ОПИСАНИЕ АППАРАТУРЫ

Аппаратура (рис. 1) содержит три камеры с полем зрения $62 \times 54^\circ$, дающих цветное RGB-изображение 3280×2464 пикселей (сенсор Sony IMX219 Exmor) в формате BMP. Перед двумя камерами были дополнительно поставлены поляризационные светофильтры с главными осями ориентированными параллельно и перпендикулярно линии горизонта. Камеры установлены на площадке телескопа Цейсс-1000 Института астрономии РАН, вблизи п. Симеиз на высоте 370 м над уровнем моря (рис. 2, см. с. 125). Ориентировочная дальность до горизонта с этой высоты составляет 72 км. Поскольку выходная информация используемых сенсоров имеет только 256 градаций яркости, то автоматический подбор экспозиции осуществляется по полосе изображения в районе линии горизонта. Экспозиции варьируются от нескольких миллисекунд днём до нескольких минут ночью. Определение ориентации камер осуществляется по ярким звёздам, изображения которых получаются ночью, и по наземным объектам. Астрометрическая калибровка камер была выполнена по звёздам путём ориентации их полей зрения в зенит.

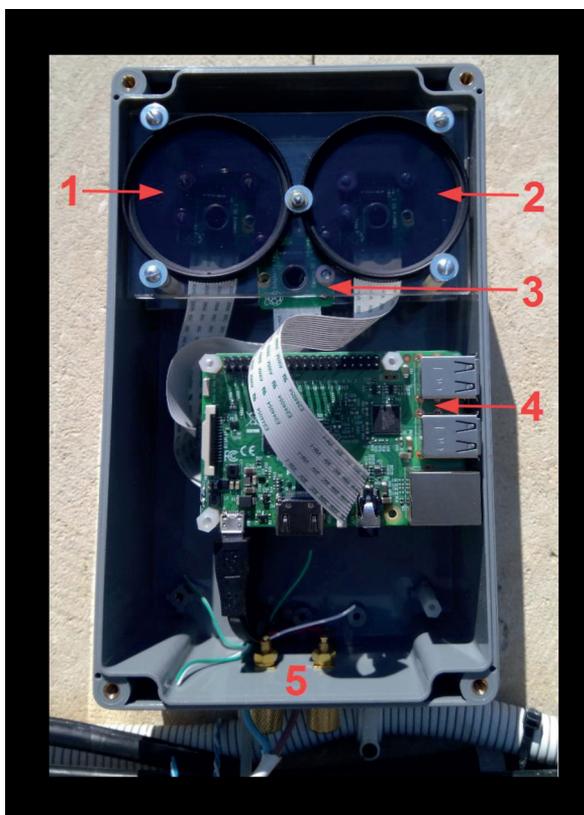


Рис. 1. Вид аппаратуры со снятой крышкой: 1, 2 — поляризационные светофильтры перед камерами, выделяющие, соответственно, горизонтальную и вертикальную поляризации света; 3 — камера без поляризационного светофильтра; 4 — три микрокомпьютера; 5 — питание 5 В



Рис. 2. Аппаратура, установленная на обсерватории

Камеры включаются автоматически при подаче на них напряжения питания (что немаловажно при наличии возможности отключения электропитания) и работают круглосуточно, записывая полученные изображения в сжатом виде (7z) на Яндекс-диск, откуда периодически переписываются исследователями на их носители информации для дальнейшей обработки (рис. 3, см. с. 126). Величина сжатого файла позволяет программно оценить количество информации в изображении и избежать ситуации, когда при передержке как таковое изображение отсутствует. Один из утренних снимков проведён на рис. 4 (см. с. 126).

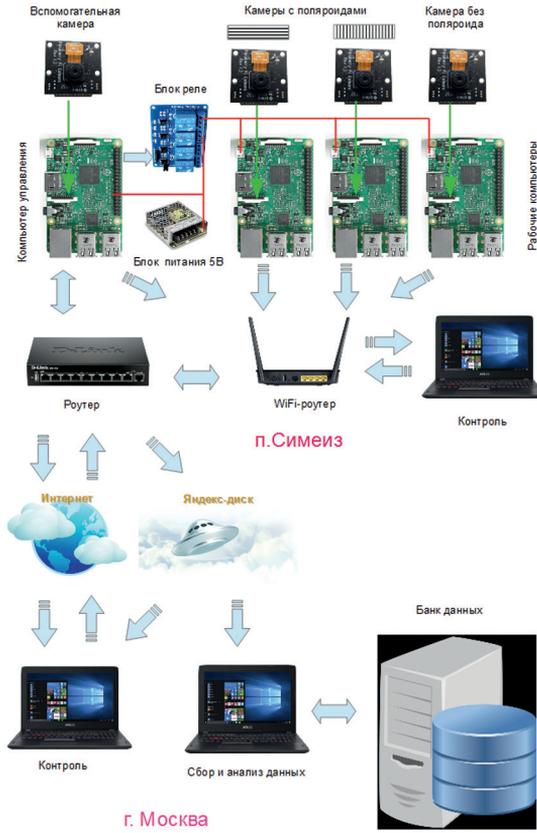


Рис. 3. Структура аппаратуры, структура управления работой аппаратуры и структура передачи информации



Рис. 4. Один из утренних снимков

Управление камерами и приём информации от них осуществляется компьютерами Raspberry Pi, использующими ARM-микропроцессор BCM2835. Компьютеры работают под управлением ОС Linux, программа управления написана на языке Python. Три компьютера с камерами включаются и начинают работать при подаче на них напряжения 5В (суммарный ток не более 3А) и, используя WiFi-соединение, через роутер связываются с сетью Интернет и с компьютерами локальной сети, с которых можно при необходимости корректно дистанционно (через интернет) включение и выключение питающего напряжения. А именно, при сбоях в электропитании компьютеры могут запуститься раньше запуска WiFi роутера и не инициализировать связь с Яндекс-дискон. Управляющий реле компьютер имеет как WiFi, так и проводное соединение с локальной сетью и это позволяет в большинстве случаев дистанционно восстановить работу аппаратуры.

РЕЗУЛЬТАТЫ

На рис. 5 приведены примеры полученной информации: снимки для моментов восхода, захода и верхней кульминации Солнца и вертикальные сканы яркости в RGB цветах в районе горизонта, используемые для оценки его положения.

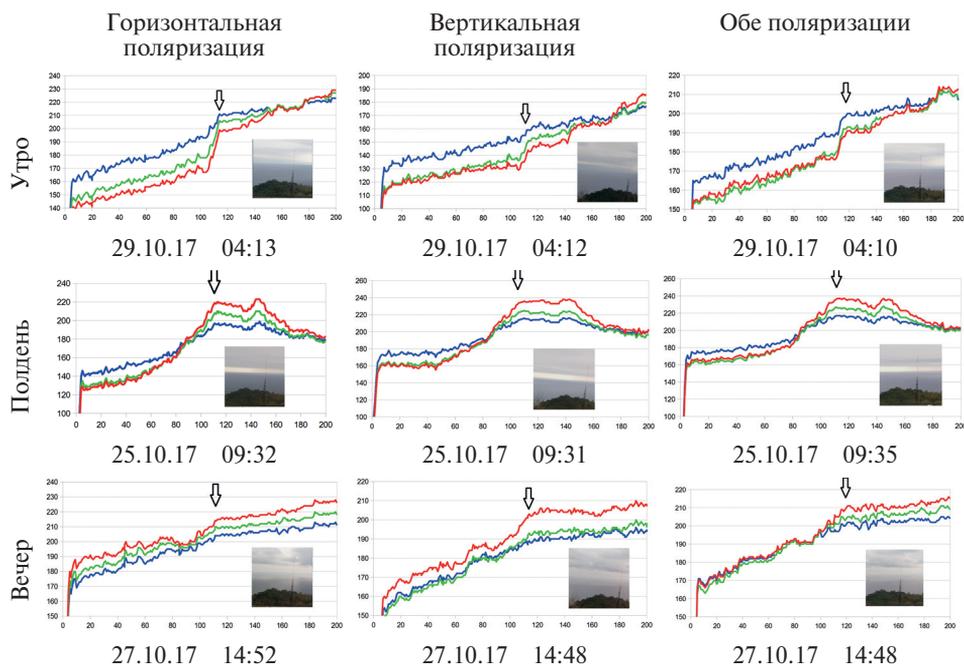


Рис. 5. Примеры полученной информации. Показаны зависимости яркости изображения для трёх цветов RGB от номера пикселя, отсчитываемого от вершины высоковольтной вышки линии электропередачи видимой на сопровождаемых кадрах. Указаны даты и всемирное время (UTC) для моментов экспозиции в местные «Утро», «Полдень» и «Вечер». Стрелкой показано положение линии горизонта

По вертикальной — оси графиков откладывается яркость, по горизонтальной оси — угловое расстояние в пикселях от наземного ориентира. Один пиксель приблизительно соответствует одной угловой минуте. Можно было бы ожидать улучшения выделения линии горизонта при уменьшении рассеянного атмосферой солнечного излучения за счёт выделения света, плоскость поляризации которого направлена на Солнце, как это происходит для облаков. То есть, при ориентации камер на юг, днём лучше выделять только вертикальную поляризацию, а утром и вечером — горизонтальную. Однако в случае наблюдения горизонта результат получается более сложным. Вероятно, это обусловлено, как близостью морской поверхности, так и многократным рассеянием света при наклонных трассах, приводящим к деполяризации света [Шмирко и др., 2015].

ВЫВОДЫ

Описанный комплекс работает с октября 2017 г. по настоящее время, передавая ежесуточно примерно по 300 снимков. Предварительный анализ полученных данных указывает на улучшение, в некоторых случаях, выделения линии горизонта при ориентации оси поляроида таким образом, чтобы исключить однократно рассеянное солнечное излучение. Анализу статистики положения линии горизонта будет посвящена отдельная статья.

ЛИТЕРАТУРА

- [Шмирко и др., 2015] Шмирко К. А., Павлов А. Н., Бобриков А. А. Расчёт отражённого взволнованной морской поверхностью диффузного солнечного излучения // Вестник ДВО РАН. 2015. № 3. С. 36–44.
- [Cornall, Egan, 2004] Cornall T., Egan G. Measuring Horizon Angle from Video Onboard a UAV // Proc. IEEE International Conference on Autonomous Robots and Agents. Palmerston North, New Zealand. 2004. P. 339–344.
- [Ettinger et al., 2002] Ettinger S. M., Nechyba M. C., Ifju P. G., Waszak M. Towards Flight Autonomy: Vision-Based Horizon Detection for Micro Air Vehicles // Proc. Florida Conf. Recent Advances in Robotics. Miami, FL. 2002. P. 1–7.
- [Grishin, Maslov, 2018] Grishin V. A., Maslov I. A. Horizon Line Stability Observations over the Sea // J. Navigation. 2018. V. 71. Iss. 2. P. 437–453. URL: <https://www.cambridge.org/core/journals/journal-of-navigation/article/horizon-line-stability-observations-over-the-sea/302BEA1AF7150C9010B238CF69456D46>.
- [Horiuchi, 2009] Horiuchi T. A Low-Power Visual-Horizon Estimation Chip // IEEE Trans. Circuits and Systems. 2009. V. 56. Iss. 8. P. 1566–1575.
- [IRES, 2005] IRES Infrared Earth Sensor / Selex ES S.p.A. A Finmeccanica Company. 2015. 2 p. URL: http://www.leonardocompany.com/documents/63265270/65642596/IRES_Infrared_Earth_Sensor_LQ_mmm07787_.pdf?download_file.
- [Meller et al., 2000] Meller D., Sripruetkiat P., Makovec K. Digital CMOS Cameras for Attitude Determination // Proc. 14th Annual AIAA/USU Conf. Small Satellites. 2000. Paper SSC00-VII-1.
- [Sampson et al., 2003] Sampson R. D., Lozowski E. P., Peterson A. E., Hube D. P. Variability in the Astronomical Refraction of the Rising and Setting Sun // Astronomical Society of the Pacific. 2003. V. 115. V. 812. P. 1256–1261.

- [Schaefer, Liller, 1990] *Schaefer B. E., Liller W.* Refraction near the Horizon // Astronomical Society of the Pacific. 1990. V. 102. Iss. 653. P. 796–805.
- [Scheidegger et al., 2007] *Scheidegger N., Krpoun R., Shea H., Niclass C., Charbon E.* A new Concept for a Low-cost Earth Sensor: Imaging Oxygen Airglow with Arrays of Single Photon Detectors // 30th Annual AAS Guidance and Control Conf. Breckenridge, Colorado. 2007. AAS 07-062.
- [Shabayek et al., 2012] *Shabayek A., Demonceaux C., Morel O., Fofi D.* Vision Based UAV Attitude Estimation: Progress and Insights // J. Intelligent and Robotic Systems. 2012. V. 65. Iss. 1. P. 295–308.
- [Shields et al., 1992] *Shields J. E., Johnson R. W., Karr M. E.* An Automated Observing System for Passive Evaluation of Cloud Cover and Visibility // Marine Physical Laboratory: Final Report. SIO Ref. 92-22, MPL-U-65/92. University of California, San Diego. 1992.
- [Shields et al., 2003] *Shields J. E., Johnson R. W., Karr M. E., Burden A. R., Baker J. G.* Whole Sky Imagers For Real-Time Cloud Assessment, Cloud Free Line of Sight Determinations and Potential Tactical Applications // Battlespace Atmospheric and Cloud Impacts on Military Operations(BACIMO) Conf. Monterey, CA. 2003.
- [Shields et al., 2013] *Shields J. E., Karr M. E., Johnson R. W., Burden A. R.* Day/night whole sky imagers for 24-h cloud and sky assessment: history and overview // Applied Optics. 2013. V. 52. Iss. 8. P. 1605–1616.
- [STD 15, 2004] STD 15 Earth Sensor / SODERN Company. 2004. URL: http://www.sodern.com/sites/docs_wsw/RUB_52/STD15.pdf.
- [STD 16, 2001] STD 16 Earth Sensor / SODERN Company. 2001. URL: http://www.sodern.com/sites/docs_wsw/RUB_52/STD16.pdf.
- [Taylor et al., 2003] *Taylor B., Bil C., Watkins S., Egan G.* Horizon Sensing Attitude Stabilisation: A VMC Autopilot // Proc. 18th International UAV Systems Conference. Bristol, UK. 2003. P. 1–9.
- [Young, 2004] *Young A. T.* Sunset Science. IV. Low-Altitude Refraction // Astronomical J. 2004. V. 127. Iss. 6. P. 3622–3637.

EQUIPMENT FOR THE STUDY OF THE POSITION OF THE SEA HORIZON LINE IN THE VISIBLE SPECTRUM USING POLARIZATION FILTERS

I. A. Maslov¹, V. A. Grishin^{1,2}, I. V. Nikolenko³

¹ Space Research Institute of the Russian Academy of Sciences (IKI RAN)

² Bauman Moscow State Technical University (BMSTU)

³ Institute of Astronomy of the Russian Academy of Sciences (INASAN)

Currently, there is a significant interest in the research of navigation methods that have stability and reliability in terms of possible electronic interference. One of the possible sources of navigation information, which has a very high resistance to electronic jamming, is the classic method of navigation, based on the use of the visible position of the Sun, moon and stars. The navigation uses information about the exact time, as well as information about the direction of the local vertical, which can be obtained from the observed horizon line. However, variations in temperature, pressure and humidity of the atmosphere give rise to variations in the degree of refraction, which causes errors in determining the orientation of the local vertical. For aircraft, these errors are further increased due to the scattering and extinction of optical radiation in atmospheric aerosols. All this leads to the fact that the main source of navigation errors is errors in determining the local vertical. Some reduction of the influence of the radiation scattered by the atmosphere can be achieved by using polarization filters. In order to assess the prospects for the use of such filters to registration of the horizon line, as well as to assess the stability of atmospheric refraction, the following equipment was developed and put into operation. A distinctive feature

of the equipment is its autonomous operation, which usually does not require external control and with its switch-on immediately after the supply of power supply voltage.

Keywords: navigation, monitoring, horizon line, sea, atmosphere, automation, remote control

Maslov, Igor Alekseevich — head of sector, imaslov@iki.rssi.ru

Grishin Vladimir Aleksandrovich — senior scientist, PhD, associate professor, vgrishin@iki.rssi.ru

Nikolenko Igor Vladimirovich — junior researcher, niki63@yandex.ru

ДИНАМИКА КОСМИЧЕСКОГО ПОЛЁТА, ТЕОРЕТИЧЕСКАЯ МЕХАНИКА, УПРАВЛЕНИЕ ДВИЖЕНИЕМ

В. В. Ивашкин¹, Р. Р. Назиров², Н. А. Эйсмонт²

¹ Институт прикладной математики им. М. В. Келдыша РАН (ИПМ РАН)

² Институт космических исследований РАН (ИКИ РАН)

В статье обсуждаются вопросы формирования, эволюции, места в космонавтике механики космического полёта как науки. Анализируются основные изучаемые проблемы и связи с другими науками. Дается анализ новых современных задач. Отмечена важность задачи управления движением естественных небесных тел для современной механики космического полёта.

Ключевые слова: динамика космического полёта, теоретическая механика, управление движением

ВВЕДЕНИЕ

Мировая космонавтика и всё человечество отмечают в апреле годовщину первого пилотируемого космического полёта — полёт советского космонавта Юрия Гагарина 12 апреля 1961 года. Время и история показали, что этот полёт стал одним из важнейших шагов (и, вероятно, самым важным шагом) в развитии Космонавтики и, может быть, всей нашей Цивилизации. Он не был превзойдён другими большими космическими достижениями. Недаром Юрия Гагарина встречали с таким восхищением, с таким энтузиазмом во всех странах, все люди, независимо от их образования, богатства, социального статуса, славы, пола, возраста — как простые рабочие и крестьяне, так и принцы, принцессы, короли, королевы, как ученики и студенты, так и профессора и доценты, как мужчины и женщины, так и дети и старики. Все люди поняли, какие огромные возможности у них есть, все видели, что Земля едина, они должны жить дружно и объединять свои силы для достижения высоких человеческих целей. Образование, наука, современные технологии получили очень сильный импульс для быстрого развития. Личность Юрия Гагарина объединила на какое-то время весь мир. Связи между народами и странами стали более тесными, открытыми и дружескими. Действительно, это был звёздный час Человечества. Космонавтика получила мощный толчок и начала развиваться гигантскими шагами. Вскоре были реализованы и другие большие достижения: полёты на Венеру и Марс, автоматическая и пилотируемая мягкая посадка на Луну, полёты к дальним планетам Солнечной системы, технологическое использование околоземного пространства. Но все эти достижения не смогли

Ивашкин Вячеслав Васильевич — главный научный сотрудник, доктор физико-математических наук, профессор, ivashkin@keldysh.ru

Назиров Равиль Равильевич — заведующий отделом, доктор технических наук, rnazirov@cosmos.ru

Эйсмонт Натан Андреевич — ведущий научный сотрудник, кандидат технических наук, neismont@rssi.ru

затмить полёт Юрия Гагарина. Более того, огромная его заслуга заключается в том, что эти последующие гигантские достижения в значительной степени были обязаны одному этому полёту Юрия Гагарина.

1. ФОРМИРОВАНИЕ МЕХАНИКИ КОСМИЧЕСКОГО ПОЛЁТА

Отмечая годовщину исторического полёта Юрия Гагарина, мы должны не только напомнить себе об огромном его значении, но и использовать эту дату как стимул для анализа, осмысления, размышлений о развитии и современных проблемах космонавтики вообще и механики космического полёта (МКП), в частности. Это важно в первую очередь для тех, кто изучает проблемы космонавтики и, возможно, изучает механику космического полёта.

МКП является одной из основных частей космонавтики. Её называют и по-другому, например, динамика космического полёта, космическая баллистика, прикладная небесная механика и т.д. На основе законов природы, которые определяют движение тел в Солнечной системе, и знания характеристик технических средств, МКП позволяет выполнять необходимые вычисления, моделировать движение космического аппарата (КА) — как движение его центра масс (ЦМ) так и движение вокруг него, — оценить реальные параметры движения КА и найти (если необходимо) наведение и управление движением КА для приведения его в заданные условия. Это даёт научную возможность проанализировать характеристики проекта, понять, можно ли достичь целей миссии и найти наилучшие способы их достижения.

МКП начала формироваться вместе с первыми исследованиями пионеров космонавтики — Циолковского, Гансвиндта, Кондратюка, Годдарда, Оберта, Цандера, Гомана, Эсно-Пельтери, Штернфельда [Ivashkin, 1993; Кондратюк, 1938; Пионеры..., 1964, 1997; Цандер, 1924; Циолковский, 1903; Штернфельд, 1937; Hohmann, 1925; Sternfeld, 1934, 1956] и т.д. Дадим некоторые примеры. Основатель научной космонавтики К.Э. Циолковский [Пионеры..., 1964; Циолковский, 1903] доказал, что принцип движения струи можно использовать для полёта в космос, изучая реактивное влияние выходящей из двигателя струи газов на скорость тела в пространстве. В частности, он доказал и использовал известную формулу, которая теперь носит его имя:

$$\Delta \mathbf{V} = -c \ln(m_i/m_f),$$

здесь $\Delta \mathbf{V}$ — изменение вектора скорости; c — скорость струи истекающего из сопла двигателя выхлопного газа; m_f и m_i — конечная и начальная массы тела. Ю.В. Кондратюк [Кондратюк, 1938; Пионеры..., 1964] предложил управляемый вход КА в атмосферу планеты, оптимальную схему для пилотируемой миссии на Луну и отметил возможности гравитации для осуществления манёвра при облёте планеты. Ф.А. Цандер [Пионеры..., 1964; Цандер, 1924] также предложил и исследовал гравитационные манёвры, изучив влияние гравитационного поля на движение КА вблизи планеты. Он изучил также использование давления солнечной радиации для межпланетных полётов. Он и В. Гоман [Пионеры..., 1964, 1977; Цандер, 1924; Hohmann, 1925] независимо предложили двухимпульсный апсидальный перелёт между круговыми компланарными орбитами планет (этот перелёт теперь часто называют перелётом Гомана —

Цандера). А. Штернфельд [Штернфельд, 1937; Sternfeld, 1934, 1956] показал трёхимпульсный перелёт между такими орбитами. В результате пионеры космонавтики заложили основы для расчёта и анализа траекторий КА, а также для космической навигации и управления [Ivashkin, 1993]. Тем самым пионеры космонавтики заложили основы МКП.

Отметим, что их книги были учебниками, настольными книгами по МКП для инженеров и учёных, которые разрабатывали первые ракеты и КА в 1950-х годах для полётов в космос. Последующий быстрый прогресс космонавтики привёл к развитию и окончательному формированию МКП. В частности, мы думаем, что проблемы, решённые при создании первого искусственного спутника Земли и осуществлении первого полёта человека в космос, внесли основной вклад в создание МКП и сформировали каркас её современной структуры. К этим проблемам относятся: оптимальные траектории запуска ракет, автономная инерциальная навигация и управление, управляемая ориентация КА, наземная космическая навигация с использованием радиотехнических и оптических измерений, вход КА в атмосферу, коррекция траектории КА, автономная навигация КА, наземные центры управления космическими полётами и т. д. Кроме того, импульс, данный в то время, до сих пор двигает нашу науку и технику. В результате теперь у нас есть фантастические возможности МКП. Она позволяет, вместе с очень совершенными техническими средствами, выполнить проектирование миссий, наведение и управление движением при полёте КА для широчайшего диапазона задач — от близкого к Земле и к Солнцу пространства Солнечной системы, до дальних её областей, чтобы изучить неизвестное ранее строение Солнечной системы, понять биологические проблемы возникновения жизни, проверить основные законы природы.

Ниже мы дадим краткий методологический анализ основных проблем современной механики космического полёта.

2. КЛАССИФИКАЦИЯ ПРОБЛЕМ МЕХАНИКИ КОСМИЧЕСКОГО ПОЛЁТА

В настоящее время существует много задач, которые решает и исследует МКП (см., например, [Ивашкин, 1993; Основы..., 1972; Охочимский, Сихарулидзе, 1990; Эльясберг, 1965; Battin, 1964; Chobotov, 1991; European..., 1998]). Конечно, невозможно представить их все подробно в данном кратком очерке. Попробуем дать здесь их краткое описание. В таблице показаны в матричном виде основные проблемы МКП.

Здесь три строки представляют три основных предмета исследований МКП:

- 1) движение центра масс (ЦМ) КА — траектория КА $\mathbf{r}(t)$, $\mathbf{V}(t)$;
- 2) движение КА вокруг его ЦМ — ориентация КА $\psi(t)$; $\theta(t)$; $\varphi(t)$; $\omega_x(t)$; $\omega_y(t)$; $\omega_z(t)$;
- 3) физическая среда, определяющая движение и активность КА: характеристики КА и его систем, окружающей среды, в частности, небесных тел, которые влияют на движение КА, а также на работу автономных и наземных систем.

Три колонки представляют три основные проблемы, которые должны быть решены для каждого предмета анализа:

- 1) моделирование движения по известным параметрам, определяющим это движение;
- 2) оценки этих параметров на основе проводимых измерений;
- 3) управление, а также его оптимизация.

Классификация проблем механики космического полёта

	Моделирование движения	Оценка параметров состояния	Управление, его оптимизация
Траектория КА	Расчёт траектории КА	Оценка параметров траектории КА	Управление движением центра масс КА, его оптимизация
Ориентация КА	Расчёт ориентации КА	Оценка параметров ориентации КА	Управление ориентацией КА, его оптимизация
Среда	Моделирование среды и её эволюции	Оценка параметров среды	Проектирование миссии, КА и его систем. Управление движением естественных небесных тел, его оптимизация

Кроме того, мы имеем также ось в перпендикулярном направлении. По третьей оси мы можем дать реальные проекты — от исследований пионеров космонавтики до современных проектов, таких как, например, проекты полётов на астероиды и кометы, проекты снижения астероидно-кометной опасности для Земли. Это, может быть, самая интересная ось. И, наконец, надо помнить, что обычно все эти проблемы изучаются для двух вариантов:

- 1) в ходе разработки КА и его траекторий;
- 2) во время полёта КА, управления его полётом.

Таким образом, можно рассматривать две трёхмерные матрицы.

Все указанные проблемы тесно связаны друг с другом. Кроме того, здесь показаны только основные проблемы. Данная классификация, конечно, является приблизительной и условной. С учётом этого мы ниже кратко дадим методологический анализ этих проблем и отметим их связь с другими науками и отраслями современной техники.

Расчёт движения ЦМ (траектории) КА, вероятно, является центральной задачей МКП. Одновременно это классическая задача динамики материальной точки в теоретической механике и небесной механике. Законы движения планет Кеплера [Kepler, 1609, 1619], закон всемирного тяготения Ньютона и его постулаты и решения задачи динамики материальной точки [Newton, 1687] стали важнейшими элементами физики, теоретической и небесной механики, а также заложили основу и дали методы решения этой задачи в МКП. Анализ орбит КА в модели Кеплера, где они рассматриваются как конические сечения в рамках задачи двух тел, является в настоящее время основным методом качественного анализа траекторий КА, особенно при проектировании КА. Но, с другой стороны, современные проблемы МКП и её возможности далеко ушли от этих элементарных моделей. Используя классические методы механики и новые разработанные, МКП должна рассчитать движение КА с необходимой точностью, которая часто должна быть необычно высокой. Многие возмущения для этого — от притяжения других небесных тел, от несферичной структуры планет, атмосферного сопротивления, давления солнечного света,

эффектов общей теории относительности, даже кометной активности и т. д. — нужно принимать во внимание. Тем не менее, следует отметить, что возможности классического метода конических сечений очень развиты за счёт ряда факторов, например, метода сфер действия Лапласа, метода Энке точного определения орбиты с использованием малой поправки к Кеплеровской орбите, метода оскулирующих элементов Лагранжа [Абалакин и др., 1971; Основы..., 1972; Охоцимский, Сихарулидзе, 1990; Эльясберг, 1965; Battin, 1964; Chobotov, 1991 и др.]. Важным отличием задач МКП от задач классической механики является учёт реактивной тяги ракетных двигателей.

Требования оперативности расчёта орбиты КА во время его полёта, а также компактности бортового алгоритма расчёта орбиты КА при высокой точности расчёта также являются очень важными новыми факторами [Draper et al., 1965; Ivashkin, 1993]. Это приводит к высокой значимости электронных компьютеров, вычислительной математики и компьютерных наук (информатики) для МКП. С другой стороны, эти требования (наряду с требованиями атомных технологий) стали важнейшими причинами и стимулами необычайно быстрого развития электронной вычислительной техники и науки.

Расчёт движения КА вокруг центра масс — *ориентация КА* — является второй основной проблемой МКП. Теоретической основой её решения становятся соответствующие результаты теоретической механики, в частности, теорема об изменении кинетического момента для механической системы, а также динамические и кинематические уравнения Л. Эйлера [La Vallée Poussin, 1925]. Космонавтика поставила новые задачи, для их решения наряду с использованием классических методов, МКП выработала новые подходы, что позволило получить важные результаты. Так, для нахождения ориентации КА с хорошей точностью должны быть учтены некоторые возмущающие моменты (гравитационные, аэродинамические, магнитные, из-за давления солнечного света и др.) и моменты реактивных двигателей. Проблему зачастую осложняет сложная структура КА, её отличие от твёрдого тела, например, наличие полостей с жидкостью (топливом), гибкость конструкции (за счёт солнечных панелей), большие сложные составные конструкции, как тросы в МКС и т. д.), движущиеся массы (как в пилотируемых КА). В результате появилась новая наука, которая сама обогатила механику, например, см. [Белецкий, 1965; Сарычев, 1978].

Моделирование среды является ещё одним важным элементом МКП. Необходимо моделировать силовые поля, влияющие на движение КА, наземные и бортовые измерения, работу некоторых систем КА (ориентация, навигация, реактивные двигатели, например), космические эксперименты и т. д. Что касается моделей движения естественных небесных тел, то мы имеем классическую задачу небесной механики. Но астронавтика поставила ряд новых, очень сильных требований к точностям знания движения этих небесных тел. Это привело к разработке новых теорий движения планет и их спутников, астероидов и комет.

Реальное движение КА всегда отличается от расчётного из-за ошибок работы реактивных двигателей, а также ошибок знаний силовых полей. Поэтому навигационная задача — выполнить измерения и *оценить параметры движения КА* в соответствии с проведёнными измерениями — тоже очень важна. Во-первых, это проблема оценки элементов орбиты КА. Астрономы давно решали эту задачу, оценивая элементы орбит планет и их спутников, астероидов, комет по данным наземных оптических наблюдений. А. М. Лежандр и К. Ф. Гаусс поставили эту задачу на научную основу, разработав метод наименьших квадратов

для статистической обработки данных измерений. Этот метод стал основным статистическим методом космической навигации и основой для разработки новых методов. В этих навигационных методах используются, в частности, разность между наблюдаемыми и расчётными значениями измеряемых параметров. Поэтому для решения навигационной задачи необходимо выполнить моделирование движения КА. В отличие от классической астрономической навигации, в настоящее время основными измерениями для наземной космической навигации являются радиолокационные данные по дальности от наземной станции до КА и доплеровской скорости, а также углы, определяющие направление от этой станции к КА. Всё более значимой становится автономная навигация КА. Она использует бортовые (оптические и/или радио-) измерения и их последующую (бортовую или наземную) обработку. Бортовые инерциальные гироскопические системы разрабатываются для навигации и управления на активных участках при работе ракетных двигателей и при входе в атмосферу. Они измеряют ускорение КА и интегрируют его, получают скорость КА и положение. Эта теория навигации основана на результатах механики в задаче движения системы тел. Космонавтика предъявляла высокие требования к точности, надёжности, скорости работы и (для автономной навигации) реализуемости алгоритмов на малых бортовых компьютерах. Это практически привело к развитию новой научной отрасли — космической навигации параллельно с использованием обычных результатов физики, астрономии, математической статистики и информатики [Основы..., 1972; Ivashkin, 1993; Battin, 1964; Draper et al., 1965; Akim, Eneev, 1963 и др.].

Аналогичная ситуация имеет место и с задачей оценки параметров ориентации и вращения КА по данным измерений. Астрономия также изучает эту проблему, анализируя вращение естественных небесных тел по некоторым наземным наблюдениям. Но не всегда классический астрономический подход является хорошим для решения задач космонавтики, где требуется быстрая и высокоточная оценка ориентации и вращения КА. Для этой цели обычно используются данные бортовых датчиков (например, оптических). Они позволяют напрямую связать бортовую систему координат с инерциальной системой. Для решения этой навигационной задачи требуются эффективные методы статистической обработки данных измерений, моделирования движения КА, моделирования среды (в частности, характеристик бортовых датчиков).

Необходимость высокой точности для моделирования движения КА и измерений требует моделирования среды с соответствующей высокой точностью как для систем КА, так и для внешнего окружения. Это приводит к проблеме **оценки параметров среды** по некоторым наземным и бортовым измерениям. Такая проблема важна не только для космонавтики, но и для фундаментальных наук, поскольку позволяет получать фундаментальную информацию о планетах, всей Солнечной системе и Вселенной. Так, хорошо известно, что первые полёты спутников позволили существенно скорректировать наши знания о земной атмосфере. В проекте ФОБОС ставилась задача более точно определить некоторые параметры Солнечной системы, а также изучить постоянство универсальной гравитационной постоянной на основе оптических и радиоизмерений, выполненных на поверхности Фобоса. Проекты ВЕГА, ДЖОТТО, ROSETTA дали богатую информацию о кометном ядре и характеристиках его активности. Оценка параметров среды может осуществляться как совместно с оценкой параметров движения КА (путём расширения вектора оцениваемого состояния), так и раздельно.

Подготовка данных для **наведения и управления** движением КА, вероятно, является наиболее важной проблемой МКП с точки зрения практической космонавтики. Так, после определения реального движения КА мы должны рассчитать необходимую коррекцию траектории и ориентации КА в соответствии с целями и возможностями проекта. Обычно это осуществляется путём включения ракетных двигателей КА, которые расходуют топливо во время их работы. Эти управляемые изменения движения КА обычно называют манёврами. Они должны быть выполнены с минимальными затратами топлива, потому что оно на борту КА, как правило, в остром дефиците. Так у нас встаёт задача оптимизации. Из-за этого, в частности, очень важны для космических полётов гравитационные манёвры, так как они позволяют изменять орбиту КА с помощью гравитационного поля небесного тела. Для снижения затрат топлива всё чаще применяются электрореактивные двигатели. Для наведения и управления движением КА механика космического полёта, конечно, использует методы и результаты общей теории управления и теории оптимизации. С другой стороны, проблемы управления КА стимулировали развитие современной теории управления и теории оптимизации в силу их важности, актуальности, поэтому были разработаны и новые, очень эффективные методы [Гродзовский и др., 1966; Ивашкин, 1975; Ильин, Кузмак, 1976; Константинов и др., 2015; Магес, 1979; Pontryagin et al., 1962 и др.].

Следует также отметить, что объединение процесса управления движением КА и процесса измерений и оценки реальных параметров движения КА, а также параметров внешней среды позволяет разработать адаптивную систему управления космическим полётом [European..., 1998].

И, наконец, кратко остановимся на последнем элементе матрицы. Здесь есть два момента.

Во-первых, МКП играет важную роль в **разработке космического проекта**: КА, схемы его полёта, траектории, его основных систем (например, маневрирования, навигации, ориентации), в оптимизации их характеристик. Это, вероятно, главная проблема МКП, потому что хороший проект миссии и КА является базой, необходимым условием успеха миссии.

Во-вторых, может быть, это кажется странным и неожиданным на первый взгляд, но **проблема управления движением естественных небесных тел** была в последнее время поставлена на передний план перед МКП. Было показано, что вероятность столкновения Земли с астероидами и кометами не так уж и мала. Эти столкновения могут вызвать земную катастрофу [Morrison, 1992]. Поэтому перед наукой и техникой стоит сейчас проблема предотвращения или уменьшения Астероидно-Кометной опасности. В связи с этим важно изучение ряда проблем, в частности: обнаружение опасного небесного тела, которое может столкнуться с Землёй, и осуществление управляемого воздействия КА на орбиту опасного небесного тела для предотвращения этого столкновения [Данхэм и др., 2013а, б; Ледков и др., 2015а, б; Эйсмонт и др., 2013, 2015; Ahrens, Harris, 1992; Ivashkin, Smirnov, 1995; Ivashkin et al., 1997, 1999; Melosh, Nemchinov, 1993; Morrison, 1992]. Эти проблемы требуют комплексного анализа ряда задач, в частности: создание системы мониторинга, поиск и анализ эффективных воздействий на опасное небесное тело и его орбиту, моделирование движения КА и естественного небесного тела, навигация и наведение, проектирование и создание специальной научно-технической системы. Это может быть важным также для увода небесного тела на специальную орбиту с целью, например, использования его минеральных

и энергетических ресурсов. МКП должна внести важный вклад в решение этих проблем.

Основная идея предлагаемого метода состоит в выполнении следующих операций. Сначала космический аппарат совершает посадку на выбранный относительно небольшой астероид и закрепляется на нём. Далее с помощью двигателя аппарата астероид переводится на орбиту гравитационного манёвра у Земли или другой планеты таким образом, чтобы в результате этого манёвра астероид-снаряд оказался на орбите столкновения с опасным небесным телом. Последующее соударение этих тел изменяет первоначальную траекторию опасного небесного объекта, что позволяет избежать его встречи с Землёй. Проведённые исследования показали принципиальную возможность выполнения таких операций. При этом моделировалось управление реальными астероидами в рамках реальных технических возможностей современной ракетно-космической техники [Данхэм и др., 2013а, б; Эйсмонт и др., 2013, 2015; Ahrens, Harris, 1992; Ivashkin, Smirnov, 1995; Ivashkin et al., 1997, 1999; Melosh, Nemchinov, 1993; Morrison, 1992; Ледков и др., 2015а, б]. Были также подтверждены возможности управления астероидами с помощью аналогичной технологии для перевода астероидов на новые орбиты, более удобные для исследований, включая захват на орбиты спутников Земли. В последнем случае в сочетании с облётом Земли использовались гравитационные манёвры у Луны.

Из этого краткого обзора и анализа можно сделать вывод, что МКП, будучи одной из наиболее важных частей космонавтики, в настоящее время является практически самостоятельной наукой. Она вобрала в себя методы и результаты целого комплекса наук, к которым можно отнести: механику (теоретическая, небесная, прикладная механика), астрономию, физику, навигацию, математическую статистику, теорию и методы наведения и управления, теорию оптимизации, информатику (вычислительная математика, компьютерная наука и техника, теория информации), машиностроение. Эти научные связи условно показаны на рисунке [Ivashkin, 1994, 1999, 2003]. И обратно, МКП, со своей стороны, оказала плодотворное влияние на развитие этих наук.

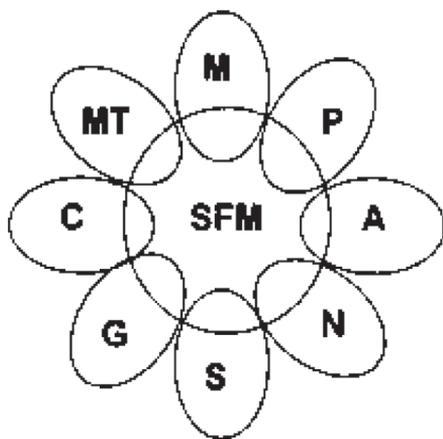


Схема научных связей механики космического полёта: SFM — механика космического полёта; М — теоретическая, прикладная, небесная механика; Р — физика; А — астрономия; N — навигация; S — математическая статистика; G — управление, оптимизация; С — вычислительная математика, компьютерные науки; МТ — машиностроение

3. НЕКОТОРЫЕ СОВРЕМЕННЫЕ ПРОБЛЕМЫ МЕХАНИКИ КОСМИЧЕСКОГО ПОЛЁТА

Только 60 лет прошло с запуска первого Спутника, а космонавтика уже достигла необычайного расцвета. МКП, развиваясь параллельно, добилась также очень высоких результатов. Изобилие и сложность решаемых ею проблем поражают и удивляют. С развитием человечества, науки и космонавтики, всё новые и новые проблемы постоянно ставятся перед космонавтикой и МКП, соответственно. Не претендуя на полноту, мы отметим ниже некоторые такие проблемы, стоящие реально сейчас перед нами, важные и интересные с нашей личной точки зрения:

1. Астероидно-кометная опасность для Земли.
2. Исследование Луны, лунные станции, точки либрации.
3. Экспедиции Земля – небесное тело – Земля.
4. Исследование астероидов, комет, спутников планет.
5. Изучение внешних областей Солнечной системы.
6. Миссии с новыми источниками энергии.
7. Малые КА.
8. Большие космические конструкции.
9. Гравитационные и аэродинамические манёвры.
10. Проблемы навигации.
11. Управление движением естественных небесных тел.
12. Поддержка новых научных экспериментов, решение вопросов фундаментальных наук, проверка законов природы.
13. Изучение происхождения и эволюции жизни.
14. Предотвращение засорения околоземного космического пространства.

Последний пункт приобретает всё большее значение. К настоящему времени в околоземном пространстве насчитывается более 20 000 космических объектов искусственного происхождения — космический мусор. Это отработавшие космические аппараты и ракетные ступени, а также их фрагменты с размерами более 10 см, орбиты которых определены. Иначе говоря, это каталогизированные объекты. Объектов с меньшими размерами по оценкам более 100 000.

Актуальность решения проблемы не вызывает сомнений. Однако практически реализуемых технологий для удаления накопившегося мусора пока не предложено, особенно с учётом стоимости таких операций. Пока что речь идёт о предотвращении дальнейшего засорения, и роль МКП здесь исключительно значима.

Относительно этого списка сделаем два замечания. Методически эти проблемы соответствуют третьей оси с проектами в нашей матрице, приведённой выше. И, конечно, порядок в этом списке не имеет отношения к приоритетности или важности проблем.

В заключение хотелось бы ещё раз напомнить, что, на наш взгляд, большие достижения МКП и всей космонавтики стали возможными в значительной степени благодаря первым успешным космическим полётам, особенно благодаря полёту первого Спутника и полёту Ю.А. Гагарина. И мы должны оправдать надежды, возлагавшиеся на космонавтику в то время.

Авторы признательны аспиранту Гуо Пэн за помощь в подготовке статьи.

ЛИТЕРАТУРА

- [Абалакин и др., 1971] *Абалакин В. К., Аксенов Е. П., Гребеников Е. А., Рябов Ю. А.* Справочное руководство по небесной механике и астродинамике. М.: Наука, 1971. 584 с.
- [Белецкий, 1965] *Белецкий В. В.* Движение искусственного спутника относительно центра масс / ред. Абашева Д. А. М.: Наука, 1965. 415 с.
- [Гродзовский и др., 1966] *Гродзовский Г. Л., Иванов Ю. Н., Токарев В. В.* Механика космического полета с малой тягой. М.: Наука, 1966. 679 с.
- [Данхэм и др., 2013а] *Данхэм Д. У., Ретсема Х. Д., Лу Э., Аренци Р., Линфилд Р., Чапман К., Фаркуар Р., Ледков А., Эйсмонт Н., Чумаченко Е.* Метод предупреждения столкновения малых астероидов с Землёй // *Астроном. вестн.* 2013. Т. 47. № 4. С. 341–351.
- [Данхэм и др., 2013б] *Данхэм Д. У., Назиров Р. Р., Фаркуар Р. У., Чумаченко Е. Н., Эйсмонт Н. А., Симонов А. В.* Космические миссии и планетарная защита. М.: Физматлит, 2013. 276 с.
- [Ивашкин, 1975] *Ивашкин В. В.* Оптимальное космическое маневрирование при свободном времени перехода и ограничениях на расстоянии до планеты: препринт. М.: ИПМ, 1975. № 40.
- [Ивашкин, 1993] *Ивашкин В. В.* Механика космического полёта: препринт. М.: ИПМ, 1993. № 19. 14 с.
- [Ильин, Кузмак, 1976] *Ильин В. А., Кузмак Г. Е.* Оптимальные перелеты космических аппаратов. М.: Наука, 1976. 744 с.
- [Кондратюк, 1938] *Кондратюк Ю. В.* Тем, кто будет читать, чтобы строить // *Пионеры ракетной техники: соч. 1918–1919; публ. 1938 / под ред. Т. М. Мелькумова; Институт истории естествознания и техники АН СССР.* М.: Наука, 1964.
- [Константинов и др., 2015] *Константинов М. С., Петухов В. Г., Тейн М.* Оптимизация траекторий гелиоцентрических перелетов. М.: Изд-во МАИ, 2015. 260 с.
- [Ледков и др., 2015а] *Ледков А. А., Эйсмонт Н. А., Назиров Р. Р., Боярский М. Н.* Способ захвата астероидов на орбиту спутника Земли // *Письма в Астроном. журн.* 2015. Т. 41. № 1–2. С. 72–89.
- [Ледков и др., 2015б] *Ледков А. А., Эйсмонт Н. А., Боярский М. Н., Федяев К. С., Назиров Р. Р.* Управление движением околоземных астероидов // *Письма Астроном. журн.* 2015. Т. 41. № 8. С. 482–489.
- [Основы..., 1972] *Основы теории полета космических аппаратов / под ред. Г. С. Нариманова, М. К. Тихонравова.* М.: Машиностроение, 1972. 607 с.
- [Охоцимский, Сихарулидзе, 1990] *Охоцимский Д. Е., Сихарулидзе Ю. Г.* Основы механики космического полёта. М.: Наука, 1990. 448 с.
- [Пионеры..., 1964] *Пионеры ракетной техники: Кибальчич, Циолковский, Цандер, Кондратюк. Избранные труды (1891–1938).* М.: Наука, 1964. 670 с.
- [Пионеры..., 1977] *Пионеры ракетной техники. Гансвиндт. Годдард. Эсно-Пельтри. Оберт. Гоман: Избранные труды (1891–1938).* М.: Наука, 1977. 632 с.
- [Сарычев, 1978] *Сарычев В. А.* Вопросы ориентации искусственных спутников. Итоги науки и техники. Сер.: Исследование космического пространства, М.: ВИНТИ, 1978. Т. 11.
- [Цандер, 1924] *Цандер Ф. А.* Перелеты на другие планеты // *Техника и жизнь.* 1924. № 13. С. 15–16.
- [Циолковский, 1903] *Циолковский К. Е.* Исследование Вселенной с ракетно-двигательными аппаратами // *Научное обозрение.* 1903. Т. 5.
- [Штернфельд, 1937] *Штернфельд А. А.* Введение в космонавтику. М.; Л.: изд-во ОНТИ, 1937. 318 с.
- [Эйсмонт и др., 2013] *Эйсмонт Н. А., Боярский М. Н., Ледков А. А., Назиров Р. Р., Данхэм Д., Шустов Б. М.* О возможности наведения малых астероидов на опасные небесные объекты с использованием гравитационного манёвра // *Астроном. вестн.* 2013. Т. 47. № 4. С. 352–360.

- [Эйсмонт и др., 2015] *Эйсмонт Н. А., Ледков А. А., Назиров Р. Р.* Укroшение астероидов: задачи и способы управления движением малых небесных тел // Природа. 2015. № 9. С. 57–65.
- [Эльясберг, 1965] *Эльясберг П. Е.* Введение в теорию полета искусственных спутников Земли. М.: Наука, 1965. 540 с.
- [Ahrens, Harris, 1992] *Ahrens T. J., Harris A. W.* Deflection and fragmentation of near-earth asteroids // Nature. 1992. V. 360. P. 429–433.
- [Akim, Eneev, 1963] *Akim E. L., Eneev T. M.* Determination of space vehicle's motion parameters by trajectory measurements data // Cosmic Research. 1963. No. 1. P. 5–55.
- [Battin, 1964] *Battin R. H.* Astronautical Guidance. N. Y., McGraw-Hill Book Company, 1964.
- [Chobotov, 1991] *Chobotov V. A.* Orbital mechanics. AIAA Education Series. 1991.
- [Draper et al., 1965] *Draper C. S., Wrigly W., Hoag D. G., Battin R.* et al. Space Navigation, Guidance and Control / ed. J. E. Miller, W. Mackey, J. Mackey and Co. Ltd. London, 1965.
- [European..., 1998] European Space Operations Center. Flight Dynamics. N7, ESA f-15. 1998.
- [Hohmann, 1925] *Hohmann W.* Die Erreichbarkeit der Himmelskörper. Untersuchungen über das Raumfahrtproblem. München, 1925.
- [Ivashkin, 1993] *Ivashkin V. V.* On the history of space navigation development // 44th Congress of the Intern. Astronautical Federation. Oct. 16–22, 1993, Graz, Austria. 1993. Paper IAA.
- [Ivashkin, 1999] *Ivashkin V. V.* Análisis de Problemas en Mecánica Espacial, Ingeniería Aeronáutica y Astronautica // Revista de la Asociación de Ingenieros Aeronáuticos. Madrid, Spain. 1999. No. 357/4-1999. P. 40–46 (in Spanish).
- [Ivashkin, 2003] *Ivashkin V. V.* Analysis of space flight mechanics problems // Acta Astronautica. 2003. V. 52. P. 663–670.
- [Ivashkin, Smirnov, 1995] *Ivashkin V. V., Smirnov V. V.* An analysis of some methods of asteroid hazard mitigation for the earth // Planetary and Space Science. 1995. V. 43. Iss. 6. P. 821–825.
- [Ivashkin et al., 1997] *Ivashkin V. V., Chernov A. V., Zaytsev A. V.* Optimal flights to near-earth asteroid for impact-kinetic influence // 12th Intern. Symp. Space Flight Dynamics, ESOC. Darmstadt, Germany, 2–6 June 1997: proc. 1997. SP-403.
- [Ivashkin et al., 1999] *Ivashkin V. V., Chernov A. V., Zaytsev A. V.* Optimal flights to near-earth asteroid // Acta Astronautica. 1999. V. 44. Iss. 5. P. 219–225.
- [Kepler, 1609] *Kepler J.* Astronomia Nova Aitiologhtos, seu Physica Coelestis. Heidelberg, 1609.
- [Kepler, 1619] *Kepler J.* Harmonices Mundi. Linc, 1619.
- [La Vallée Poussin, 1925] *La Vallée Poussin Ch.-J. de.* Leçons de mécanique analytique. Louvain: Librairie universitaire, 1925.
- [Marec, 1979] *Marec J.-P.* Optimal Space Trajectories. Amsterdam, Oxford, New York: Elsevier Science Publishing Company, 1979.
- [Melosh, Nemchinov, 1993] *Melosh M. J., Nemchinov I. V.* Solar asteroid diversion // Nature. 1993. V. 366. P. 21–22.
- [Morrison, 1992] *Morrison D.* The Spaceguard Survey: Report of the NASA International Near-Earth-Object Detection Workshop, January 25, 1992. 64 p.
- [Newton, 1687] *Newton I.* Philosophiae Naturalis Principia Mathematica. London, 1687.
- [Pontryagin et al., 1962] *Pontryagin L. S., Boltyanskii V. G., Gamkrelidze R. V., Mischenko E. F.* The Mathematical Theory of Optimal Processes. N. Y.: Wiley, Inter-science Publishers, 1962.
- [Sternfeld, 1934] *Sternfeld A. Ja.* Comptes rendus de l'Académie des sciences. 1934. V. 198. P. 711–713.
- [Sternfeld, 1956] *Sternfeld A. Ja.* Artificial Satellites of the Earth, GITTL, Moscow, 1956.

SPACE FLIGHT DYNAMICS FROM GAGARIN FIRST ORBIT TO CONTROL MOTION OF NATURAL SKY OBJECTS.

V. V. Ivashkin¹, R. R. Nazirov², N. A. Eismont²

¹ Keldysh Institute of Applied Mathematics of Russian Academy of Sciences

² Space Research Institute of Russian Academy of Sciences

The questions of forming, evolution of space flight dynamics as a science and its role in cosmonautics at large are discussed. Principal problems of studies are explored connected with other natural sciences. Analysis of new tasks generated by contemporary space projects is presented. Importance of the problems of natural objects motion control problems is underlined.

Keywords: space flight dynamics, theoretical mechanics, motion control

Ivashkin Vyacheslav Vasilievich — chief scientist, doctor of physical and mathematical sciences, professor, ivashkin@keldysh.ru

Nazirov Ravil Ravilyevich — head of department, doctor of technical sciences, rnazirov@cosmos.ru

Eismont Natan Andreevich — leading scientist, PhD, neismont@rssi.ru

ЯДЕРНАЯ ДВИГАТЕЛЬНАЯ УСТАНОВКА С ЗАМКНУТЫМ ЭНЕРГОБЛОКОМ

В. С. Буриков, И. В. Буриков, Ю. В. Буриков

ООО АЛЬТИКЭН «Альтернативная и криогенная энергетика»

В докладе рассмотрена схема ракетного двигателя с ядерным источником энергии. Двигатель включает в себя замкнутый энергетический контур с ядерным источником тепла. В контуре вырабатывается механическая энергия, а не использованное тепло отводится к рабочему телу двигателя, находящемуся на космическом аппарате. Рабочее тело после испарения и предварительного нагрева дополнительно подогревается в ядерном источнике энергии, сжимается в компрессоре и с помощью газодинамического сопла создаёт тягу двигателя. Компрессор осуществляет сжатие рабочего тела за счёт работы, полученной в замкнутом контуре.

Использование в качестве рабочего тела жидкого водорода позволяет получить большую тягу двигателя и высокий удельный импульс.

Ключевые слова: ракетный двигатель, ядерный источник энергии, замкнутый энергетический контур, жидкий водород

В настоящее время на повестку дня вновь встал вопрос о возможности и необходимости возрождения работ по созданию ядерных ракетных двигателей (ЯРД). Это связано с тем, что дальнейшее освоение космического пространства затруднено при использовании химического топлива. Энергетические ограничения не позволяют получить удельный импульс жидкостного ракетного двигателя (ЖРД) более 450 с, что приводит к неоправданному росту массы разгонных блоков ракет для достижения характеристических скоростей, необходимых для осуществления космических полётов. В большей степени это относится к пилотируемым полётам, что связано с временным фактором при осуществлении космической экспедиции, например, на Марс или астероиды.

При создании ЯРД определились два направления — двигатели с прямым нагревом рабочего тела и электрореактивные (с предварительным преобразованием тепловой энергии реактора в электрическую). Первая схема реализована в двигателе КБХА РД0410 [www.kbkha.ru]. Примерная схема двигателя представлена на рис. 1 (см. с. 144). Там же для сравнения приведена схема ЖРД открытой схемы.

Основные характеристики двигателя РД 0410 (рис. 2, см. с. 144) (www.kbkha.ru):

- тяга в пустоте — 3,59 т (35,2 кН)
- тепловая мощность реактора — 196 МВт;
- удельный импульс тяги (в пустоте) — 910 с;
- ресурс работы (до 10 включений) — 3600 с;
- масса с радиационной защитой — 2000 кг;
- высота — 3500 мм;
- диаметр — 600 мм.

Буриков Юрий Вячеславович — директор, alticen@mail.ru

Буриков Игорь Вячеславович — ведущий инженер

Буриков Владислав Сергеевич — научный консультант, кандидат технических наук

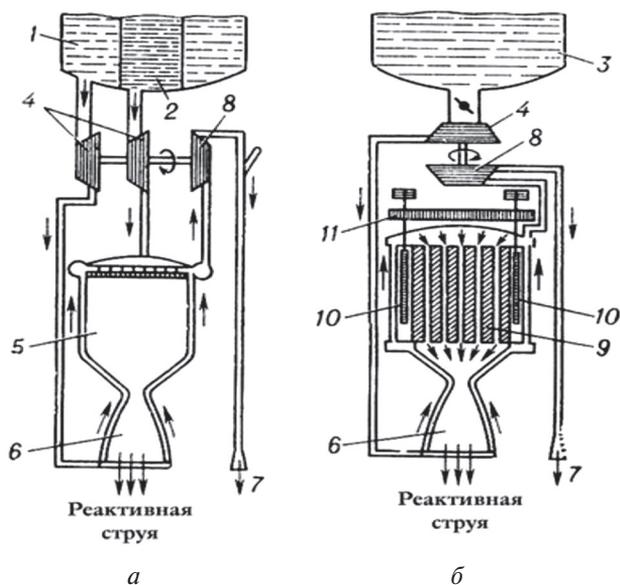


Рис. 1. Схема ЖРД (*a*): 1 — бак горючего; 2 — бак окислителя; 4 — насосы окислителя и горючего; 5 — камера сгорания; 6 — газодинамическое сопло; 7 — выхлоп турбины; 8 — турбина привода насосов. Схема прямогочного ЯРД (*б*): 3 — бак жидкого водорода; 4 — насос; 6 — газодинамическое сопло; 7 — выхлоп турбины; 8 — турбина привода насоса; 9 — ТВЭЛы; 10 — регулирующие стержни; 11 — корпус системы регулирования



Рис. 2. Двигатель РД 0410

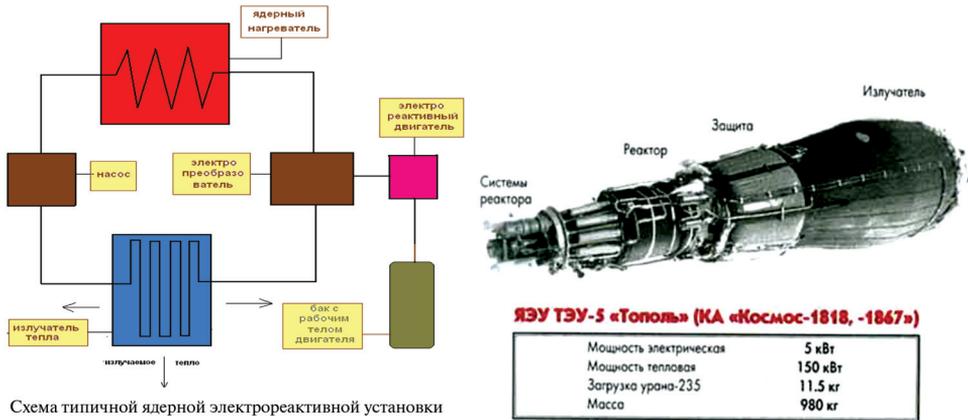


Рис. 3

На рис. 3 представлена типичная схема ЯЭРД (ядерный электрореактивный двигатель), и космический аппарат «Тополь» с подобной двигательной установкой (некоторые его технические характеристики).

В настоящее время в РФ создаётся энергодвигательная установка мегаваттного класса с использованием ядерного реактора. Ожидаемая электрическая мощность на борту космического аппарата (КА) — 1 МВт, удельный импульс ЭРД до 7000 с.

Примерные характеристики КА и ДУ представлены на сайте www.roznovatelno.ru (Ядерные энергетические и энергодвигательные установки. Космос. Невероятно, но факт).

В РНЦ «Курчатовский институт» разработан проект ЯЭУ с турбомашинным преобразователем энергии, рассчитанной на следующие параметры:

- тепловая мощность — до 10 МВт;
- электрическая мощность — около 2 МВт;
- система преобразования энергии — турбомашинная (цикл Брайтона);
- суммарное время работы — не менее 104 ч;
- число включений за ресурс — до 30;
- максимальная температура рабочего тела — до 1500 К.

В результате проведённых проработок космического аппарата (КА) определены основные проектные характеристики ЯЭУ:

- масса газоохлаждаемого реактора — 1000 кг;
- топливо — UC (U, Zr) C, UN с 90%-м обогащением по U235, оболочка топлива — Zr, W184, отражатель — Be;
- масса радиационной защиты (LiH, W, B4C) — 1000 кг;
- масса преобразователя энергии (турбина, компрессор и униполярный генератор) — 3500 кг;
- рабочее тело — гелий-ксеноновая смесь (1–3 % Xe);
- холодильники-излучатели на тепловых трубах при средней температуре около 700 К, масса — 3000 кг;
- площадь холодильника-излучателя (эффективная) — около 300 м²;
- масса системы автоматического управления, системы электроснабжения — 1000 кг;

- масса конструкции ЯЭУ — 1500 кг;
- суммарная масса ЯЭУ — 11 000 кг;
- удельная масса — 5,5 кг/кВт (11 000 кг/2000 кВт).

При сравнении двух схем ЯРД очевидно, что схема с прямым нагревом рабочего тела двигателя в реакторе в энергетическом и весовом плане предпочтительнее, удельная масса (2000 кг/196 000 кВт) 0,01 кг/кВт. Основной недостаток — удельный импульс ниже 1000 с, ресурс работы не более 1 ч, что недостаточно для осуществления межпланетных перелётов. Небольшой ресурс двигателя обусловлен высокой температурой в рабочей зоне ядерного реактора (более 2500 К).

В случае использования ЯЭУ с электрореактивным двигателем температура в рабочей зоне ядерного реактора близка к 1500 К за счёт чего ресурс установки возрастает до 1000 ч и более. Ожидаемая тяга ЭРД (при КПД 50 % и скорости истечения струи рабочего тела до 70 км/с) порядка 18 Н (1,8 кг), т. е. она в 2000 раз меньше, чем в двигателе РД0410. При осуществлении дальних пилотируемых полётов этот фактор может оказаться определяющим.

В процессе проведения научно-исследовательской работы по созданию эффективных экологически чистых двигателей в ООО «АЛЬТИКЭН — альтернативная и криогенная энергетика» длительное время исследовалась схема двигательной установки, использующей бинарный термодинамический процесс, в котором один процесс замкнутый, а второй открытый (Патент РФ № 2116460 от 6.02.1997). В такой термодинамической системе удаётся наиболее эффективно использовать тепло источника энергии — нагревателя для получения полезной работы. Особенно ярко проявляются достоинства рассматриваемой схемы в криогенном двигателе. В этом случае отводимое из замкнутого контура тепло после получения в нём полезной работы не выбрасывается в окружающую среду, а поглощается криогенной жидкостью для её испарения и предварительного нагрева. Пары криогенной жидкости являются рабочим телом открытого термодинамического контура, в котором после нагрева и расширения возникает дополнительная полезная работа. Использование в замкнутом контуре в качестве поглотителя тепла низкотемпературной жидкости позволяет получить высокий термический коэффициент преобразования низкотемпературного тепла нагревателя в полезную работу.

По нашему мнению эта схема двигательной установки может быть реализована и в космическом реактивном двигателе. При этом она может оказаться эффективной, если в качестве источника энергии используется ядерный реактор, а в качестве рабочего тела двигателя — жидкий водород (Патент РФ № 2447311 от 17.09.2008). Двигатель, который может быть выполнен по предлагаемой схеме, представлен на рис. 4 (см. с. 147).

Он состоит из источника тепла (1), нагревателя рабочего тела замкнутого контура (2), турбины (3), холодильника-испарителя (4), компрессора (5), соединённого с турбиной, второго охлаждаемого компрессора (6*) с рубашкой охлаждения (11), вход которой связан с выходом холодильника-испарителя (4), а выход — с входом второго теплообменника-нагревателя (10), газодинамического разгонного устройства (7*), бака с криогенным рабочим телом (8), насоса (9), второго теплообменника-нагревателя (10).

Двигатель работает следующим образом. В замкнутом контуре преобразования тепла в полезную работу циркулирует рабочее тело, нагреваясь от источника тепла (1) в теплообменнике-нагревателе (2), и отдаёт энергию турбине (3).

Схема космического разгонного блока

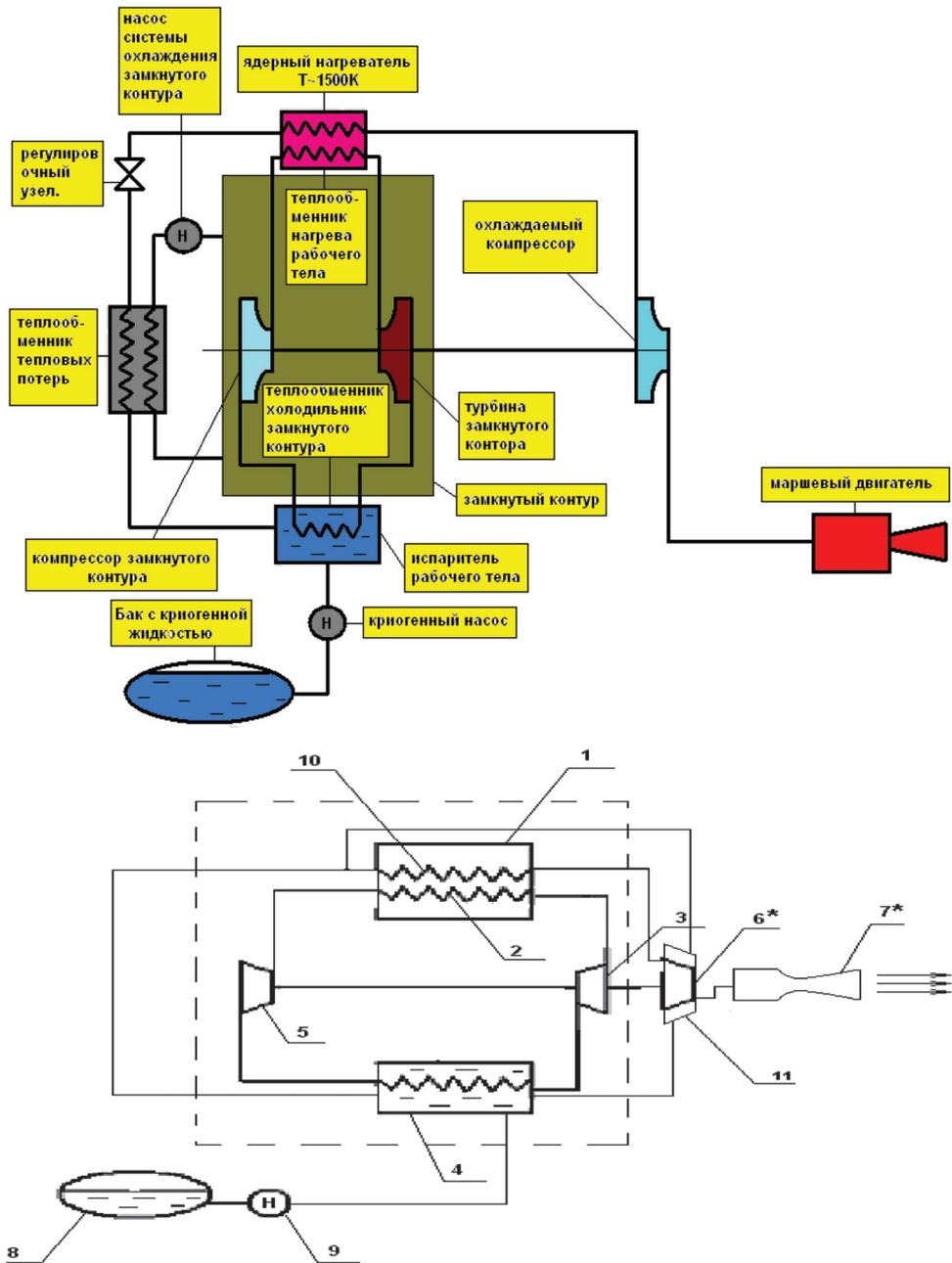


Рис. 4

После получения полезной работы рабочее тело замкнутого контура охлаждается в холодильнике-испарителе (4), сжимается компрессором (5) и снова поступает в теплообменник-нагреватель. Отводимое из контура тепло идёт на испарение и нагрев жидкого криогенного рабочего тела двигателя в холодильнике-испарителе (4). Испарившееся рабочее тело дополнительно подогревается в теплообменнике-нагревателе рабочего тела двигателя (10), сжимается охлаждаемым компрессором (6) и разгоняется в газодинамическом разгонном устройстве для достижения максимально возможной скорости истечения. Так как температура испарения криогенного рабочего тела очень низкая (примерно 20 К), термический КПД преобразования тепловой энергии в механическую работу близок к единице. Эта полезная работа преобразуется в охлаждаемом компрессоре при сжатии рабочего тела двигателя близком к адиабатическому в рост температуры и энтальпии рабочего тела перед его истечением из сопла, что позволяет получить скорость истечения более чем в 2–2,5 раза выше, чем у наиболее совершенных ЖРД.

В качестве примера рассматривается вариант двигателя со следующими основными параметрами:

- расход рабочего тела двигателя (водорода) — 1 кг/с;
- рабочее тело двигателя — жидкий водород;
- начальная температура водорода — 20 К;
- рабочее тело в замкнутом контуре — неон;
- температура неона на входе в турбину — 1500 К;
- начальная температура неона — 30 К.

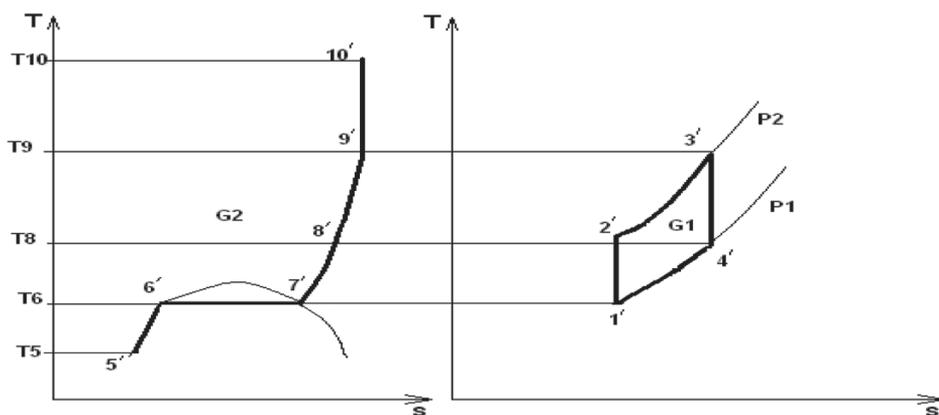


Рис. 5. Диаграмма ЯРД с замкнутой энергоустановкой: 1' — конечная точка охлаждения рабочего тела замкнутого контура энергоустановки; 2' — точка завершения сжатия рабочего тела в замкнутом контуре; 3' — точка завершения нагрева рабочего тела замкнутого контура в ядерном реакторе; 4' — точка завершения расширения рабочего тела в замкнутом контуре после получения полезной работы; 5' — начальное состояние рабочего тела двигателя; 6' — состояние рабочего тела двигателя перед началом его испарения; 7' — состояние рабочего тела двигателя после его испарения; 8' — точка завершения нагрева рабочего тела двигателя за счёт тепла, отводимого из замкнутого контура энергоустановки; 8'-9' — нагрев рабочего тела двигателя в ядерном реакторе; 9'-10' — адиабатическое сжатие рабочего тела двигателя и его нагрев в компрессоре рабочего тела двигателя (за счёт полезной работы, полученной в замкнутом контуре энергоустановки)

На рис. 5 (см. с. 148) замкнутый цикл (цикл Брайтона) неона обозначим точками 1'-2'-3'-4', а цикл нагрева и подвода полезной работы к рабочему телу (водороду) обозначим точками 5'-6'-7'-8'-9'-10'. Подвод тепла в цикле неона осуществляется изобарно на участке 2'-3' при давлении P_2 , отвод тепла также осуществляется изобарно на участке 4'-1' при давлении P_1 . Сжатие рабочего тела 1'-2' и расширение 3'-4' — адиабатическое.

Испарение и нагрев водорода осуществляется на участке 5'-6'-7'-8' за счёт отводимого тепла 4'-1' из замкнутого цикла. Дополнительный нагрев водорода от бортового источника тепла происходит на участке 8'-9' изобарно при давлении P_6 . Дальнейший рост температуры водорода на участке 9'-10' осуществляется за счёт адиабатического сжатия, которое зависит от полезной работы, полученной в замкнутом цикле неона.

Обозначим: Q — количество тепла, участвующее в этих циклах; G_1 — масса рабочего тела в замкнутом цикле; G_2 — масса водорода (в нашем примере равна 1 кг).

Запишем уравнение теплового баланса между отводимым теплом от неона и подводимым теплом к водороду.

$$\begin{aligned} Q_{5-6-7-8} &= Q_{1-4}, \\ Q_{5-6-7-8} &= Q_{5-6} + Q_{6-7} + Q_{7-8}, \\ Q_{5-6} &= (S_6 - S_5) \frac{T_6 + T_5}{2} G_2, \\ Q_{6-7} &= (i_7 - i_6) G_2, \\ Q_{7-8} &= (i_8 - i_7) G_2, \\ Q_{1-4} &= (i_4 - i_1) G_1, \end{aligned}$$

где S и i — энтропия и энтальпия рабочих тел.

При условии, что $G_2 = 1$ кг,

$$G_1 = \frac{(S_6 - S_5) \frac{T_6 + T_5}{2} + (i_8 - i_7)}{i_4 - i_1}.$$

Количество подведённого тепла в замкнутом контуре будет

$$Q_{3-2} = (i_3 - i_2) G_1.$$

Количество тепла, подведённое к водороду на участке 8-9, будет

$$Q_{8-9} = (i_9 - i_8) G_2 = (i_9 - i_8).$$

Полезная работа, полученная в замкнутом цикле, будет равна

$$Q_{\text{полезн}} = Q_{3-2} + Q_{4-1}, \quad Q_{\text{полезн}} = Q_{10-9}.$$

Дополнительный нагрев водорода на участке 9-10 находим из условия адиабатического сжатия за счёт подвода полезной работы.

$$Q_{9-10} = k \frac{R}{k-1} (T_{10} - T_9),$$

где k — отношение теплоёмкостей; R — газовая постоянная водорода.

$$\text{Отсюда } T_{10} - T_9 = Q_{\text{полезн}} \frac{k-1}{kR}.$$

Результаты расчётов с использованием начальных данных представлены в таблице.

	№ точки	T[K]	P[Па]	S[кДж/кг*К]	i[кДж/кг]
	0	1	2	3	4
0	0	0	0	0	0
1	1	30	$1 \cdot 10^5$	3.484	94.24
2	2	270	$5 \cdot 10^7$	3.484	320
3	3	$1.5 \cdot 10^3$	$5 \cdot 10^7$	6.36	$1.674 \cdot 10^3$
4	4	350	$1 \cdot 10^5$	6.36	360.5
5	5	20	$1 \cdot 10^5$	11.085	261
6	6	30	$8.07 \cdot 10^5$	16.495	409.5
7	7	30	$8.07 \cdot 10^5$	26.685	715.5
8	8	350	$8.07 \cdot 10^5$	59.02	$4.989 \cdot 10^3$
9	9	$1.5 \cdot 10^3$	$8.07 \cdot 10^5$	80	$2.22 \cdot 10^4$
10	10	$2.667 \cdot 10^3$		80	

Значение параметров рабочего тела взяты из источников [Зубарев и др., 1989; Варгафтик, 1972].

Расчёты показывают, что *предельное* значение приращения температуры, которое можно получить с использованием рассмотренной выше схемы, $T_{10} - T_9 = 1167$ К и достигаемая скорость истечения W , вычисленная по формуле $W = \sqrt{2kR(T/(k-1))}$, где $T = T_{10}$, $W = 9,969$ км/с, КПД замкнутого цикла 84 %.

Предельная тяга двигателя 1000 кг.

Тепловая мощность источника тепла не более 30 МВт (без учёта тепла на подогрев водорода в реакторе). Степень повышения давления в водородном компрессоре 7,5.

В случае использования реактора с тепловой мощностью 10 МВт и удельных затратах энергии на получение килограмма тяги как и в двигателе РД0410 (65 кВт/1 кг тяги) ожидаемая тяга рассматриваемого двигателя 154 кг. С учётом массовых характеристик представленных выше (за вычетом веса электрогенератора, излучателя и ЭРД) массовый коэффициент не превысит 0,7 кг/кВт, т. е. в 8 раз меньше, чем в ЯЭУ с ЭРД.

Таким образом, решается основная задача — достижение большой тяги и высокой скорости истечения рабочего тела ракетного двигателя при относительно невысокой рабочей температуре ядерного реактора и приемлемых массовых характеристиках. При использовании в качестве рабочего тела жидкого водорода скорость истечения рабочего тела будет в 2–2,5 раза превышать скорость истечения (и соответственно удельный импульс) наиболее совершенных водородно-кислородных ЖРД.

Необходимо отметить, что речь идёт о возможности *дополнительного* подогрева рабочего тела двигателя за счёт произведённой в замкнутом контуре ядерного реактора полезной работы.

Создание ядерной двигательной установки, выполненной по предлагаемой схеме, позволяет эффективно использовать низкотемпературное тепло, вырабатываемое реактором с рабочей температурой порядка 1500 К. При этом возможно достижение в двигателе удельного импульса 900 с и более. По конструкции предлагаемый двигатель подобен ЖРД замкнутой схемы.

В нём камера сгорания заменена ядерным источником тепла, а *дополнительный* нагрев рабочего тела осуществляется за счёт предварительного сжатия рабочего тела двигателя в компрессоре. Проблема охлаждения компрессора близка проблеме охлаждения камер сгорания ЖРД, которая успешно решена, хотя температура газов в камере превышает 3500 К.

При исследовании в 1960–1980 гг. возможности создания ядерного двигателя, выполненного по схеме с прямым нагревом рабочего тела, проведён большой объём исследований по нагреву газообразного водорода в ядерном реакторе. В предлагаемой схеме нагревается газообразный неон и водород, что позволит использовать накопленный научный задел и ускорить разработку двигателя. По нашему мнению существенно, что при относительно низкой температуре ядерного реактора, возможно его многократное включение и выключение без опасности разрушения из-за перегрева, что позволит создать надёжный космический аппарат длительного использования с многократным включением и многочасовым ресурсом. Кроме того, при нагреве водорода вне зоны реактора (что достигается посредством теплообмена с нагретым неоном в отдельном теплообменнике) радиоактивность выбрасываемого рабочего тела двигателя может быть существенно уменьшена. Это улучшит общую радиационную безопасность аппарата.

ВЫВОДЫ

Предлагаемая схема космической ядерной двигательной установки по нашему мнению является перспективной и реализуемой при достигнутом уровне техники.

Существенным достоинством предлагаемого технического решения стало следующее:

- все элементы конструкции двигательной установки могут быть отработаны на земле, что позволит подтвердить надёжность, высокий удельный импульс и многочасовой ресурс его работы;
- при отработке могут применяться моделирующий источник энергии и рабочие тела замкнутого контура и двигателя;
- отработка двигателя в целом так же может быть осуществлена на стендах в земных условиях, на орбиту ИСЗ он может быть доставлен в виде завершённого отработанного блока;
- эффективное использование тепла ядерного реактора позволит существенно увеличить тягу двигателя, и таким образом снизить гравитационные потери при разгоне и торможении космического аппарата вблизи планет;
- отказ от электрогенератора, излучателя тепла и электрореактивного двигателя с запасом рабочего тела для его работы увеличивает надёжность энергоустановки и приведёт к снижению веса космического аппарата.

ЛИТЕРАТУРА

- [Зубарев и др., 1989] *Зубарев В. Н., Козлов А. Д., Кузнецов В. Д.* Теплофизические свойства технически важных газов при высоких температурах и давлениях: справочник. М.: Энергоатомиздат, 1989. 232 с.
- [Варгафтик, 1972] *Варгафтик Н. Б.* Справочник по теплофизическим свойствам газов и жидкостей. М.: Наука, 1972.

NUCLEAR PROPULSION SYSTEM WITH A CLOSED POWER UNIT

V. S. Burikov, I. V. Burikov, V. Yu. Burikov

ООО ULTICAN "Alternative and cryogenic energy"

In the report the scheme of the rocket engine with a nuclear power source is considered. The engine includes a closed energy circuit with a nuclear heat source. The circuit produces mechanical energy, and not the used heat is transferred to the working body of the engine, located on the spacecraft. The working fluid after evaporation and preheating is further heated in a nuclear power source, compressed in the compressor and using a gas-dynamic nozzle creates engine thrust. The compressor compresses the working fluid by working in a closed loop.

The use of liquid hydrogen as a working fluid allows to obtain a large engine thrust and a high specific impulse.

Keywords: rocket engine, nuclear power source, closed energy circuit, liquid hydrogen.

Burikov Yury Vyacheslavovich — director, alticen@mail.ru

Burikov Igor Vyacheslavovich — senior engineer

Burikov Vladislav Sergeevich — scientific consultant, PhD

АНАЛИЗ ВОЗМОЖНОСТЕЙ КОСМИЧЕСКОЙ ТРАНСПОРТНОЙ СИСТЕМЫ ПРИ ИСПОЛЬЗОВАНИИ ЭЛЕКТРОРАКЕТНОЙ ДВИГАТЕЛЬНОЙ УСТАНОВКИ И ГРАВИТАЦИОННОГО МАНЁВРА У ЗЕМЛИ В НАЧАЛЕ МЕЖПЛАНЕТНОГО ПЕРЕЛЁТА

М. С. Константинов

Московский авиационный институт
(национальный исследовательский университет) (МАИ)

Проводится анализ увеличения транспортных возможностей космической системы при реализации межпланетных перелётов благодаря использованию (на начальном этапе полёта) гелиоцентрического перелёта Земля–Земля (на нём предполагается работа электроракетной двигательной установки) и гравитационного манёвра у Земли. Предполагается, что химический разгонный блок обеспечивает уход КА из окрестности Земли с относительно небольшим гиперболическим избытком скорости. Затем химический разгонный блок отделяется от КА. Гелиоцентрический перелёт Земля–Земля реализуется с использованием электроракетной двигательной установки. Благодаря её работе КА подлетает к Земле (для гравитационного манёвра у неё) с относительно большим гиперболическим избытком скорости (8...10 км/с). Таким образом, введение в схему межпланетного полёта участка гелиоцентрического перелёта Земля–Земля и гравитационного манёвра у Земли позволяет вывести на гиперболическую траекторию отлёта от Земли (с большим гиперболическим избытком скорости) КА довольно большой массы. В работе проанализирована зависимость массы КА на гиперболе отлёта от Земли (после гравитационного манёвра у неё) как функция величины гиперболического избытка скорости. Показано, что анализируемый манёвр позволяет существенно увеличить диапазон реализуемых гиперболических избытков скоростей и увеличить массу КА на гиперболе отлёта от Земли для фиксированных значений гиперболического избытка скорости. Анализируемый манёвр должен рассматриваться как начальный этап дальнейшего перелёта к исследуемому небесному телу или планете, у которой будет осуществляться гравитационный манёвр. То есть, как часть анализируемого маршрута межпланетного перелёта.

Ключевые слова: межпланетный перелёт, гиперболический избыток скорости, гравитационный манёвр, электроракетная двигательная установка

ВВЕДЕНИЕ

Проектно-баллистические возможности космической транспортной системы (ракеты-носителя (РН) и химического разгонного блока (ХРБ)) при межпланетных полётах во многом определяются массой КА, выведенного на гиперболическую отлётную от Земли траекторию, и величиной гиперболического избытка скорости на этой траектории. Принципиально для каждой существующей транспортной системы на сайте её производителя можно найти информацию о возможности системы в виде зависимости массы КА на гиперболической траектории отлёта от Земли как функции величины гиперболического избытка скорости. В наиболее простом случае, такая функция является монотонно убывающей во всём диапазоне возможных величин гиперболического избытка скорости.

Константинов Михаил Сергеевич — профессор, доктор технических наук, mkonst@bk.ru

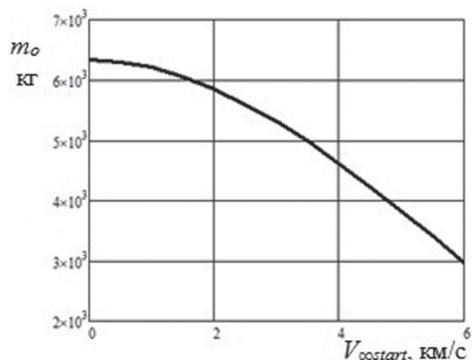


Рис. 1. Масса (кг) КА, выведенного на гиперболическую отлётную траекторию, как функция величины гиперболического избытка скорости (км/с); РН «Протон-М» и ХРБ «Бриз-М»

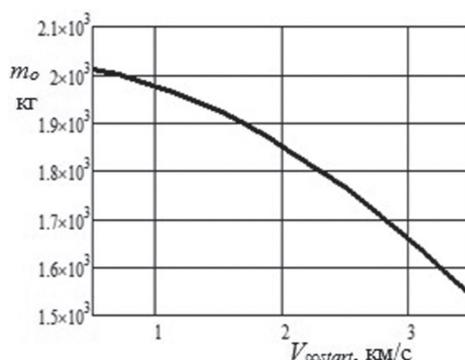


Рис. 2. Масса (кг) КА, выведенного на гиперболическую отлётную траекторию, как функция величины гиперболического избытка скорости (км/с); РН «Союз 2-16» и ХРБ «Фрегат»

На рис. 1, как пример, показана зависимость массы КА на гиперболической отлётной от Земли траектории для транспортной системы РН «Протон-М» и ХРБ «Бриз-М». Эта зависимость построена по данным взятым из работы [Proton..., 2009].

Из представленного графика видно, при использовании РН «Протон-М» и ХРБ «Бриз-М» можно обеспечить выведение КА на гиперболическую отлётную с величиной гиперболического избытка скорости до 6 км/с. Масса выведенного на гиперболическую траекторию КА m_0 является монотонно убывающей функцией величины гиперболического избытка скорости $V_{\infty start}$. При максимальном гиперболическом избытке скорости (6 км/с) эта масса равна 2961 кг. При гиперболическом избытке 5 и 4 км/с масса КА, выведенного на гиперболическую траекторию, оказывается равной 3821 и 4595 кг соответственно.

Важной характеристикой анализируемой функции является максимально достижимое значение гиперболического избытка скорости. Можно утверждать, что чем больше максимально достижимое значение гиперболического избытка скорости, тем более широкий круг задач может решить КА. Безусловно, очень важна величина массы КА на гиперболической траектории отлёта от Земли. Чем она больше, тем более широкий круг исследовательских задач может решить проектируемый КА. Величина гиперболического избытка скорости в 6 км/с легко позволяет реализовывать прямые перелёты КА к Марсу и Венере, но прямые перелёты к более дальним от Земли небесным телам (например, Юпитеру, Меркурию и ко многим астероидам главного астероидного пояса) реализовать невозможно.

Если проанализировать транспортные возможности транспортной системы на базе РН среднего класса, то они ещё более скромные. На рис. 2 представлена зависимость массы КА, выведенного на гиперболу отлёта от Земли как функция величины гиперболического избытка скорости, для транспортной системы на базе РН «Союз 2-16» и одного из вариантов ХРБ «Фрегат». На рисунке видно, что максимальная величина гиперболического избытка скорости в этом случае равна только 3,5 км/с. При таком гиперболическом избытке скорости масса КА на гиперболе отлёта от Земли равна 1540 кг.

ОБЩАЯ СХЕМА АНАЛИЗИРУЕМОГО КОСМИЧЕСКОГО МАНЁВРА

Существует возможность увеличить диапазон, реализуемых величин гиперболического избытка скорости за счёт реализации с использованием электроракетной двигательной установки (ЭРДУ) гелиоцентрического перелёта Земля – Земля и гравитационного манёвра у Земли. В настоящей работе проанализирована такая возможность на примере транспортной системы на базе РН «Союз 2-1б», ХРБ «Фрегат». Общая схема межпланетного перелёта при этом такова. РН и ХРБ обеспечивают выход КА из окрестности Земли с некоторым вектором гиперболического избытка скорости. Затем ХРБ отделяется от КА, и он осуществляет гелиоцентрический перелёт Земля – Земля. На этом перелёте может работать ЭРДУ. Она обеспечивает подлёт КА к Земле с довольно большой величиной гиперболического избытка скорости. Она существенно больше величины гиперболического избытка скорости при старте от Земли. Полагается, что характеристики гравитационного манёвра у Земли выбираются из условий выполнения последующего маршрута межпланетного перелёта, который в настоящей работе не анализируется. Идея использования на начальном этапе межпланетной траектории перелёта Земля – Земля с гравитационным манёвром у Земли широко использовалась при оптимизации схем выведения КА на гелиоцентрическую орбиту с малым перигелием и большим наклоном [Константинов, Тейн, 2013а, б; 2017; Konstantinov et al., 2014; Platov et al., 2016] в рамках проекта «Интергелиозонд» и при оптимизации траектории перелёта к Юпитеру [Константинов, Орлов, 2014; Константинов, Тейн, 2015, 2016; Константинов и др., 2015; Konstantinov, Orlov, 2014; Konstantinov, Thein, 2017a, b].

Основным итогом представляемой работы является зависимость массы КА на гиперболе подлёта к Земле для гравитационного манёвра у неё от величины гиперболического избытка скорости при гравитационном маневре.

ВОЗМОЖНЫЕ ПОСТАНОВКИ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ АНАЛИЗИРУЕМОГО МАНЁВРА

Эффективность использования на начальном этапе межпланетного перелёта манёвра с гелиоцентрическим перелётом Земля – Земля и гравитационного манёвра у Земли можно анализировать, используя многие постановки. Среди них можно указать такие, которые анализируют возможные диапазоны определённых характеристик гелиоцентрической траектории после гравитационного манёвра у Земли как функции величины массы КА на отлётной траектории. Среди этих характеристик можно использовать следующие характеристики:

- минимальная и максимальная достижимая энергии гелиоцентрической траектории (константа энергии или большая полуось);
- максимальный радиус афелия гелиоцентрической траектории;
- минимальный радиус перигелия гелиоцентрической траектории;
- максимальное наклонение гелиоцентрической траектории к плоскости эклиптики;

- максимальное удаление КА на гелиоцентрической траектории от плоскости эклиптики.

Можно рассматривать каждую из перечисленных характеристик как оптимизируемую и анализировать транспортные возможности космической системы, находя зависимость этой характеристики как функции массы КА на траектории полёта к Земле. Каждая из таких задач, безусловно, интересна. Но смущает большое число перечисленных (и не перечисленных) задач.

ВЫБРАННАЯ ПОСТАНОВКА ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ АНАЛИЗИРУЕМОГО МАНЁВРА

Для ответа на вопрос об эффективности манёвра мы посчитали целесообразным анализировать саму величину гиперболического избытка скорости при гравитационном манёвре у Земли, а не выше перечисленные характеристики. Эта оценка является более универсальной оценкой эффективности анализируемого манёвра. В общем случае, чем больше эта величина, тем большие возможности представляются КА для реализации межпланетной миссии. С другой стороны, наш подход не учитывает ограничения по возможному направлению гиперболического избытка скорости. Ясно, что во многих случаях при реализации межпланетного перелёта не будет целесообразно использовать максимальную величину гиперболического избытка скорости. Более того, для любого межпланетного перелёта существует оптимальная величина гиперболического избытка скорости при гравитационном манёвре у Земли, и эту величину следует определять.

При использованном в настоящей работе подходе задача была сформулирована так. Для заданных характеристик РН и ХРБ (заправка которого оптимизируется), заданных характеристиках ЭРДУ найти зависимость максимальной массы КА на гиперболической траектории полёта к Земле для гравитационного манёвра от величины гиперболического избытка скорости во всём допустимом диапазоне этого избытка.

Такая постановка нуждается в одном важном дополнении, которое связано с допущением о времени выполнения манёвра. Это время оптимизируется в диапазоне, ограниченном сверху некоторым предельным временем. Считалось, что время гелиоцентрического перелёта Земля—Земля ограничено сверху 1,5 годами. При этом траектория гелиоцентрического перелёта Земля—Земля имеет угловую дальность меньшую двух витков относительно Солнца. Как правило, эта угловая дальность немного меньше 450° . Такая угловая дальность соответствует рассматриваемой в работе постановке задачи.

МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

Предполагаем известными:

- характеристики транспортной системы РН + ХРБ. Зависимость массы КА на гиперболе отлёта от Земли при старте от неё (m_0) как функция величины гиперболического избытка скорости при старте ($V_{\infty start}$):

- $m_o = m(V_{\infty start})$. При численном анализе использовалась зависимость для РН «Союз 2-1б» и ХРБ «Фрегат», представленная на рис. 2;
- характеристики ЭРДУ (тяга P , удельный импульс W , массовый расход q).
 - имея в виду малый эксцентриситет орбиты Земли, мы посчитали возможным зафиксировать дату гравитационного манёвра у Земли: 10 апреля 2028 г. (юлианская дата 2461871.5). Такая дата была использована при анализе проекта солнечного зонда «Интергелио-Зонд» для эпохи запуска в 2026 г. и объясняется относительным положением Земли и Венеры в используемую дату.

Требуется найти оптимальные характеристики:

- дату старта от Земли,
- величину и направление гиперболического избытка скорости при старте от Земли,
- закон включения-выключения ЭРДУ на траектории перелёта Земля—Земля,
- углы ориентации вектора тяги ЭРДУ на активных участках траектории гелиоцентрического перелёта,

которые обеспечат заданную величину гиперболического избытка скорости при подлёте к Земле (V_{∞} , она перебирается в допустимом диапазоне) и максимальную величину массы КА в этот момент времени.

Таким образом, критерий оптимальности — масса КА в момент полёта к Земле для гравитационного манёвра у неё. Величина гиперболического избытка скорости при подлёте к Земле для гравитационного манёвра у неё перебирается от совсем малых значений, которые обеспечиваются РН + ХРБ, до максимально возможного значения.

МОДЕЛЬ ОПТИМАЛЬНОГО ДВИЖЕНИЯ КА НА ГЕЛИОЦЕНТРИЧЕСКОМ ПЕРЕЛЁТЕ. ФОРМУЛИРОВКА КРАЕВОЙ ЗАДАЧИ

Движение КА на гелиоцентрическом участке перелёта анализируется в гелиоцентрической эклиптической системе координат. Учитывается притяжение КА Солнцем и тяга ЭРДУ. Общий используемый метод оптимизации — принцип максимума Понтрягина. Анализируется полный набор условий трансверсальности. Уравнения оптимального движения записываются в следующем безразмерном виде:

$$\begin{cases} \frac{d\mathbf{r}}{dt} = \mathbf{V}; & \frac{d\mathbf{V}}{dt} = -\frac{\mathbf{r}}{r^3} + \frac{P\delta}{m} \cdot \frac{\vec{\lambda}_v}{\lambda_v}; & \frac{dm}{dt} = -\delta q; \\ \frac{d\vec{\lambda}_v}{dt} = -\vec{\lambda}_r; & \frac{d\vec{\lambda}_r}{dt} = \frac{\vec{\lambda}_v}{r^3} - \frac{3\mathbf{r}\lambda_{xyz}}{r^5}; & \frac{d\lambda_m}{dt} = \frac{P\lambda_v}{m^2} \delta. \end{cases}$$

где \mathbf{r} — безразмерный радиус-вектор КА (за единицу расстояния принимается астрономическая единица); \mathbf{V} — безразмерный вектор скорости КА (за единицу

скорости принимается местная круговая гелиоцентрическая скорость на расстоянии астрономической единицы от Солнца); t — безразмерное время; m — безразмерная масса; P — безразмерная тяга ЭРДУ (за единицу тяги рассматривалось произведение единицы массы на единицу ускорения); δ — функция включения-выключения двигателя; $\vec{\lambda}_r$ — вектор, сопряжённый к радиус-вектору \mathbf{r} ; $\vec{\lambda}_V$ — вектор, сопряжённый к вектору скорости \mathbf{V} (базис-вектор); λ_m — сопряжённая переменная к массе КА m ; λ_{xyz} — скалярное произведение векторов $\vec{\lambda}_V$ и \mathbf{r} : $\lambda_{xyz} = (\vec{\lambda}_V, \mathbf{r})$. При численном анализе за единицу массы мы посчитали целесообразным использовать не начальную массу на гиперболе отлёта от Земли при старте от неё (она зависит от величины гиперболического избытка скорости), а некоторую фиксированную массу. Было использовано значение 8200 кг — масса, выводимая РН на низкую околоземную орбиту, но возможно использование любой другой величины.

Оптимальный закон функции включения-выключения двигателя δ получен из условия максимума гамильтониана и используется в виде:

$$\delta = \begin{cases} 1, & \text{если } \psi > 0; \\ 0, & \text{если } \psi \leq 0, \end{cases} \quad \text{где функция переключения } \psi \text{ имеет вид: } \psi = (w/m)\lambda_V - \lambda_m.$$

В последнем соотношении используется безразмерный удельный импульс w .

Анализируется полный набор краевых условий, включая условия трансверсальности. Перечислим эти условия.

Условия оптимальности даты старта

$$\left(\vec{\lambda}_r(T_o), \mathbf{V}_{\infty start} \right) - \frac{P}{w} \psi(T_o) \delta(T_o) = 0.$$

В последнем равенстве использовано скалярное произведение вектора, сопряжённого к радиус-вектору КА в начальный момент времени (в дату старта) T_o , и вектора гиперболического избытка скорости при старте (при оптимальном его значении этот вектор коллинеарен базис-вектору $\vec{\lambda}_V$). В равенстве используется функция переключения и функция включения-выключения ЭРДУ в момент T_o . В том случае, если гелиоцентрический перелёт начинается пассивным участком, то из приведённого условия следует ортогональность вектора $\vec{\lambda}_r$ и базиса-вектора в начальной точке оптимального гелиоцентрического перелёта. Заметим, что такой случай является типичным для большого диапазона исследуемых значений гиперболического избытка скорости при гравитационном манёвре у Земли.

Условия оптимальности величины гиперболического избытка скорости при старте от Земли

$$\lambda_V(T_o) + \lambda_m(T_o) \cdot \frac{dm_o}{dV_{\infty start}} = 0.$$

В записанное условие входит производная массы КА в начальной точке гелиоцентрического перелёта (массы на гиперболе отлёта от Земли при старте от неё) по величине гиперболического избытка скорости при старте. Кроме отмеченной производной в соотношении присутствуют величина базис вектора и сопряжённой к массе переменной в начальной точке гелиоцентрического перелёта.

Векторное условие оптимальности направления вектора гелиоцентрической скорости при подлёте к Земле для гравитационного манёвра

$$\bar{\lambda}_V(T_f) - \frac{\mathbf{V}_{\infty f}}{|\mathbf{V}_{\infty f}|} = 0.$$

В соотношении T_f — дата подлёта КА к Земле для гравитационного манёвра у неё. Соотношение используется в виде, который обеспечивает коллинеарность вектора гиперболического избытка скорости при подлёте к Земле и базис-вектора КА и одновременно масштабирует сопряжённые переменные. (Система дифференциальных уравнений для сопряжённых переменных является однородной, семейство решений которой определяется с точностью до произвольной постоянной. Приведённое соотношение фиксирует эту постоянную.)

Кроме перечисленных пяти скалярных условий, на траектории перелёта должны быть удовлетворены краевые условия в момент подлёта КА к Земле:

- условия попадания КА в окрестность Земли для гравитационного манёвра у неё $\mathbf{r}(T_f) = \mathbf{r}_{Earth}(T_f)$. Используется модель движения Земли из DE-405;
- обеспечение зафиксированной величины гиперболического избытка скорости при гравитационном манёвре у Земли $V_{\infty f}$.

Таким образом, рассматриваемая задача имеет девять краевых условий. Количество неизвестных параметров краевой задачи тоже равно девяти: $V_{\infty start}$, T_o , $\bar{\lambda}_r(T_o)$, $\bar{\lambda}_V(T_o)$, $\lambda_m(T_o)$. Краевая задача сформулирована.

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО АНАЛИЗА

Рассматривается транспортная система на базе РН «Союз 2-1б», ХРБ «Фрегат» и ЭРДУ с тягой 0,234 Н и удельным импульсом 1670 с (соответствует удельному импульсу СПД-140). Гиперболический избыток скорости при гравитационном манёвре у Земли перебирался от значения 1,5 км/с до максимально возможного значения, которое требовалось определить.

На рис. 3–6 (см. с. 160) показаны проекции на плоскость эклиптики траекторий оптимального гелиоцентрического перелёта Земля–Земля для нескольких значений гиперболического избытка скорости при подлёте к Земле $V_{\infty f}$. Начальные и конечные точки перелёта показаны квадратом и треугольником соответственно. Активные участки перелёта показаны сплошной линией, пассивные — пунктиром. На активных участках показано направление тяги.

На рис. 3 величина $V_{\infty f}$ мала (2,5 км/с). Оптимальная траектория перелёта содержит всего один активный участок достаточно малой продолжительности (62 сут). Вся траектория состоит из двух пассивных участков, разделённых одним активным участком. Оптимальная величина гиперболического избытка скорости при старте от Земли небольшая (642,8 м/с). Требуемая масса рабочего тела для ЭРДУ всего 76,8 кг.

Увеличение $V_{\infty f}$ до 6 км/с (см. рис. 4) приводит к появлению нового активного участка, к увеличению суммарного времени работы ЭРДУ до 174,6 сут,

к увеличению требуемой массы ксенона до 216,1 кг, к увеличению оптимальной величины гиперболического избытка скорости при старте от Земли до 1391,4 м/с. Оптимальная траектория включает три пассивных участка, разделённые двумя активными участками.

При увеличении $V_{\infty f}$ до 8 км/с (см. рис. 5) оптимальной оказалась экстремаль новой типа. Характерной её особенностью является появление активного участка в начале гелиоцентрического перелёта. Это обстоятельство объясняет уменьшение оптимальной величины гиперболического избытка скорости при старте от Земли. Эта величина уменьшилась до 1332,1 м/с (при $V_{\infty f} = 6$ км/с она была равна 1391,4 м/с). Оптимальная траектория включает три активных и три пассивных участка. Время времени работы ЭРДУ увеличилось до 294,7 сут. Требуемая масса ксенона увеличилась до 364,7 кг.

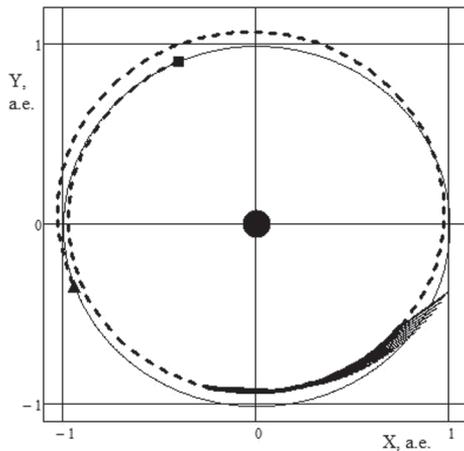


Рис. 3. Оптимальная траектория гелиоцентрического перелёта. $V_{\infty f} = 2,5$ км/с

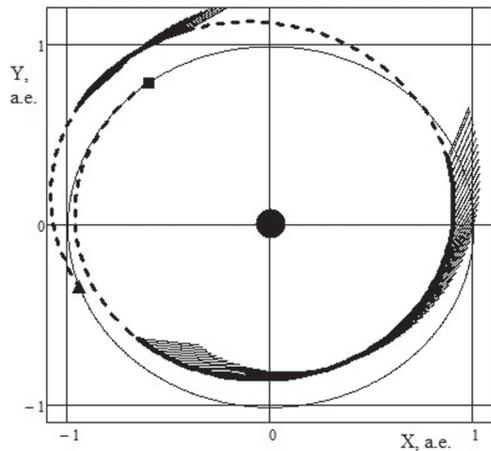


Рис. 4. Оптимальная траектория гелиоцентрического перелёта. $V_{\infty f} = 6$ км/с

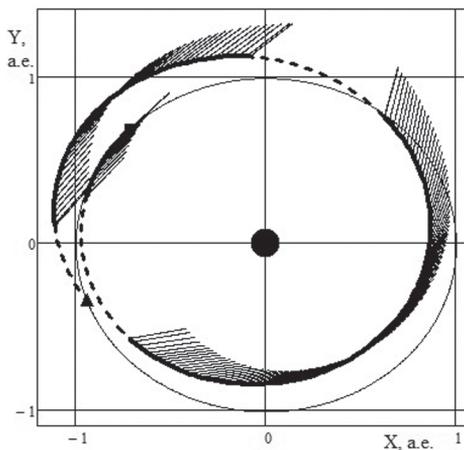


Рис. 5. Оптимальная траектория гелиоцентрического перелёта. $V_{\infty f} = 8$ км/с

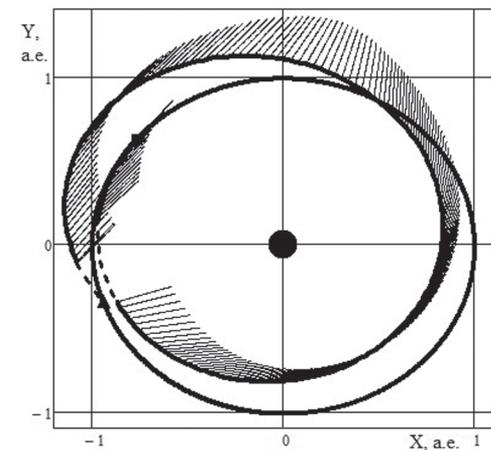
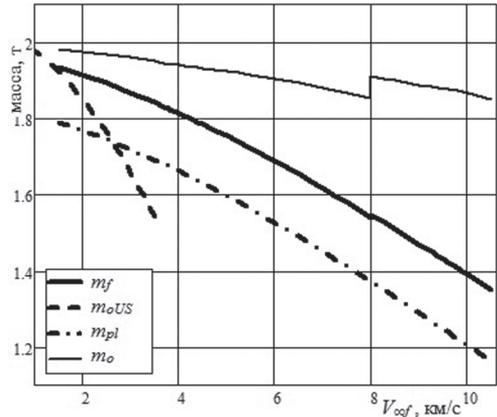


Рис. 6. Оптимальная траектория гелиоцентрического перелёта. $V_{\infty f} = 9,8$ км/с

Рис. 7. Масса КА в момент подлёта к Земле (m_f), масса КА на гиперболической траектории отлёта от Земли (при старте), обеспечиваемая ХРБ (m_{oUS}), масса полезной нагрузки КА (m_{pl}) и масса КА в начале гелиоцентрического перелёта при использовании ЭРДУ (m_o) как функции $V_{\infty f}$



Увеличение $V_{\infty f}$ до 9,8 км/с (см. рис. 6) приводит к увеличению времени работы ЭРДУ до 376,6 сут (при общем времени перелёта 424,8 сут), к увеличению требуемой массы ксенона до 466,1 кг,

к увеличению оптимальной величины гиперболического избытка скорости при старте от Земли до 1638,9 м/с. Оптимальная траектория включает два активных и два малой продолжительности пассивных участка. Траектория начинается активным участком и заканчивается очень небольшим (17,1 сут) пассивным участком.

На рис. 7 приведены массовые характеристики анализируемого проекта как функции величины гиперболического избытка скорости при подлёте к Земле $V_{\infty f}$. Пунктиром показана масса КА на гиперболической траектории отлёта от Земли (при старте), обеспечиваемая ХРБ. Видно, что рассматриваемая транспортная система (РН «Союз 2-1б», ХРБ «Фрегат») может обеспечить максимальный гиперболический избыток скорости около 3,5 км/с. При таком гиперболическом избытке скорости масса КА при отлёте от Земли (после отделения ХРБ) равна 1540 кг. Использование ЭРДУ на начальном этапе гелиоцентрического перелёта (перелёте Земля-Земля) позволяет значительно увеличить массу КА на отлёте от Земли после гравитационного манёвра у неё m_f . Эта зависимость показана на рисунке сплошной жирной линией. Анализ зависимости показывает, что максимальная величина $V_{\infty f}$ увеличивается до 10,6 км/с. При этом при $V_{\infty f} = 4, 6, 8, 10$ км/с масса КА при подлёте к Земле равна 1812,8, 1687,8, 1546,1, 1390,6 кг соответственно.

Сплошной тонкой линией на рис. 7 показана масса КА в начале гелиоцентрического перелёта при использовании ЭРДУ (масса КА после отделения ХРБ от КА). Эта зависимость оказалась разрывной. При $V_{\infty f} = 8$ км/с изменяется тип экстремали. При меньших $V_{\infty f}$ оптимальная траектория реализовывалась с большей величиной гиперболического избытка скорости при старте от Земли, и гелиоцентрическая траектория начиналась с пассивного участка. При больших значениях $V_{\infty f}$ оптимальная величина гиперболического избытка скорости при старте скачкообразно уменьшается, а гелиоцентрическая траектория начинается с активного участка. То есть оказывается целесообразным увеличение скорости при старте от Земли за счёт работы ЭРДУ, а не за счёт ХРБ.

Если использование ЭРДУ не будет предполагаться на дальнейшем этапе межпланетного перелёта (после гравитационного манёвра у Земли), целесообразно рассмотреть возможную модель ЭРДУ и ввести понятие массы полезной нагрузки. Это было выполнено так. Масса ЭРДУ m_{EP} полагалась линейной функцией входной электрической мощности ЭРДУ N_{input} и массы требуемого топлива m_{Xe} . Рассматривалась такая зависимость: $m_{EP} = (\gamma_{EP}/\eta_{EP})N_{input} + a_{10}m_{Xe}$,

где γ_{EP} — удельная масса ЭРДУ (вместе с энергетической установкой); a_{to} — удельная масса системы хранения и подачи топлива ЭРДУ; η_{EP} — КПД ЭРДУ. При расчётах принималось: $\gamma_{EP} = 40$ кг/кВт, $a_{to} = 0,1$, $\eta_{EP} = 0,55$. Массой полезной нагрузки m_{pg} считалась масса КА в момент окончания гелиоцентрического перелёта без массы ЭРДУ: $m_{pg} = m_f - m_{EP}$. Зависимость $m_{pg}(V_{\infty f})$ на рис. 7 показана штрихпунктирной линией. При $V_{\infty f} = 4, 6, 8, 10$ км/с масса полезной нагрузки m_{pg} равна 1660,8, 1526,5, 1369,9, 1203,2 кг соответственно.

На рис. 8 показаны зависимости требуемой массы топлива ЭРДУ m_{Xe} и времени работы ЭРДУ t_m как функций гиперболического избытка скорости при подлёте к Земле $V_{\infty f}$. Зависимости являются монотонно возрастающими со скачком при значении $V_{\infty f} = 8$ км/с. При этом значении $V_{\infty f}$ происходит изменение типа экстремали, пояснённое выше. Диапазон требуемого запаса топлива ЭРДУ 46,4—500,5 кг. Другой интересной особенностью функций на рис. 8 является угловая точка при значении аргумента 3805 м/с. Она мало выражена, но существует и будет хорошо видна при анализе зависимостей на рис. 9.

На рис. 9 показаны зависимости оптимальной величины гиперболического избытка скорости при старте от Земли $V_{\infty start}$ и оптимального времени гелиоцентрического перелёта t как функций гиперболического избытка скорости при подлёте к Земле $V_{\infty f}$.

Зависимость $V_{\infty start}(V_{\infty f})$ является монотонно возрастающей в трёх диапазонах непрерывности этой зависимости. Границами диапазонов являются точки $V_{\infty f} = 3805$ и 8000 м/с. Вторая из этих точек выше была проанализирована. В ней многие исследуемые характеристики оптимальной траектории имеют разрывы первого рода. Что касается первой точки (3805 м/с), то в ней исследуемые функции имеют разрыв первой производной. В этой точке при увеличении аргумента $V_{\infty f}$ появляется новый активный участок. Скорость изменения характеристик оптимальной траектории изменяется при переходе от экстремали с одним активным участком к экстремали с двумя активными участками. Зависимость $t(V_{\infty f})$ является монотонно убывающей в тех же трёх диапазонах. Она имеет разрыв первого рода в точке $V_{\infty f} = 8000$ м/с и разрыв производной в точке $V_{\infty f} = 3805$ м/с. Оптимальное время перелёта изменяется в узком диапазоне 423,0—452,8 сут.

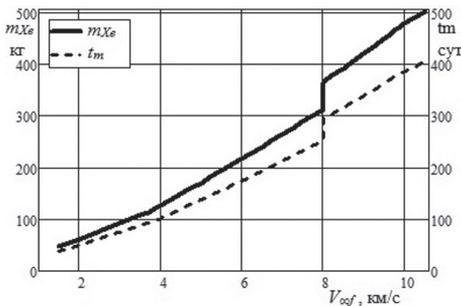


Рис. 8. Масса требуемого топлива ЭРДУ (m_{Xe} , кг) и время работы ЭРДУ (t_m , сут) как функции $V_{\infty f}$

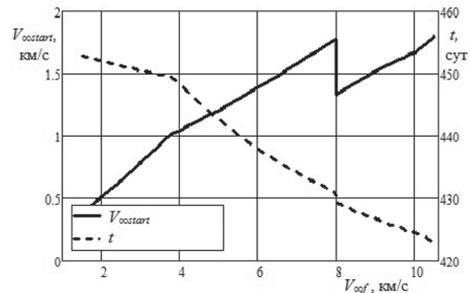
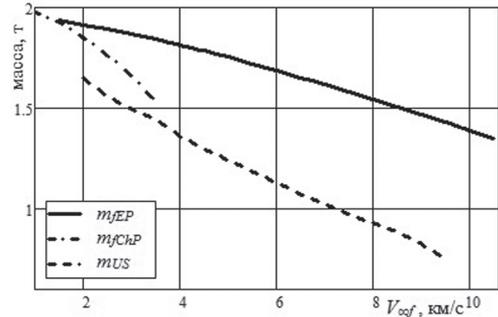


Рис. 9. Оптимальная величина гиперболического избытка скорости при старте от Земли ($V_{\infty start}$, км/с) и время гелиоцентрического перелёта (t , сут) как функции $V_{\infty f}$

Рис. 10. Масса КА в момент подлёта к Земле при использовании ЭРДУ m_{fEP} , масса КА в момент подлёта к Земле при использовании химического двигателя КА m_{fChP} и масса КА на гиперболической траектории отлёта от Земли (при старте), обеспечиваемая ХРБ m_{US} как функции $V_{\infty f}$



В заключение приведём результаты анализа возможной целесообразности использования на первом этапе межпланетной траектории гелиоцентрического перелёта Земля — Земля, на котором большая величина гиперболического избытка скорости при подлёте к Земле обеспечивается не ЭРДУ (как ранее рассматривалось), а химической двигательной установкой. Удельный импульс химического двигателя был принят равным 310 с. На рис. 10 показаны результаты такого анализа. Сплошная линия показывает массу КА при подлёте к Земле при использовании ЭРДУ на траектории перелёта Земля — Земля. Пунктирная линия показывает ту же массу, если на траектории перелёта используется химический двигатель КА. Для справки на рисунке показана (штрихпунктирная линия) и зависимость массы КА на гиперболической траектории отлёта от Земли (при старте), обеспечиваемая ХРБ. Видно, что возможный диапазон гиперболических избытков скорости при использовании манёвра Земля — Земля с использованием химического ракетного двигателя существенно увеличил диапазон возможных гиперболических избытков скорости (с 3,5 до почти 9,5 км/с). Но величина массы КА при подлёте к Земле при использовании ЭРДУ значительно больше, чем при использовании химического двигателя КА. При небольших значениях $V_{\infty f}$ использование ЭРДУ вместо химического двигателя позволяет увеличить массу КА при подлёте к Земле на ~20 %. При увеличении $V_{\infty f}$ выигрыш от использования ЭРДУ увеличивается и при $V_{\infty f} = 9,45$ км/с составляет 90 %.

ЗАКЛЮЧЕНИЕ

Использование на начальном этапе гелиоцентрического перелёта ЭРДУ и гравитационного манёвра у Земли увеличивает возможности транспортной системы. Оно позволяет существенно увеличить величину гиперболического избытка скорости при отлёте от Земли (после гравитационного манёвра у неё). При этом новых ограничений на дату старта не появляется. Масса КА при гравитационном манёвре у Земли существенно больше массы, которую могут обеспечить химические разгонные блоки при старте от Земли.

Большие гиперболические избытки скорости при гравитационном манёвре у Земли позволяют реализовывать достижение небесных тел (в частности, астероидов) с относительно небольшими затратами характеристической скорости.

Использование химических двигателей на траектории перелёта Земля — Земля может значительно увеличить гиперболический избыток скорости при

гравитационном манёвре у Земли по отношению к гиперболическому избытку скорости, обеспечиваемому ХРБ при старте от Земли. Но такой вариант при больших величинах гиперболического избытка скорости при гравитационном манёвре существенно проигрывает варианту использования ЭРДУ.

Выявлены интересные свойства оптимальных характеристик траектории перелёта Земля—Земля как функций величины гиперболического избытка скорости при подлёте к Земле. Оказалось, что существует несколько типов экстремалей траектории перелёта. Переход от экстремали одного типа к экстремали другого типа в некоторых случаях может вызывать разрывы первого рода в таких характеристиках как оптимальная масса требуемого топлива, оптимальная величина гиперболического избытка скорости при старте от Земли и некоторых других характеристиках. В случаях связанных с появлением новых активных участков на траектории перелёта перечисленные зависимости имеют разрыв производных. Учёт этих обстоятельств очень важен для разработки эффективных методов оптимизации траекторий КА.

Исследование выполнено при поддержке Российского научного фонда. Проект № 16-19-10429.

ЛИТЕРАТУРА

- [Константинов, Орлов, 2014] *Константинов М. С., Орлов А. А.* Оптимизация траектории к Юпитеру космического аппарата с малой тягой с использованием двух гравитационных манёвров у Земли // Вестник Моск. авиац. ин-та. 2014. Т. 21. № 1. С. 58–69.
- [Константинов, Тейн, 2013а] *Константинов М. С., Тейн М.* Оптимизация траектории выведения космического аппарата на рабочую гелиоцентрическую орбиту // Электрон. журн. «Труды МАИ». 2013. № 67. С. 20.
- [Константинов, Тейн, 2013б] *Константинов М. С., Тейн М.* Анализ одной схемы полёта космического аппарата для исследования Солнца // Электрон. журн. «Труды МАИ». 2013. № 71. С. 24.
- [Константинов, Тейн, 2015] *Константинов М. С., Тейн М.* Квазиоптимальные траектории полёта к Юпитеру с последовательностью гравитационных манёвров у Земли // Вестник НПО им. С. А. Лавочкина. 2015. Т. 22. № 4. С. 70–76.
- [Константинов, Тейн, 2016] *Константинов М. С., Тейн М.* Оптимизация перелёта КА с электроракетной двигательной установкой к Юпитеру при использовании последовательности гравитационных манёвров // Полёт. 2016. Т. 2–3. С. 30–37.
- [Константинов, Тейн, 2017] *Константинов М. С., Тейн Мин* Оптимизация траектории выведения КА на систему гелиоцентрических орбит // Космические исследования. 2017. Т. 55. № 2. С. 1–10.
- [Константинов и др., 2015] *Константинов М. С., Петухов В. Г., Тейн М.* Оптимизация траекторий гелиоцентрических перелётов: монография. М.: Изд-во МАИ, 2015. 259 с.
- [Konstantinov, Orlov, 2014] *Konstantinov M. S., Orlov A. A.* Optimization of the transfer trajectory of a low-thrust spacecraft for research of Jupiter using an Earth gravity-assist maneuver // Solar System Research. 2014. V. 48. Iss. 7. P. 605–610.
- [Konstantinov, Thein, 2017a] *Konstantinov M. S., Thein Min* Preliminary Optimization of the Complicated Interplanetary Flight Path of the Spacecraft with Electric Propulsion // Procedia Engineering. 2017. V. 185. P. 246–253.
- [Konstantinov, Thein, 2017b] *Konstantinov M. S., Thein Min.* Method of interplanetary trajectory optimization for the spacecraft with low thrust and swing-bys // Acta Astronautica. 2017. V. 136. P. 297–311.

- [Konstantinov et al., 2014] *Konstantinov M. S., Petukhov V. G., Thein M.* Optimization of the spacecraft insertion into the system of heliocentric orbits for Sun exploration // Proc. 65th Intern. Astronautical Congress. Toronto, Canada, 2014. Paper IAC-14.C1.9.4. P. 11.
- [Platov et al., 2016] *Platov I. V., Simonov A. V., Konstantinov M. S.* Choosing an Efficient Option of the Combined Propulsion System and Flight Profile of the Interhelio-Probe Spacecraft // Solar System Research. 2016. V. 50. No. 7. P. 581–586.
- [Proton..., 2009] Proton Launch System Mission Planner's Guide. Revision 7, October 2009. <http://www.ilslaunch.com/launch-services/proton-mission-planners-guide>.

ANALYSIS OF THE CAPABILITIES OF THE SPACE TRANSPORTATION SYSTEM USING AN ELECTRIC PROPULSION AND A GRAVITATIONAL MANEUVER NEAR THE EARTH AT THE BEGINNING OF AN INTERPLANETARY FLIGHT

M. S. Konstantinov

Moscow Aviation Institute (National Research University) (MAI)

An analysis is made of the increase in the transport capabilities of the space system in the implementation of interplanetary flights due to the use (at the initial stage of the flight) of the heliocentric Earth – Earth flight (on which electric propulsion (EP) is working) and of the gravitational maneuver near the Earth. It is assumed that the chemical upper stage provides relatively small hyperbolic excess of velocity. Then the chemical upper stage separates from the spacecraft (SC). Heliocentric flight Earth – Earth is realized using EP. A large hyperbolic excess of velocity (8...10 km/s) for a gravitational maneuver near the Earth is provided due to work of EP. Thus, the introduction into the scheme of interplanetary flight of the heliocentric Earth–Earth flight and of the gravitational maneuver at the Earth makes it possible to deduce a rather large mass of spacecraft on the hyperbolic trajectory of departing from the Earth (with a large hyperbolic excess of velocity). The SC mass in the hyperbola of flight from the Earth (after the gravitational maneuver near Earth) as a function of the magnitude of the hyperbolic excess of velocity is analyzed. It is shown that the analyzed maneuver allows significantly increasing the range of realized hyperbolic excesses of velocities and increasing the mass of the spacecraft on the hyperbola of flight from the Earth for fixed values of the hyperbolic excess of velocity. The analyzed maneuver should be considered as the initial stage of further flight to the investigated celestial body (as a part of whole mission).

Keywords: interplanetary flight, hyperbolic excess of velocity, gravitational maneuver, electric propulsion

Konstantinov Mikhail Sergeevich — professor, doctor of technical sciences, mkonst@bk.ru

УПРАВЛЕНИЕ ДВИЖЕНИЕМ КОСМИЧЕСКОГО АППАРАТА НА ГАЛО-ОРБИТЕ ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ НА НАПРАВЛЕНИЯ КОРРЕКТИРУЮЩИХ МАНЁВРОВ

С. А. Аксенов, С. А. Бобер

Институт космических исследований Российской академии наук (ИКИ РАН)

Работа посвящена проблеме управления космическим аппаратом, находящимся на гало-орбите при наличии ограничений на направления корректирующих манёвров. Для исследования выбрана орбита вокруг точки либрации L2 системы Солнце–Земля. Предложена стратегия, позволяющая удерживать аппарат в окрестности точки либрации, осуществляя коррекцию по направлениям, проекции которых на ось Солнце–Земля положительны. Другими словами, любая коррекция осуществляется в направлении «от Солнца». Разработанная стратегия управления космическим аппаратом проверена на имитационной модели, учитывающей неточности определения положения и скорости аппарата, а также наличие неточностей исполнения корректирующих манёвров. Показано, что удержание аппарата на орбите вокруг точки либрации при указанных ограничениях возможно.

Ключевые слова: гало-орбиты, управление движением, система Солнце–Земля, точки либрации

ВВЕДЕНИЕ

Размещение космических аппаратов на орбитах вокруг точек либрации системы Солнце – Земля позволяет решать задачи, недоступные для круговых или эллиптических орбит. Особенностью таких орбит является то, что движение космического аппарата происходит вокруг точки, положение которой относительно Солнца и Земли остаётся неизменным, что было впервые использовано в миссии ISEE-3 [Dunham et al., 2015] к точке либрации L1 системы Солнце – Земля. Это свойство орбит вокруг точек либрации незаменимо для астрометрических миссий Европейского космического агентства и NASA [Hechler, 2009]. Первой российской миссией, использующей орбиту вокруг точки либрации должна стать миссия «Спектр-РГ» к точке либрации L2 системы Солнце – Земля [Prin et al., 2014; Korobtsev et al., 2017].

Движение вокруг коллинеарных точек либрации является неустойчивым [Маркеев, 1978, Farquhar, Kamel, 1973], поэтому для долговременного пребывания на орбитах вокруг этих точек необходимо периодически осуществлять корректирующие манёвры. Величина необходимого для управления КА характеристического импульса зависит от точности определения положения и скорости КА и погрешности исполнения корректирующих манёвров. Современный обзор стратегий управления КА на орбитах вокруг точек либрации представлен в работе [Shirobokov et al., 2017].

Настоящая статья посвящена исследованию возможности управления КА при наличии ограничений на направления, по которым осуществляется кор-

Аксенов Сергей Алексеевич — ведущий математик, кандидат технических наук, aksenov.s.a@gmail.com

Бобер Станислав Алексеевич — программист, stas.bober@gmail.com

рекция. Такие ограничения могут возникнуть вследствие технических требований к аппарату при выполнении миссии. В работе исследуется вопрос о возможности удержания космического аппарата в окрестности точки либрации L2 системы Солнце – Земля, избегая корректирующих манёвров по направлению в сторону Солнца.

Для расчёта ограниченных орбит в окрестности коллинеарных точек либрации существует множество методов: аналитические [Farquhar, Kamel, 1973; Gomez et al., 1998; Richardson, 1980], полуаналитические [Howell, Pernicka, 1987; Jorba, Villanueva, 1998], численные [Aksenov, Bober, 2018; Kolemen et al., 2012; Ren, Shan, 2014, Zhang, Li, 2016].

В работе авторы использовали численный метод, предложенный ими ранее в статье [Aksenov, Bober, 2018], так как он достаточно прост в реализации и гибок в использовании: позволяет получать решения как в ограниченной задаче трёх тел, так и в эфемеридной постановке с учётом всех необходимых возмущений.

ПОСТАНОВКА ЗАДАЧИ

В качестве номинальной орбиты для исследования возможности контроля движения КА при наличии ограничений на направления манёвров коррекции была выбрана южная гало-орбита вокруг точки L2 системы Солнце – Земля с максимальным выходом из эклиптики в северном направлении равном 200 000 км. Во вращающейся системе координат, центр которой расположен в точке либрации, ось X сонаправлена с осью Солнце – Земля, ось Z направлена на север эклиптики и ось Y дополняет систему до правой тройки, амплитуда такой орбиты по оси Y составляет 737 146 км. Выбор начальных условий и расчёт номинальной орбиты осуществлялся в системе GMAT (General Mission Analysis Tool (NASA), URL: <http://gmatcentral.org/>, <https://sourceforge.net/projects/gmat/>) с помощью алгоритмов, описанных в работе [Aksenov, Bober, 2018].

На рис. 1 (см. с. 168) представлен вид номинальной орбиты, рассчитанной в ограниченной задаче трёх тел, моделирующей систему Солнце – Земля с учётом эксцентриситета Земной орбиты. Орбита представлена в проекции на эклиптику во вращающейся системе координат. Поскольку орбита Земли вокруг Солнца имеет ненулевой эксцентриситет, полученная гало-орбита, строго говоря, не является периодической. В верхней части рисунка хорошо видно раздвоение гало-орбиты, которое происходит вследствие эллиптичности орбиты Земли. Положение космического аппарата на орбите определяется углом (α) между проекцией радиус-вектора КА на эклиптику и направлением Солнце – Земля. Коррекция движения осуществляется с помощью манёвров, направления которых параллельны плоскости эклиптики. Направление корректирующего манёвра определяется углом (β), который оно составляет с направлением Солнце – Земля.

Целью работы является проверка возможности осуществления контроля движения КА вокруг точки либрации L2 с началом на номинальной орбите при осуществлении периодических коррекций, направление которых удовлетворяет следующим ограничениям (см. рис. 1): $\beta \in [-75^\circ, -5^\circ] \cup [5^\circ, 75^\circ]$. Минимальное время между коррекциями – 40 сут. Положение КА известно с точностью 5 км, скорость – 15 см/с. Ошибка исполнения корректирующего манёвра составляет не больше 6 %.

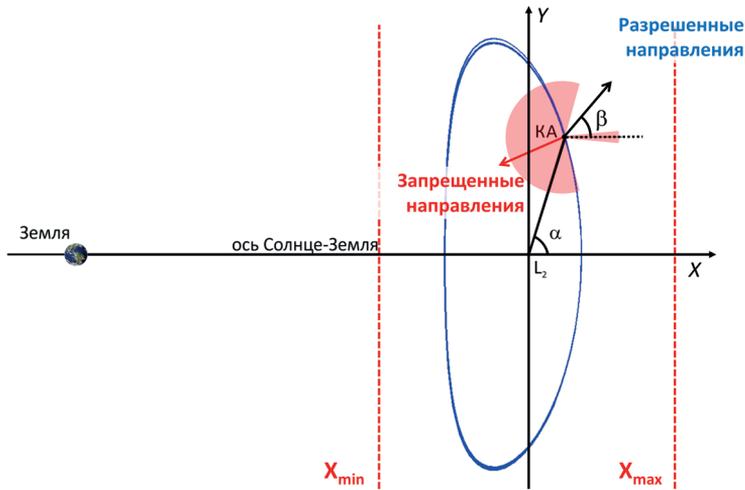


Рис. 1. Вид номинальной гало-орбиты в проекции на плоскость эклиптики и ограничения, накладываемые на направление коррекции

СТРАТЕГИЯ КОРРЕКЦИЙ

Хорошо известно [Zhang, Li, 2016], что гало-орбиты являются седловыми траекториями и лежат на пересечении устойчивого и неустойчивого многообразий решений задачи трёх тел. Бесконечно-малое возмущение вектора состояния, принадлежащего гало-орбите, приводит к переходу на одну из двух, соответствующих ему, орбит неустойчивого многообразия. Другими словами, движение приобретает неустойчивую составляющую, которая в зависимости от направления возмущения может быть положительной или отрицательной. Для точки L_2 системы Солнце – Земля, наличие в движении положительной неустойчивой составляющей будет соответствовать движению в сторону от Земли, отрицательной — к Земле. Этот факт использован в работе [Aksenov, Vober, 2018], для разработки численного алгоритма расчёта ограниченных орбит вокруг точки либрации. Суть алгоритма заключается в отделении окрестности, в которой лежит ограниченная орбита двумя плоскостями, ортогональными направлению Солнце – Земля. Численное интегрирование производится до момента пересечения одной из плоскостей, что сигнализирует о наличии в решении неустойчивой составляющей, которая нивелируется посредством корректирующего манёвра.

Для исследования влияния направления коррекции на величину неустойчивой компоненты вектор состояния КА, соответствующий движению по гало-орбите, подвергался возмущению величиной $\Delta v = 10\text{--}5$ км/с в направлении β , которое варьировалось от -180 до 180° , после чего производилось интегрирование орбиты до пересечения любой из плоскостей, показанных на рис. 1. Зависимость времени нахождения на орбите после возмущения вектора состояния и плоскости, которую он пересёк в момент окончания интегрирования от направления возмущения, показана на рис. 2 (см. с. 169). Положение КА соответствовало значению $\alpha = 180^\circ$.

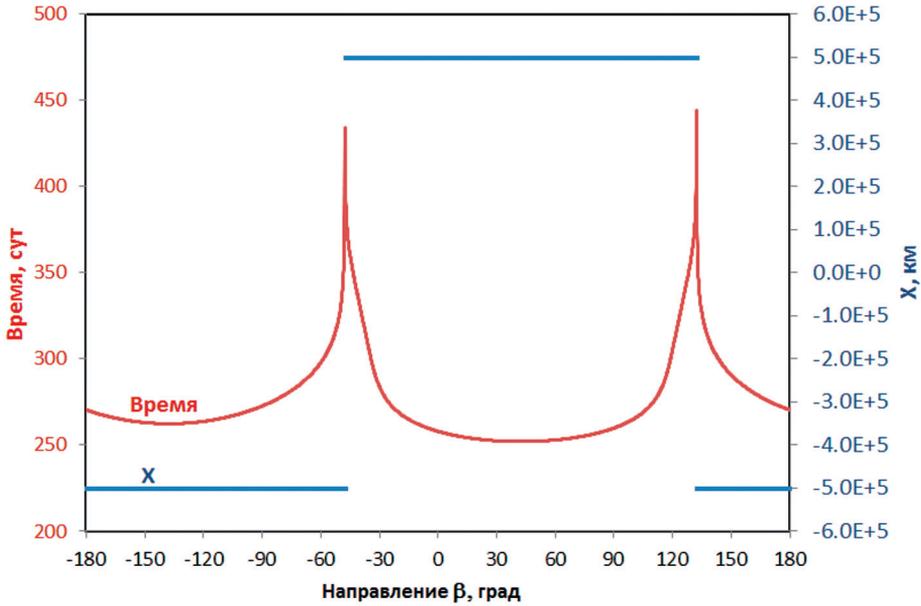


Рис. 2. Зависимости финальной координаты X (синяя линия) и времени движения аппарата между плоскостями $X = X_{\min} = -50\,000$ км и $X = X_{\max} = 50\,000$ км (красная линия) после возмущения скорости в направлении β в точке $\alpha = 180^\circ$

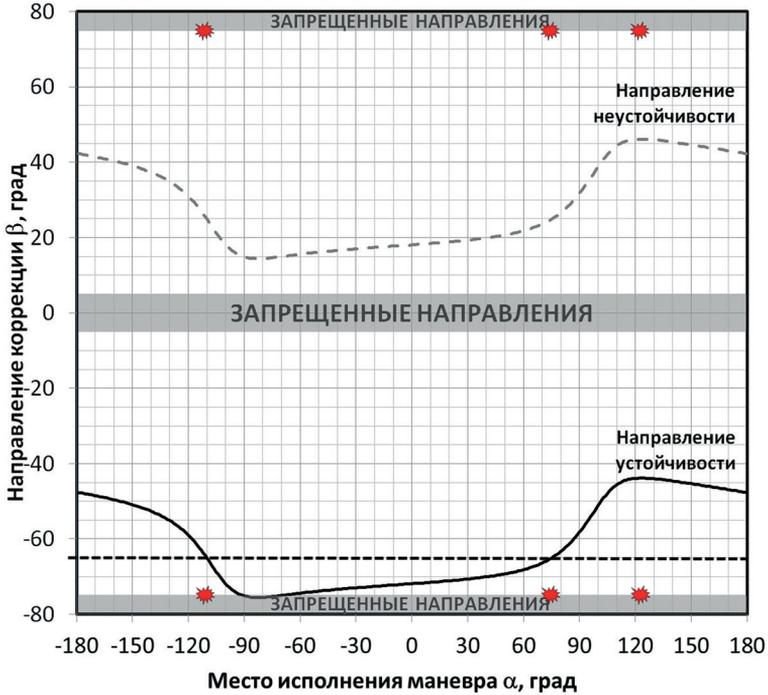


Рис. 3 Зависимость устойчивого и неустойчивого направлений от положения КА на орбите, наложенная на ограничения направлений коррекций

Результаты, проиллюстрированные на рис. 2, свидетельствуют, что время нахождения КА между двумя выделенными плоскостями существенно зависит от направления возмущения. Данная зависимость имеет два минимума в точках $\beta = -137,6^\circ$ и $\beta = 42,4^\circ$ и две вертикальные асимптоты в точках $\beta = -47,6^\circ$ и $\beta = 132,4^\circ$ с точностью до $0,1^\circ$. Минимумы отстают друг от друга на 180° , т. е. соответствуют одному направлению, возмущение по которому (положительное или отрицательное) окажет наибольшее влияние на движение КА. Возмущение по направлению, которому соответствуют асимптоты, напротив, приводит к минимальным изменениям орбиты. Таким образом, в каждой точке орбиты существуют два ортогональных направления (устойчивое и неустойчивое), которые необходимо принимать во внимание при осуществлении коррекции. Малые возмущения гало-орбиты по устойчивому направлению не приводят к возникновению неустойчивых компонент движения. Для рассмотренной точки орбиты, возмущения по направлениям в диапазоне $-47,6 < \beta < 132,4^\circ$ приводят к появлению положительной неустойчивой компоненты движения. Как видно из графика “X” на рис. 2, орбита в этом случае пересекает правую ограничивающую плоскость. Возмущения, выполняемые по направлениям, лежащим вне данного диапазона, приведут к появлению отрицательной неустойчивой компоненты движения. Таким образом, для эффективного управления КА, необходимо иметь возможность выполнения коррекций в обоих диапазонах направлений.

На рис. 3 (см. с. 169) приведена зависимость устойчивого и неустойчивого направления от местонахождения КА на орбите, наложенная на разрешённый диапазон ограничений. Можно видеть, что при нахождении КА на участке $-90 < \alpha < -60^\circ$ эффективное управление КА при заданных ограничениях невозможно, поскольку любая коррекция, удовлетворяющая ограничениям, может компенсировать только отрицательную неустойчивую компоненту. Для того чтобы можно было устранить как положительную, так и отрицательную неустойчивую компоненту движения, необходимо иметь возможность осуществлять коррекцию в направлениях, лежащих по разные стороны от устойчивого направления, отмеченного непрерывной линией на рис. 3.

Предлагаемая стратегия коррекции КА строится следующим образом. Корректирующие манёвры осуществляются по направлениям $\beta = \pm 75^\circ$ в трёх точках орбиты ($\alpha = -110^\circ$, $\alpha = 75^\circ$, и $\alpha = 123^\circ$), отмеченных на рис. 3. При попадании КА в одно из этих положений производится интегрирование его дальнейшего движения до пересечения одной из ограничивающих плоскостей для определения знака неустойчивой компоненты движения. Если неустойчивая компонента положительна, коррекция осуществляется по направлению $\beta = -75^\circ$, в противном случае — $\beta = +75^\circ$. Расчёт величины коррекции осуществляется согласно методике, изложенной в статье [Aksenov, Bober, 2018]. Положение $\alpha = 123^\circ$ позволяет выполнять коррекцию по направлению, наиболее удалённому от направления устойчивости. Манёвры в точках $\alpha = -110^\circ$ и $\alpha = 75^\circ$, необходимы для уменьшения интервала между коррекциями. Они выбирались таким образом, чтобы разница между возможным направлением коррекции и устойчивым направлением составляла как минимум 10° .

Алгоритм коррекции по направлениям, близким к неустойчивому, является наиболее эффективным, поскольку оказывает наибольшее влияние на характер движения.

РЕЗУЛЬТАТЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Для проверки предлагаемой стратегии проводилось имитационное моделирование движения КА на гало-орбите по методике Монте-Карло. Интегрирование осуществлялось в системе N тел, моделирующей влияние планет Солнечной системы и Луны с помощью системы GMAT. Для моделирования неопределенностей определения положения и скорости КА, его вектор состояния в начальный момент времени и в моменты коррекций возмущался случайным образом. Для возмущения радиус-вектора и вектора скорости к ним прибавлялся случайный вектор из шара, соответствующего радиуса. Расчёт значения корректирующего импульса осуществлялся для КА, положение которого изменено на случайный вектор из шара радиуса: 5 км — для координат аппарата, 15 см/с — для скорости аппарата. Рассчитанное значение корректирующего импульса Δv изменялось на случайный вектор из шара радиуса $0,06\Delta v$, применялось к невозмущённому вектору состояния КА и интегрирование продолжалось до следующего момента коррекции.

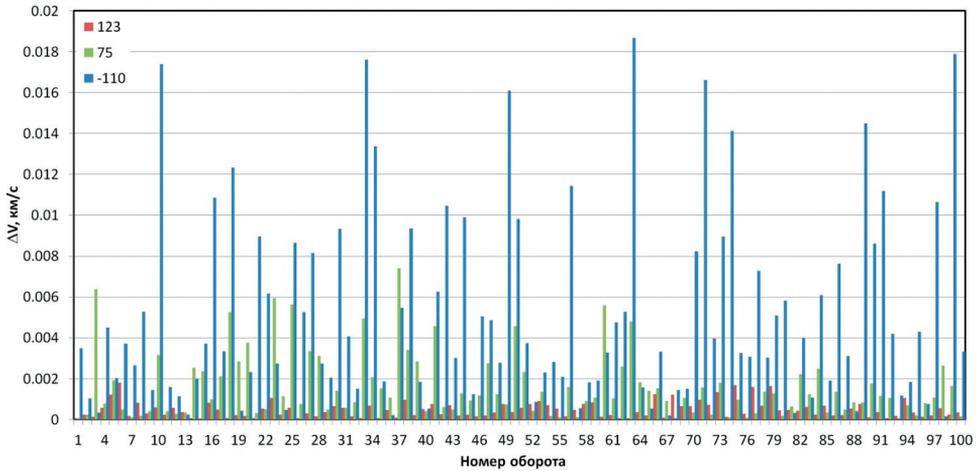


Рис. 4. Значения корректирующих импульсов: красный цвет в точке $\alpha = 123^\circ$, зелёный — $\alpha = 75^\circ$, синий — $\alpha = -110^\circ$

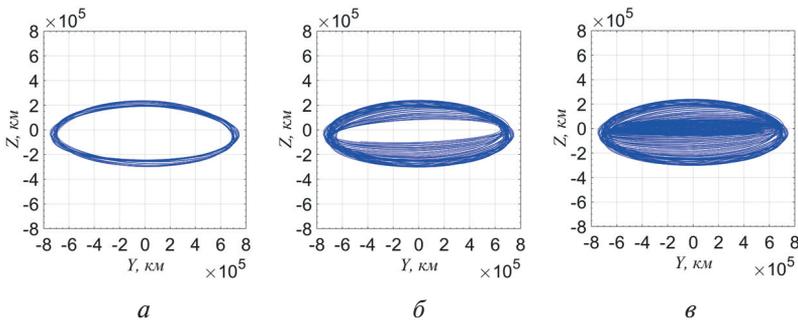


Рис. 5. Проекция рассчитанной орбиты на плоскость YZ :
 a — 20 оборотов; $б$ — 50 оборотов; $в$ — 100 оборотов

Суммарный корректирующий импульс при моделировании 100 оборотов вокруг точки либрации оказался равным 718 м/с, среднее значение импульса — 2,39 м/с, максимальное — 18,7 м/с, годовой расход характеристического импульса — 14,6 м/с. На рис. 4 (см. с. 171) приведена столбчатая диаграмма, иллюстрирующая величины Δv манёвров. Максимальные импульсы приходятся на точку $\alpha = -110^\circ$, где направление осуществления коррекции отстоит на 10° от неустойчивого и манёвр совершается после 84 сут неконтролируемого полёта. Минимальные значения корректирующих импульсов приходятся на точку $\alpha = 123^\circ$, в которой направление коррекции наиболее сильно отличается от направления устойчивости.

Проекции орбит, полученных при имитационном моделировании движения КА с периодическим осуществлением коррекций, рассчитанных согласно описанной методике с указанными разбросами по скоростям и положению аппарата, представлены на рис. 5 (см. с. 171). Рисунок 5а иллюстрируют траекторию, рассчитанную на 10 лет, рис. 5б — на 25 лет и рис. 5в — на 50 лет.

Результаты проведённого имитационного моделирования показывают, что в течение первых 20 оборотов разброс орбиты кардинально не увеличивается. При дальнейшем пребывании на орбите, большие корректирующие импульсы в неоптимальных направлениях приводят к вырождению первоначальной формы орбиты. Это вырождение сопровождается постепенным уменьшением амплитуд, приводящим к переходу орбиты из класса квазигало в класс Лиссажу. Тем не менее, разработанная стратегия позволяет удерживать аппарат в окрестности точки либрации даже после существенного изменения амплитуды орбиты и изменения её типа.

ЗАКЛЮЧЕНИЕ

В работе рассмотрена проблема управления космическим аппаратом на орбите вокруг точки либрации при наличии ограничений на направления корректирующих импульсов. Показано, что существует возможность удержания КА на гало-орбите амплитудой 200 тысяч километров при осуществлении корректирующих манёвров по направлениям, составляющим угол 75° с направлением Солнце — Земля.

Имитационное моделирование движения КА при использовании разработанной стратегии показало, что в течение первых 20 оборотов (что соответствует 10 годам) амплитуда орбиты в направлении, ортогональном эклиптике может меняться в пределах 20 %. При этом амплитуда в эклиптике меняется в пределах 7 %.

Приведённые в работе рассуждения могут быть распространены на построение других стратегий коррекции, отличающихся частотой осуществления манёвров, ограничениями на направления исполнения манёвров или характеристиками орбиты.

ЛИТЕРАТУРА

[Маркеев, 1978] *Маркеев А. П.* Точки либраций в небесной механике и космодинамике. М.: Наука, 1978. 312 с.

- [Aksenov, Bober, 2018] *Aksenov S. A., Bober S. A.* Calculation and Study of Limited Orbits around the L2 Libration Point of the Sun–Earth System // *Cosmic Research*. 2018. V. 56. Iss. 2. P. 144–150.
- [Dunham et al., 2015] *Dunham D. W., Farquhar R. W., Loucks M., Roberts C. E., Wingo D., Cowing K. L., Garcia L. N., Craychee T., Nickel C., Ford A., Colletuori M., Folta D. C., Giorgini J. D., Nace E., Spohr J. E., Dove W., Mogk N., Furfaro R., Martin W. L.* The 2014 Earth return of the ISEE-3/ICE spacecraft // *Acta Astronautica*. 2015. V. 110. P. 29–42.
- [Farquhar, Kamel, 1973] *Farquhar R. W., Kamel A. A.* Quasi-periodic orbits about the translunar libration point // *Celestial Mechanics*. 1973. V. 7. Iss. 4. P. 458–473.
- [Gomez et al., 1998] *Gomez G., Masdemont J. J., Simo C.* Quasihalo orbits associated with libration points // *J. Astronaut. Sci.* 1998. V. 46. Iss. 2. P. 135–176.
- [Hechler, 2009] *Hechler M.* Launch windows for libration point missions // *Acta Astronautica*. 2009. V. 64. Iss. 2–3. P. 139–151.
- [Howell, Pernicka, 1987] *Howell K. C., Pernicka H. J.* Numerical determination of Lissajous trajectories in the restricted three-body problem // *Celestial Mechanics*. 1987. V. 41. Iss. 1–4. P. 107–124.
- [Il'in et al., 2014] *Il'in I. S., Zaslavsky G. S., Lavrenov S. M., Sazonov V. V., Stepanyantz V. A., Tuchin A. G., Tuchin D. A., Yaroshevsky V. S.* Ballistic design of transfer trajectories from artificial-satellite earth orbit to halo orbit in the neighborhood of the L2 point of the Sun–Earth system // *Cosmic Research*. 2014. V. 52. Iss. 6. P. 437–449.
- [Jorba, Villanueva, 1998] *Jorba A., Villanueva J.* Numerical computation of normal forms around some periodic orbits of the restricted three-body problem // *Physica D*. 1998. V. 114. Iss. 3. P. 197–229.
- [Kolemen et al., 2012] *Kolemen E., Kasdin N. J., Gurfil P.* Multiple Poincaré sections method for finding the quasiperiodic orbits of the restricted three body problem // *Celestial Mechanics and Dynamical Astronomy*. 2012. V. 112. No. 1. P. 47–74.
- [Korobtsev et al., 2017] *Korobtsev I. V., Goryashin V. E., Eselevich M. V.* Results of tracking a spacecraft in the vicinity of the L2 libration point of the Sun–Earth system // *Astronomy Reports*. 2017. V. 61. Iss. 2. P. 153–159.
- [Ren, Shan, 2014] *Ren Y., Shan J.* A novel algorithm for generating libration point orbits about the collinear points // *Celestial Mechanics and Dynamical Astronomy*. 2014. V. 120. No. 1. P. 57–75.
- [Richardson, 1980] *Richardson D. L.* Analytic construction of periodic orbits about the collinear points // *Celestial Mechanics*. 1980. V. 22. Iss. 3. P. 241–253.
- [Shirobokov et al., 2017] *Shirobokov M., Trofimov S., Ovchinnikov M.* Survey of station-keeping techniques for libration point orbits // *J. Guidance, Control, and Dynamics*. 2017. V. 40. Iss. 5. P. 1085–1105.
- [Zhang, Li, 2016] *Zhang H., Li S.* A general method for the generation and extension of collinear libration point orbits // *Celestial Mechanics and Dynamical Astronomy*. 2016. V. 126. Iss. 4. P. 339–367.

CONTROL OF A SPACECRAFT MOTION ON A HALO-ORBIT WITH RESTRICTED MANEUVER DIRECTIONS

S. A. Aksenov, S. A. Bober

Space Research Institute of Russian Academy of Science (IKI RAN)

The paper is oriented to the problem of a spacecraft control in the halo-orbit with restrictions on the directions of corrective maneuvers. The Halo-orbit around the libration point L2 of the Sun–Earth system was subjected to the study. A station-keeping technique is proposed that allows to hold the spacecraft in vicinity of libration point, making corrections in directions whose

projections on the Sun-Earth axis are positive. In other words, all corrections are made in the direction “from the Sun”. The developed station-keeping strategy is verified on a simulation model that takes into account inaccuracies in determining the position and speed of a spacecraft, as well as the inaccuracies in the execution of corrective maneuvers. It is shown that holding the device in an orbit around the libration point under the indicated limitations is possible.

Keywords: Halo-orbits, motion control, Sun-Earth system, libration points

Aksenov Sergey Alekseevich — lead mathematician, PhD, aksenov.s.a@gmail.com

Bober Stanislav Alekseevich — programmer, stas.bober@gmail.com

ИССЛЕДОВАНИЕ ТРАЕКТОРИЙ КОСМИЧЕСКОГО АППАРАТА ДЛЯ ЭКСПЕДИЦИИ ЗЕМЛЯ – АПОФИС – ЗЕМЛЯ С УЧЁТОМ ВЫБОРА ОРБИТ ПРЕБЫВАНИЯ У АСТЕРОИДА

А. Лан¹, В. В. Ивашкин^{1,2}

¹ Московский государственный технический университет имени Н. Э. Баумана (МГТУ им. Н. Э. Баумана)

² Институт прикладной математики имени М. В. Келдыша Российской академии наук (ИПМ им. М. В. Келдыша РАН)

В работе исследованы траекторные характеристики экспедиции к потенциально опасному астероиду Апофис с пребыванием космического аппарата (КА) и специального мини-спутника у астероида в течение некоторого времени и возвратом к Земле — для изучения этого астероида и уменьшения астероидной опасности. Разработаны методика построения энергетически оптимальных межпланетных траекторий Земля – астероид и астероид – Земля, а также методика анализа орбитального движения спутника астероида. Определены и проанализированы оптимальные по максимуму полезной массы КА траектории для экспедиции Земля – Апофис – Земля с двигательными установками большой тяги. Проанализировано пассивное орбитальное движения основного КА и мини-спутника с радиомаяком около астероида Апофис с учётом возмущений от притяжения небесных тел, несферичности астероида и давления солнечного света (ДСС).

Ключевые слова: экспедиции к астероиду, оптимальные межпланетные траектории, орбитальное движение спутника астероида, астероид Апофис, «время жизни» спутника астероида

ВВЕДЕНИЕ

Активные исследования малых тел Солнечной системы (астероидов, комет) с помощью автоматических межпланетных станций начались в 1980-х гг. К главным причинам этих исследования можно отнести изучение механизмов происхождения и эволюции Солнечной системы, проблемы астероидно-кометной опасности (АКО), а также проблемы использования в будущем материальных ресурсов этих тел. Сейчас возрастает роль экспедиций к малым небесным телам с взятием образцов их вещества и возвращением КА к Земле. К данному моменту реально разработаны четыре космические миссии — Stardust (США), Hayabusa (Япония), Hayabusa-2 (Япония) и OSIRIS-REx (США), в первых трёх миссиях осуществлён или осуществляется полёт КА с двигательной установкой малой тяги к ядру кометы или астероиду и возврат к Земле. Для российской космонавтики в настоящее время более реально использование обычных химических двигательных установок большой тяги (ДУБТ). Однако эти ДУ приводят к большому расходу топлива, что делает особенно актуальной

Лан Аньци — стажёр, магистр, аспирант, seatu_angel@126.com

Ивашкин Вячеслав Васильевич — главный научный сотрудник, доктор физико-математических наук, профессор Ivashkin@Keldysh.ru

оптимизацию межпланетных траекторий экспедиции. Кроме того, при проектировании траекторий полёта КА для такой экспедиции также надо выбрать рациональные орбиты спутника астероида — для повышения «времени жизни» этого спутника у астероида. В работе разработана общая методика построения траекторий для экспедиции Земля — астероид — Земля, и на этой основе определены траектории экспедиции к астероиду Апофис.

Рассмотрена следующая схема экспедиции. Ракета-носитель (РН) выводит КА с разгонным блоком (РБ) на опорную орбиту. После пассивного движения по ней в некоторый оптимальный момент t_0 блоком РБ сообщается импульс скорости ΔV_1 , производится разгон КА, и КА переводится на орбиту полёта к астероиду Апофис. Затем РБ отделяется от КА, и дальнейшие манёвры осуществляются с помощью второй двигательной установки большой тяги, ДУ2. В момент $t_{1\text{сд}}$ аппарат выходит из сферы действия Земли. Далее, в момент t_2 КА подлетает к астероиду. С помощью ДУ2 сообщается импульс скорости ΔV_2 , осуществляется торможение КА, и КА переходит на орбиту спутника астероида. В окрестности Апофиса КА пребывает некоторое время Δt_{23} , это — «время ожидания». В течение этого времени возможны посадка на поверхность астероида, взятие образцов его грунта и другие исследования. Одной из главных целей экспедиции считается выведение мини-спутника астероида Апофис с радиомаяком с целью более точного определения орбиты Апофиса. Поэтому полагаем, что после специальных манёвров от основного КА отделяется и оставляется на некоторой орбите спутника астероида мини-аппарат с радиомаяком, который должен летать вокруг Апофиса в течение нескольких лет. Затем, в момент t_3 основному КА сообщается импульс скорости ΔV_3 , КА разгоняется и переходит на траекторию возвращения к Земле. В момент $t_{4\text{сд}}$ КА подлетает к сфере действия Земли. От КА отделяется спускаемый аппарат, и в основном варианте анализа в момент t_f происходит его гиперболический вход в атмосферу Земли, аналогично [Ajluni et al., 2015; Atkins et al., 1997; Brownlee et al., 2003; Sandford, 2011], затем — аэродинамическое управляемое торможение, посадка. В этом случае энергетические затраты на экспедицию в номинале определяются тремя величинами импульсов скорости ΔV_1 , ΔV_2 , ΔV_3 . Дополнительно рассмотрен также вариант эллиптического входа в атмосферу Земли с предварительным ракетным торможением.

1. ПОСТРОЕНИЕ ОПТИМАЛЬНЫХ ТРАЕКТОРИЙ КОСМИЧЕСКИХ АППАРАТОВ ДЛЯ ПОЛЁТА К АПОФИСУ С ВОЗВРАЩЕНИЕМ К ЗЕМЛЕ

Задача определения оптимальных межпланетных траекторий решается в два этапа. На первом этапе гелиоцентрические траектории Земля—астероид и астероид—Земля определяются в модели точечных сфер действия Земли и астероида, в импульсной постановке.

Схема решения задачи будет следующей. При задании граничного времени экспедиции t_1 (отлёт с орбиты Земли), t_2 (подлёт к орбите астероида), t_3 (отлёт с орбиты астероида), t_4 (подлёт к орбите Земли) гелиоцентрические орбиты перелёта между небесными телами определяются путём двукратного решения задачи Эйлера—Ламберта (с учётом возможности совершения одного пассивного

витка хотя бы по одной орбите). Это позволяет найти скорости «на бесконечности» $V_{\infty 1}, V_{\infty 2}, V_{\infty 3}, V_{\infty 4}$ и требуемые импульсы скорости для перелёта $\Delta V_1, \Delta V_2, \Delta V_3$. Суммой величин этих импульсов является характеристическая скорость $V_{\text{хар}}$. Определяются также конечная m_f и полезная m_p массы КА. Полезная масса КА определяется с учётом отделяемых масс РБ и ДУ2, при этом скорости истечения газов из двигательных установок РБ и ДУ2 различны [Ивашкин и др., 2013]. Необходимо выбором значений времени t_1, t_2, t_3, t_4 (при заданных областях для значений этого времени) найти оптимальные траектории с максимальной полезной массой:

$$m_p \rightarrow \max. \quad (1)$$

Рассмотрено несколько задач оптимизации:

1. Основная задача оптимизации сформулирована следующим образом: при заданной общей продолжительности экспедиции $\Delta t_{\Sigma} = t_4 - t_1$, и заданном времени пребывания КА у астероида $\Delta t_{23} = t_3 - t_2$ оптимизируются время старта t_1 и время перелёта от Земли до астероида $\Delta t_{12} = t_2 - t_1$, чтобы выполнялось условие (1). Эта постановка близка к работе [Ильин, Кузмак, 1976].
2. При заданном времени Δt_{23} и при ограничении на общую продолжительность экспедиции Δt_{Σ} (например, $\Delta t_{\Sigma} \leq 2$ года), оптимизируются $t_1, \Delta t_{12}$ и Δt_{Σ} .
3. При заданном времени Δt_{Σ} , оптимизируются $t_1, \Delta t_{12}$, и Δt_{23} .
4. Оптимизируются все значения времени $\Delta t_{\Sigma}, t_1, \Delta t_{12}$ и Δt_{23} .
5. Полная 4-параметрическая оптимизация времени $\Delta t_{\Sigma}, t_1, \Delta t_{12}$ и Δt_{23} с учётом ограничения на скорость входа КА в атмосферу Земли $V_{\text{вх}}$:

$$V_{\text{вх}} \leq V_{\text{max}}$$

Для поиска оптимальных траекторий на данном этапе использовано несколько методов: метод И. М. Соболя [Соболь, Стагинков, 1981], генетический алгоритм (ГА) [Панченко, 2007], квазиньютоновский BFGS (Broyden-Fletcher-Goldfarb-Shanno) метод [Nocedal, Wright, 2006]. Метод И. М. Соболя с точками ЛП_τ-последовательностей, которые очень равномерно распространены в пространстве, позволяет найти области, где расположены локальные оптимумы и глобальный оптимум. Запуск ГА в этих областях определяет глобальный оптимум с точностью до суток. Для реализации простого ГА используются следующие основные операторы: рулеточный отбор для выбора «родителей»; одноточечный кроссинговер с вероятностью $P_c = 0,8$; одноточечная мутация с вероятностью $P_m = 0,01$ и элитарный отбор «особей» в следующее поколение. BFGS метод позволяет затем быстро (за 2–8 итераций) уточнить оптимум.

После определения оптимальных гелиоцентрических траекторий перелёта КА на первом этапе, эти траектории проверяем на выполнение необходимых условий оптимальности в классе многоимпульсных перелётов с помощью сопряжённых функций для функционала (1). Для этого получены выражения для базис-вектора Лоудена \mathbf{p} (сопряжённого к скорости КА вектора $\vec{\lambda}_V$) в значения граничного времени t_1, t_2, t_3, t_4 [Ивашкин, Лан, 2017а] и разработан алгоритм построения сопряжённых функций на основе этих граничных векторов. Выполнение условия

$$p(t) = |\mathbf{p}(t)| \leq 1 \quad (2)$$

необходимо для оптимальности траекторий в классе многоимпульсных перелётов. Если это условие не выполняется на некотором участке, то данную траекторию можно улучшить введением дополнительных импульсов или вариацией граничного времени.

На втором этапе анализа уточняются характеристики полученных оптимальных межпланетных траекторий. При этом траектории КА определяются — с учётом возмущений (притяжения небесных тел, сжатия Земли, давления солнечного света (ДСС)) и эфемерид для определения координат небесных тел - численным интегрированием системы дифференциальных уравнений движения КА и решением краевых задач для выполнения граничных условий. Выполнена также коррекция массово-энергетических характеристик. Для этого учтены гравитационные потери при разгоне КА у Земли из-за конечности тяги; предусмотрены дополнительные импульсы скорости на коррекцию траекторий; уточнены массовые характеристики отделяемых частей РБ; предусмотрены гарантийные запасы топлива; введены отделяемые массы — для мини-спутника астероида и посадочного устройства. Выполняется оптимизация уточнённых траекторий по граничным временам, методом координатного спуска [Аббасов, 2014].

В соответствии с разработанной методикой, для экспедиции Земля – Апофис – Земля определены и исследованы энергетически оптимальные, по максимуму полезной массы КА, межпланетные траектории с запуском КА в течение 2019–2022 гг. [Ивашкин, Лан, 2017б]. Сначала исследован вариант с использованием РН «Союз-ФГ» и РБ «Фрегат». Применительно к двигательной установке ДУ2 принято, что её удельная тяга 304 с; постоянная составляющая массы $m_{20} = 100$ кг; коэффициент массы топливных баков $a_{T2} = 0,15$. Начальный радиус низкой орбиты искусственного спутника Апофиса (ИСА) для основного КА составляет 500 м.

Для основной задачи (1) оптимизации фиксированы время ожидания $\Delta t_{23} = t_3 - t_2 = 7$ сут и суммарное время Δt_{Σ} , которое выбиралось из множества $T_S = [390; 420; 450; 510; 540; 570; 600; 630; 660; 690; 730]$ сут. Варьируются и значения t_1 и $\Delta t_{12} = t_2 - t_1$. Решение этой основной задачи на первом этапе анализа дало оптимальные траектории для разных значений времени Δt_{Σ} . Наибольшая масса $m_p = 272$ кг соответствует $\Delta t_{\Sigma} = 690$ сут, характеристики этой траектории (№ 16) даны в табл. 1.

Таблица 1. Характеристики оптимальных траекторий № 16, 19, 20

Номер траектории	Δt_{Σ} , сут	t_1	Δt_{12} , сут	Δt_{23} , сут	t_4	$V_{\text{хар}}$, км/с	m_f , кг	m_p , кг
16	690	24.05.2019	335	7	13.04.2021	6,618	527	272
19	690	23.05.2019	336	93	12.04.2021	6,519	544	293
20	716	05.05.2020	300	112	21.04.2022	6,343	545	329

Решением задачи (2) трёхмерной оптимизации при заданном времени ожидания Δt_{23} из множества [7; 30; 60; 90; 120; 130] сут при $\Delta t_{\Sigma} \leq 2$ года определено оптимальное время ожидания КА у Апофиса: $\Delta t_{23opt} \sim 90-120$ сут (рис. 1, см. с. 179). При оптимальном времени ожидания полезная масса увеличивается до ~ 328 кг. Так, для траектории с $\Delta t_{23} = 120$ сут: $t_1 = 06.05.2020$, $\Delta t_{12} = 297$ сут, $\Delta t_{\Sigma} = 716$ сут, $m_p = 328$ кг.

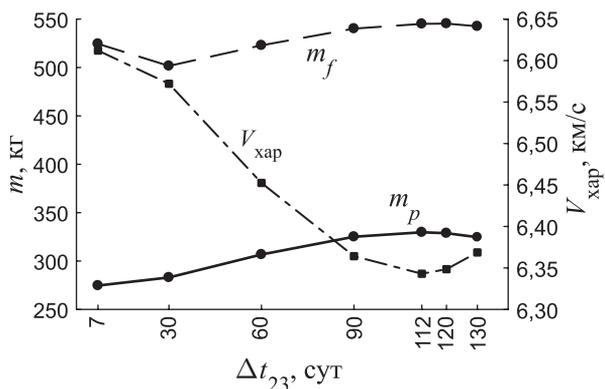


Рис. 1. Зависимость V_{xap} , m_f , m_p от времени Δt_{23}

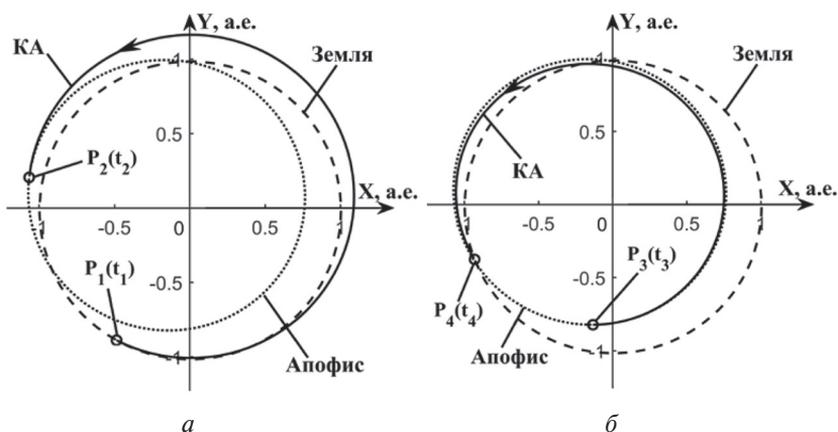


Рис. 2. Межпланетные перелёты КА для траектории № 19: а — перелёт от Земли (P_1) до Апофиса (P_2); б — перелёт от Апофиса (P_3) до Земли (P_4)

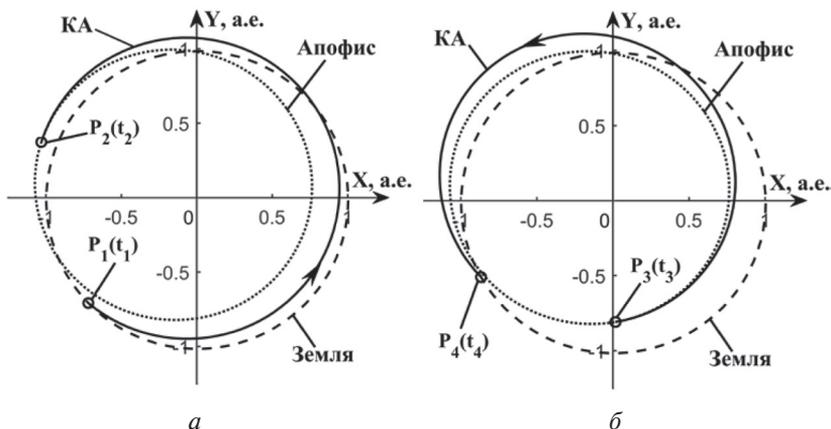


Рис. 3. Межпланетные перелёты КА для траектории № 20: а — перелёт от Земли (P_1) до Апофиса (P_2); б — перелёт от Апофиса (P_3) до Земли (P_4)

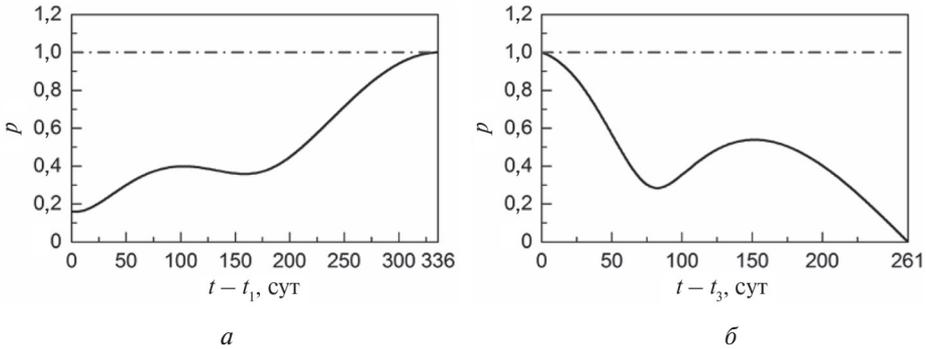


Рис. 4. Изменение модуля базис-вектора на траектории № 19: *а* — перелёт от Земли до Апофиса; *б* — перелёт от Апофиса до Земли

Решены также дополнительные задачи: (3) — трёхмерной оптимизации при задании Δt_{Σ} ($= 690$ сут) и (4) — полной четырёхмерной оптимизации при условиях $\Delta t_{\Sigma} \leq 2$ года, $\Delta t_{23} \geq 7$ сут. В табл. 1 приведены характеристики полученных оптимальных траекторий № 19, 20, соответственно. Полезная масса здесь больше, чем для траектории № 16. Для этих вариантов возврат к Земле происходит вблизи восходящего узла орбиты Апофиса. Картинки траекторий № 19, 20 в плоскости эклиптики показаны на рис. 2 и 3 (см. с. 179).

Для полученных траекторий построены сопряжённые функции, в частности, модуль базис-вектора. Так как для траектории № 16 $\Delta t_{23} = 7$ сут $< \Delta t_{23opt}$, то на некотором её участке, вблизи момента t_3 отлёта от Апофиса, $p(t) > 1$. В этом случае увеличение Δt_{23} и t_3 позволяет увеличить полезную массу. Это подтверждается характеристиками траектории № 19, у которой время ожидания $\Delta t_{23} = 93$ сут. На рис. 4 представлено изменение $p(t)$ для этой траектории, здесь $p(t) \leq 1$. Для траектории № 20 условие $p(t) \leq 1$ также выполняется.

Далее, на втором этапе анализа, выполнено уточнение полученных на первом этапе характеристик оптимальных траекторий. В окрестности значений граничного времени полученных траекторий выполнена оптимизация задачи на множестве уточнённых траекторий — методом покоординатного спуска. Получены оптимальные траектории № 16а, 19а, 20а. Характеристики этих траекторий приведены в табл. 2. Здесь $\Delta t_s = t_f - t_0$. Уточнение на втором этапе привело к некоторому уменьшению полезной массы.

Таблица 2. Характеристики траекторий № 16а, 19а, 20а

Номер траектории	Δt_s , сут	t_0	t_2	t_3	t_f	$V_{\text{хар}}$, км/с	m_f , кг	m_p , кг
16а	692	21.05.2019	24.04.2020	01.05.2020	12.04.2021	6,721	492	226
19а	691	21.05.2019	24.04.2020	22.07.2020	11.04.2021	6,624	509	245
20а	715	06.05.2020	02.03.2021	23.06.2021	21.04.2022	6,447	517	290

Кроме РН «Союз-ФГ», рассмотрено использование РН «Союз-2.16» и «Зенит» с РБ «Фрегат». В табл. 3 приведены значения масс m_f и m_p для траекторий № 16а, 19а, 20а при использовании этих РН. При этом для РН «Союз» при разгоне у Земли делается два включения РБ, для РН «Зенит» — три.

Таблица 3. Конечная и полезная масса КА для траекторий № 16а, 19а, 20а (после уточнения) с использованием разных РН

	«Союз-ФГ»			«Союз-2.1б»			«Зенит»		
m_0 (кг)	7130			8250			14000		
Номер траектории	16а	19а	20а	16а	19а	20а	16а	19а	20а
m_f , кг	492	509	517	604	624	611	1193	1233	1287
m_p , кг	226	245	290	301	325	362	700	748	877

Таблица 4. Изменение полезной массы со скоростью $V_{\text{вх}}$

$V_{\text{вх max}}$, км/с	12,8	12,9	13	13,1	13,2	13,26
m_p , кг, «Союз-ФГ»	275,3	280	284	287,5	289	290
m_p , кг, «Союз-2.1б»	343,6	349	354	358,0	361	362
m_p , кг, «Зенит»	839,6	852	861	869,8	876	877

Для полученных траекторий при возврате к Земле оценена скорость входа КА в атмосферу Земли $V_{\text{вх}}$: $V_{\text{вх}} = 12,74$ км/с, $12,32$ км/с, $13,26$ км/с, соответственно, для траекторий № 16а, 19а, 20а. Отметим, что в проекте Stardust $V_{\text{вх}} = 12,9$ км/с, в проекте Hayabusa $V_{\text{вх}} = 12,5$ км/с, в проекте OSIRIS-REx $V_{\text{вх}} = 12,2$ км/с, т.е. современные технологии позволяют приземлять КА при скорости $V_{\text{вх}}$ до $12,9$ км/с. Учитывая это, решена ещё задача (5) — полной четырёхпараметрической оптимизации с ограничением на скорость входа, $V_{\text{вх}} \leq 12,8$ км/с для траектории № 20а. В табл. 4 приведены результаты анализа зависимости полезной массы от $V_{\text{вх}}$ ($12,8-13,26$ км/с). Первая строка — ограничение по скорости входа в атмосферу Земли. Отметим, что при этом для РН «Союз-2.1б» и «Зенит» полезная масса всё же остаётся большей, чем сухая масса КА Stardust (300 кг).

2. ОРБИТАЛЬНОЕ ДВИЖЕНИЕ СПУТНИКА АПОФИСА

Здесь принят численный метод анализа с использованием дифференциальных уравнений астероидоцентрического движения КА с учётом трёх типов возмущений:

$$\frac{d\mathbf{r}}{dt} = \mathbf{V}; \quad \frac{d\mathbf{V}}{dt} = \mathbf{a}_0 + k_1 \cdot \mathbf{a}_1 + k_2 \cdot \mathbf{a}_2 + k_3 \cdot \mathbf{a}_3, \quad (3)$$

где t , $\mathbf{r}(x, y, z)$, $\mathbf{V}(V_x, V_y, V_z)$ — вектор состояния КА в невращающейся прямоугольной геоэкваatorialной системе координат $OXYZ$ с центром в центре масс астероида; \mathbf{a}_0 — центральное ускорение силы тяжести астероида, $\mathbf{a}_0 = -(\mu_A/r^3) \cdot \mathbf{r}$; \mathbf{a}_1 — суммарное возмущающее ускорение от притяжения удалённых небесных тел; \mathbf{a}_2 — возмущающее ускорение от несферичности астероида; \mathbf{a}_3 — возмущающее ускорение от ДСС; $k_1, k_2, k_3 = (0; 1)$ — коэффициенты для учёта возмущения.

Возмущающее ускорение от удалённых небесных тел вычисляется, для повышения точности расчёта, в модифицированной форме, близкой к методу Энке [Бэттин, 1966]. Векторы положения небесных тел относительно Солнца берутся из Эфемериды DE421; вектор положения Апофиса определяется по сайту Jet Propulsion Laboratory (JPL) (<http://ssd.jpl.nasa.gov/horizons.cgi>). Для определения возмущающего ускорения от несферичности астероида использована приближённая модель однородного вытянутого эллипсоида вращения [Дубошин, 1975; Ивашкин, 1998]. На данном этапе анализа полагаем, что астероид равномерно вращается вокруг малой оси \mathbf{a}_A с периодом P_A , причём эта ось вращения имеет постоянную ориентацию в пространстве по вектору \mathbf{L} , определяемому в эклиптической системе координат долготой λ_L и широтой β_L . При определении коэффициента тени K_{SH} используется цилиндрическая модель тени (не учитывается полутень), также учитывается несферичная форма астероида как эллипсоида вращения. Подробно расчёт этих возмущающих ускорений описан в работах [Ивашкин, Лан, 2016, 2017в; Lang, Ivashkin, 2016].

Для анализа движения спутника в окрестности Апофиса за начальную дату выбрана дата t_2 прилёта КА к Апофису для оптимальной траектории № 16 — 23 апреля 2020 г. Начальная орбита КА взята круговой с радиусом r_0 в диапазоне 0,5–2 км. По рекомендации АО «НПО Лавочкина», взято для основного КА: миделевы площади КА и солнечных панелей $F_{SC} = 5 \text{ м}^2$, $F_{SP} = 10 \text{ м}^2$, коэффициенты, учитывающие отражение: $C_{SC} = 1,4$, $C_{SP} = 1,1$, масса КА $m_{SC} = 600 \text{ кг}$. Для мини-спутника принято, что он имеет форму шара с диаметром $D = 40 \text{ см}$, массой $m = 10 \text{ кг}$, для него $C_{SC} = 1,4$. На основе результатов последних наблюдений Апофиса [Pravec et al., 2014], для анализа принято, что гравитационный параметр Апофиса $\mu_A = 1,8\text{--}2,86 \text{ м}^3/\text{с}^2$, средний радиус Апофиса $R_A = 160 \text{ м}$; длина малой и средней полуоси астероида равна $a_A = b_A$, удлинение $\alpha = c_A/a_A = (1,3; 1,5; 1,7)$. В основном варианте анализа астероид вращается вокруг малой оси \mathbf{a}_A с периодом $P_A = 30,56 \text{ ч}$, для оси вращения $\lambda_L = 250^\circ$ и $\beta_L = -75^\circ$. Для оценки «времени жизни» космического аппарата T оно определено условиями: $r/R_\odot \geq 1,1$ и $r \leq 10 \text{ км}$, где r/R_\odot — отношение расстояния от центра масс астероида до КА к расстоянию от центра масс астероида до подспутниковой точки на поверхности астероида-эллипсоида.

Основные результаты анализа орбитального движения ИСА приведены в статьях [Ивашкин, Лан, 2016, 2017в; Lang, Ivashkin, 2016]. Анализ показал, что главными параметрами, лимитирующими стабильность орбиты ИСА, являются начальный радиус и ориентация плоскости орбиты. Здесь под словом «стабильность» понимаем, что спутник Апофиса не сталкивается с Апофисом, и не улетает далеко от него. На низкой орбите ИСА главное возмущение — от несферичности Апофиса, а на орбите с большим радиусом — от ДСС. Возмущение от удалённых небесных тел обычно на несколько порядков меньше, чем от ДСС. Следует отметить, что в апреле 2029 г. будет тесное сближение Апофиса с Землёй (~38 000 км), и возмущение за счёт притяжения Земли резко увеличится, орбита спутника сильно изменится и КА улетит от астероида или ударится о его поверхность.

При анализе влияния несферичности Апофиса найдены два типа орбит: плоскость орбиты будет либо вращаться по долготе восходящего узла ДВУ Ω (в ротационном режиме) (рис. 5а), либо колебаться по наклону и узлу (в либрационном режиме) (рис. 5б). Выявлены две стационарные точки:

$$(1) \Omega_0 \approx 208,6^\circ, i_0 \approx 169,35^\circ; (2) \Omega_0 \approx 28,6^\circ, i_0 \approx 10,65^\circ. \quad (4)$$

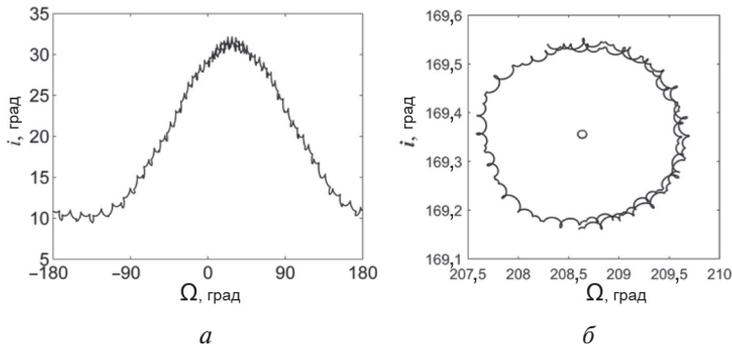


Рис. 5. Изменение углов Ω и i за 30 сут при учёте возмущения только от несферичности; $r_0 = 0,5$ км, $\mu_A = 1,8 \text{ м}^3/\text{с}^2$, $\alpha = 1,7$, вращение вокруг малой оси a : a — для случая вращательного движения по Ω ($\Omega_0 = 45^\circ$, $i_0 = 30^\circ$); b — для случая колебательного движения в координатах (Ω, i) у стационарной точки (1)

Они соответствуют экваториальной плоскости астероида. Первая точка указана в центре «окружности» на рис. 5б.

С другой стороны с ориентацией плоскости орбиты тесно связано влияние давления солнечного света. С точки зрения «времени жизни» КА рациональной начальной ориентацией плоскости орбиты КА является перпендикулярная к направлению на Солнце (угол между нормалью к орбите КА и направлением от КА к Солнцу $\beta \sim 0$). При этом меняются, в основном, угловые параметры орбиты, а её размеры остаются довольно стабильными. Отметим, что из-за вращения Апофиса вокруг Солнца, давление солнечного света вызывает прецессию орбиты КА, поворот плоскости орбиты КА по долготе восходящего узла Ω , так что если начальная ориентация плоскости орбиты КА выбрана должным образом, то угол β будет не только в начальный момент, но и за долгое время оставаться небольшим. Вследствие этого линейные параметры орбиты слабо меняются. Рассмотрены также случаи ориентации плоскости орбиты ИСА в «равновесных точках» (4), полученных из результатов анализа движения КА при учёте только несферичности астероида. Для них при $t_0 = 23$ апреля 2020 г.: $\beta_0 \approx 85^\circ$ и $\beta_0 \approx 95^\circ$, тогда с учётом влияния ДСС «равновесное» состояние вскоре будет нарушено, и КА быстро врежется в поверхность астероида.

Для случая совместного действия всех возмущений в табл. 5 приведено «время жизни» T основного КА на низкой орбите спутника Апофиса с начальным радиусом $0,5$ км, $i_0 = \Omega_0 = 90^\circ$ ($\beta \sim 10^\circ$) при разных значениях параметра μ_A и удлинения α . Анализ показал, что «время жизни» T основного КА при $r_0 = 0,5$ км можно обеспечить достаточно большим, свыше 30 сут.

Анализ движения основного КА на орбите спутника Апофиса с бóльшим радиусом, при $r_0 = 0,9\text{--}1,2$ км, показал, что в этом случае для всех рассмотренных значений μ_A и α , «время жизни» основного КА $T > 120$ сут. То есть увеличением r_0 можно использовать оптимальное время ожидания КА у Апофиса (90–120 сут), полученное во втором разделе.

Сделана также оценка необходимых затрат характеристической скорости для коррекции орбиты основного КА при необходимости обеспечения большего времени жизни, ~ 4 мес, при реализации оптимального времени ожидания КА у астероида до отлёта к Земле. Получено, что необходимая скорость коррекции мала, < 1 м/с.

Таблица 5. «Время жизни» основного КА на орбите спутника Апофиса при разных значениях μ_A и удлинения α для случая $r_0 = 0,5$ км, $e_0 = 0$, $i_0 = \Omega_0 = 90^\circ$ со всеми возмущениями

$\mu_A, \text{м}^3/\text{с}^2$	α	T , сут
1,8	1,3	136
	1,5	88
	1,7	38
2,86	1,3	71
	1,5	333
	1,7	60

В табл. 6 приведено «время жизни» T около Апофиса для мини-спутника с учётом всех трёх возмущений при разных значениях μ_A , удлинения α и начального радиуса орбиты r_0 ($i_0 = \Omega_0 = 90^\circ$). Видно, что в окрестности радиуса $r_0 = 1,5$ км, для рассмотренных значений μ_A , α , мини-спутник всегда может стабильно держаться у Апофиса в течение ~ 9 лет — до тесного сближения Апофиса с Землёй в 2029 г.

Таблица 6. «Время жизни» мини-аппарата на орбите спутника Апофиса при разных значениях μ_A , удлинения α и начального радиуса r_0 для случая $e_0 = 0$, $i_0 = \Omega_0 = 90^\circ$, со всеми возмущениями

$\mu_A, \text{м}^3/\text{с}^2$	α	r_0 , км	T
1,8	1,3–1,5	1,5	~ 9 лет
		0,9	$\sim 1,2$ года
	1,7	1,5	~ 9 лет
		2,0	~ 80 сут
2,86	1,7	1,0	~ 2 года
		1,5	~ 9 лет
	1,3–1,7	2,0	~ 9 лет
		2,5	~ 90 сут

На рис. 6а (см. с. 185) во вращающейся плоскости, перпендикулярной направлению с астероида на Солнце, приведена картина движения миниспутника с 2020 по 2029 г. для варианта $\mu_A = 1,8 \text{ м}^3/\text{с}^2$, $\alpha = 1,7$, $r_0 = 1,5$ км, $i_0 = \Omega_0 = 90^\circ$. Отлёт КА от астероида происходит вблизи показанной на рис. 6а точки P_0 сближения Апофиса с Землёй до расстояния ~ 38 тыс. км, 13.04.2029. Точки P_n ($n = 1, 2, 3$) на дуге отлёта соответствуют времени полёта n суток после сближения в точке P_0 . Стрелки E_2, M_2 показывают на Землю и Луну через 2 сут после отлёта. На рис. 6б дано относительное расстояние r/R_\odot для данного варианта.

Более подробный анализ при варьировании начального радиуса орбиты r_0 показал, чт.е. некоторый диапазон значений радиуса r_0 , для которого, при начальной ориентации орбиты, перпендикулярной к направлению на Солнце, миниспутник в течение 9 лет (2020–2029) остаётся в окрестности астероида. При $\alpha = 1,5–1,7$, $\mu_A = 1,8 \text{ м}^3/\text{с}^2$, $P \approx 30$ ч для этого диапазона $r_0 \approx 1,3–1,6$ км. Для более тяжёлого астероида, $\mu_A = 2,86 \text{ м}^3/\text{с}^2$, при $\alpha = 1,7$ указанный диапазон значений радиуса орбиты миниспутника составляет $r_0 \approx 1,4–2$ км.

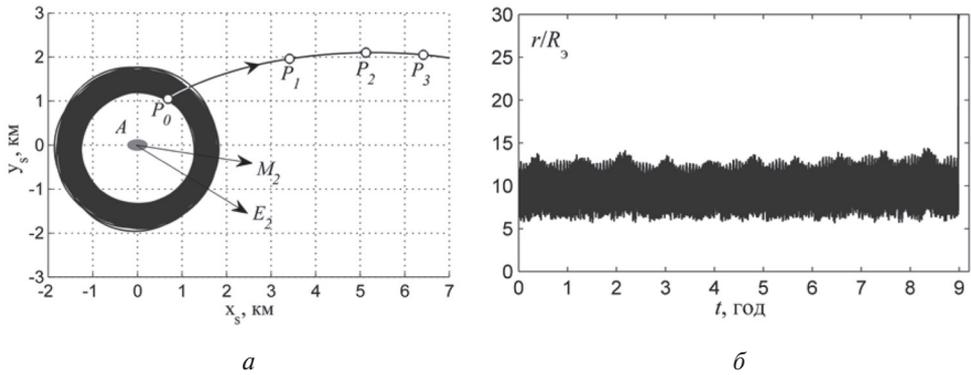


Рис. 6. Движение миниспутника около Апофиса в течение 2020–2029 гг. для случая $\mu_A = 1,8 \text{ м}^3/\text{с}^2$, $\alpha = 1,7$, $r_0 = 1,5 \text{ км}$, $i_0 = \Omega_0 = 90^\circ$, со всеми возмущениями

Для анализа надёжности результатов варьировался также ряд других основных характеристик, в частности: масса, удлинение, параметры вращения астероида, масса и параметры начальной орбиты КА (её плоскость, начальная большая полуось, эксцентриситет), начальная дата полёта. Основные результаты анализа, представленные выше, сохраняются в этих случаях.

ВЫВОДЫ

В работе рассмотрена методика определения оптимальных траекторий экспедиции КА для полёта от Земли к астероиду, пребывания КА у астероида и последующего возвращения КА к Земле с помощью ДУБТ, включая формирование орбиты искусственных спутников астероида. Разработанная методика является двухэтапной. Она максимизирует полезную массу межпланетных траекторий КА для данной экспедиции. Разработана также методика определения характеристик орбитального движения спутника астероида с учётом возмущений от небесных тел, несферичности астероида, давления солнечного света и собственного вращения астероида вокруг малой оси. Приведены численные результаты для экспедиции к астероиду Апофис. Построены и исследованы оптимальные траектории для экспедиции Земля–Апофис–Земля с суммарной продолжительностей до 2 лет при полёте в 2019–2022 гг. Показано, что существует принципиальная возможность реализации космической экспедиции Земля–Апофис–Земля с использованием обычных двигателей большой тяги и РН «Союз-ФГ», «Союз-2.1б», «Зенит». Определены характеристики орбит движения основного КА и мини-аппарата как спутников Апофиса. Выявлены параметры начальной круговой орбиты (ориентация и радиус), при которых «время жизни» спутника Апофиса будет большим: более месяца для основного КА на низкой орбите и несколько лет, до сближения с Землёй в 2029 г., для мини-спутника. При этом «оптимальной» является начальная ориентация плоскости орбиты КА, перпендикулярная к направлению на Солнце.

ЛИТЕРАТУРА

- [Аббасов, 2014] *Аббасов М. Э.* Методы оптимизации: учеб. пособие. СПб.: Изд-во «ВВМ», 2014. 64 с.
- [Бэттин, 1966] *Бэттин Р.* Наведение в космосе / пер. с англ. М.: Машиностроение, 1966. 448 с.
- [Дубошин, 1975] *Дубошин Г. Н.* Небесная механика: Основные задачи и методы. М.: Наука, 1975. 799 с.
- [Ивашкин и др., 2013] *Ивашкин В. В., Крылов И. В., Лан А.* Оптимальные траектории для экспедиции КА к астероиду Апофис с возвращением к Земле // *Астрономический вестник*. 2013. Т. 47. № 4. С. 361–372.
- [Ивашкин, 1998] *Ивашкин В. В.* Модель орбитального движения КА вблизи ядра кометы. I: препринт. М.: ИПМ им. М. В. Келдыша, 1998. № 60. 32 с.
- [Ивашкин, Лан, 2016] *Ивашкин В. В., Лан А.* Анализ орбитального движения космического аппарата вокруг астероида Апофис // *Доклады Академии наук*. 2016. Т. 468. № 4. С. 403–407.
- [Ивашкин, Лан, 2017а] *Ивашкин В. В., Лан А.* Анализ оптимальности траекторий экспедиции Земля-астероид-Земля: препринт. М.: ИПМ им. М. В. Келдыша, 2017. № 113. 25 с. doi: 10.20948/prepr-2017-113.
- [Ивашкин, Лан, 2017б] *Ивашкин В. В., Лан А.* Определение и анализ оптимальных космических траекторий для организации экспедиции Земля–Апофис–Земля с применением двигательных установок большой тяги // *Космонавтика и ракетостроение*. 2017. Вып. 5(98). С. 63–71.
- [Ивашкин, Лан, 2017в] *Ивашкин В. В., Лан А.* Анализ орбитального движения спутника астероида Апофис // *Космические исследования*. 2017. Т. 55. № 4. С. 268–277.
- [Ильин, Кузмак, 1976] *Ильин В. А., Кузмак Г. Е.* Оптимальные перелёты космических аппаратов. М.: Наука, 1976. 744 с.
- [Панченко, 2007] *Панченко Т. В.* Генетические алгоритмы: учебно-методическое пособие / под ред. Ю. Ю. Тарасевича. Астрахань: Изд. дом «Астраханский ун-т», 2007. 87 с.
- [Соболь, Стагинков, 1981] *Соболь И. М., Стагинков Р. Б.* Выбор оптимальных параметров в задачах со многим критериями. М.: Наука, 1981. 110 с.
- [Ajluni et al., 2015] *Ajluni T., Everett D., Linn T.* et al. OSIRIS-REx, returning the asteroid sample // *IEEE Aerospace Conf.* 2015. P. 1–15.
- [Atkins et al., 1997] *Atkins K. L., Brownlee D. E., Duxbury T.* et al. STARDUST: Discovery's InterStellar dust and cometary sample return mission // *IEEE Aerospace Conf.* 1997. 1997. V. 4. P. 229–245.
- [Brownlee et al., 2003] *Brownlee D. E., Tsou P., Anderson J. D.* et al. Stardust: Comet and interstellar dust sample return mission // *J. Geophysical Research*. 2003. V. 108. No. E10. P. 8111. doi: 10.1029/2003JE002087.
- [Lang, Ivashkin, 2016] *Lang A., Ivashkin V. V.* Dynamics of Spacecraft Orbital Motion around Asteroid Apophis // *Intern. Astronautical Congress IAC-2016*. Guadalajara, Mexico, 26–30 September 2016: Proc. 2016. Paper IAC-16-C1,6,2,x33922. 12 p.
- [Nocedal, Wright, 2006] *Nocedal J., Wright S. J.* Numerical Optimization. 2nd ed. Springer, 2006.
- [Pravec et al., 2014] *Pravec P., Scheirich P., Durech J.* et al. The tumbling spin state of (99942) Apophis // *Icarus*, 2014. V. 233. P. 48–60.
- [Sandford, 2011] *Sandford S. A.* The Power of Sample Return Missions-Stardust and Hayabusa // *Proc. Intern. Astronomical Union*. 2011. V. 7. No. S280. P. 275–287.

INVESTIGATION OF TRAJECTORIES FOR THE EARTH – APOPHIS – EARTH MISSION TAKING INTO ACCOUNT THE CHOICE OF THE NEAR-ASTEROID STAY ORBITS

A. Lang¹, V. V. Ivashkin^{1,2}

¹ Bauman Moscow State Technical University

² Keldysh Institute of Applied Mathematics Russian Academy of Sciences

This paper investigates spacecraft (SC) trajectories for the mission to a potentially dangerous asteroid, staying there for some time and following returning its samples to Earth in order to study the asteroid and reduce the asteroid hazard. This paper develops a two-step method for constructing interplanetary trajectories of the Earth-asteroid and asteroid-Earth transfers, as well as a method for analyzing the asteroid's satellite orbits. Optimal interplanetary trajectories for the Earth – Apophis – Earth mission using high thrust engines are found by maximizing the spacecraft payload mass. The near Apophis orbital motions of the main SC and the mini-satellite with a radio beacon have been studied taking into account three types of perturbations: the gravitational effects from distant celestial bodies, the nonsphericity of Apophis and the solar radiation pressure.

Keywords: space mission to an asteroid, optimal interplanetary trajectories, asteroid's satellite orbital motion, asteroid Apophis, “lifetime” of asteroid's satellite

Ivashkin Vyacheslav Vasilievich — chief scientist, doctor of physical and mathematical sciences, professor, Ivashkin@Keldysh.ru

Lang Anqi — trainee, master, PhD candidate, seatu_angel@126.com

КОМПЛЕКС УПРАВЛЕНИЯ ТЕЛЕСКОПОМ ZEISS-1000 С ВОЗМОЖНОСТЬЮ ПРОВЕДЕНИЯ НАБЛЮДЕНИЙ В РЕЖИМЕ УДАЛЁННОГО ДОСТУПА

С. В. Крючков¹, И. А. Маслов², И. В. Николенко^{1,3}, Угольников О. С.²

¹ Институт астрономии Российской академии наук (ИНАСАН)

² Институт космических исследований Российской академии наук (ИКИ РАН)

³ Крымская астрофизическая обсерватория Российской академии наук (КраО)

Телескоп Zeiss-1000 в заводском исполнении включает в себя оптические и механические узлы и приводы, управление которыми осуществляется специализированной ЭВМ Vilati (Венгрия). Сложности в содержании и ремонте возникли из-за устаревшей элементной базы. Это послужило для работы по модернизации всей управляющей системы. Работы были проведены в несколько этапов. С появлением устройств на микроконтроллерах стало возможно дополнить функционально и осуществить возможность управления телескопом с использованием удалённого доступа и новой элементной базы. При этом обеспечивается полное управление куполом, а также контроль за погодой и состоянием неба.

Ключевые слова: телескоп, автоматизация, удалённое управление

ВВЕДЕНИЕ

В 80-е гг. XX в. телескоп Zeiss-1000 был установлен на Симеизской научной базе Астросовета АН СССР, расположенной в посёлке Голубой Залив, вблизи Симеиза, на высоте 368 м. Он был смонтирован в башне, построенной в 1926 г. для 1-метрового рефлектора английской фирмы Grabbe, и реконструированной под установку телескопа и 8-метрового купола предприятия Carl Zeiss Jena.

Однометровый зеркальный телескоп предприятия Carl Zeiss Jena представляет собой оптическую систему Ричи–Кретьена–Куде с зеркалом диаметром 1016 мм и фокусом 13,3 м (рис. 1, см. с. 189). В наблюдения используется только система Ричи–Кретьена. Оптический тракт системы фокуса Куде не используется. Телескоп смонтирован на английской монтировке, оснащённой приводами, позволяющими вращать его по часовому углу и склонению с четырьмя дискретными скоростями, от 1 до 120 град/мин. Отдельный двигатель используется для часового ведения. Положение телескопа определялось 24-разрядными абсолютными угловыми датчиками. Система управления базировалась на специализированной ЭВМ Vilati (Венгрия) и имела три пульта управления — главный, в фокусе Кудэ и в фокусе Кассегрена.

Крючков Сергей Владимирович — ведущий электроник, skruch@inasan.ru

Маслов Игорь Алексеевич — заведующий сектором, imaslov@iki.rssi.ru

Николенко Игорь Владимирович — младший научный сотрудник КраО, инженер по внедрению новой техники и технологий ИНАСАН, niki63@yandex.ru

Угольников Олег Станиславович — старший научный сотрудник, ougolnikov@gmail.com



Рис. 1. Однометровый зеркальный телескоп предприятия Carl Zeiss Jena с установленным на нём поляриметром ИКИ РАН

МОДЕРНИЗАЦИЯ

С первых лет эксплуатации телескопа появились проблемы с управляющей системой *Vilati*, проявившиеся в периодических отказах в работе. Было принято решение переделать систему управления с использованием персонального компьютера. В 1992–1993 гг. такая система была создана на компьютере на базе процессора Intel 80286 и модулей КАМАК. Однако в ней оставалось большое количество узлов от старой системы, которые работали ненадёжно. В 2000 г. она была заменена новой системой управления, разработанной в Институте астрономии РАН. В настоящее время система управления телескопом Zeiss-1000 включает в себя релейный шкаф от *Vilati*, персональный компьютер на базе процессора Intel Core2, с блоком интерфейсных плат и комплекс управляющих программ.

Много проблем было из-за неудовлетворительной работы цейссовских 24-разрядных абсолютных угловых датчиков, которые многократно подвергались ремонту. В 2007 г. были установлены новые, немецкой фирмы *Kübler*. Это абсолютный многооборотный энкодер с разрядностью 25 бит, что позволяет вводить координаты в компьютер с точностью не хуже 1 угл. с (реальная точность позиционирования ниже из-за механических характеристик телескопа). Связь энкодеров с компьютером осуществляется по последовательному интерфейсу RS-485.

Для выдачи управляющих сигналов и ввода информации с датчиков использованы адаптеры дискретного ввода-вывода для шины PCI фирмы ICPDAS.

Адаптеры имеют высокую нагрузочную способность по выходным линиям, что позволяет управлять реле без дополнительных согласующих устройств. Оптроны, стоящие на входах, позволяют подключать устройства, удалённые более чем на 20 м.

Блок пусковых реле и часового ведения оставлен от Vilati и модернизирован частично. В нём заменён генератор часового ведения. Для синхронизации времени используется GPS приёмник Trimble Accutime.

Для управления телескопом написана программа на языке Delphi. Программа работает под системами Windows9X, 2000, XP, Windows7 и имеет удобный пользовательский интерфейс (рис. 2). Программа осуществляет индикацию на мониторе всемирного и звёздного времени, координат наведения телескопа (часовой угол, прямое восхождение, склонение и зенитное расстояние). Опрос и индикация датчиков положения производится раз в 0,2 с. Оператор, нажимая соответствующие кнопки, может управлять движением телескопа с разными скоростями, включать часовое ведение, фокусировать телескоп, управлять движением купола. Предусмотрено управление телескопом и с пульта Кассегрена. Программа осуществляет блокировку движения телескопа в случае возникновения опасных ситуаций и информирует об этом оператора.

Система управления постоянно улучшается и превосходит по своим сервисным возможностям оригинальную систему на базе Vilati. В частности, реализованы ведение журнала наблюдений с фиксацией времени и координат телескопа в файл, автоматическое позиционирование телескопа по координатам [Крючков, Николенко, 2008; Крючков и др., 2009; Смирнов и др., 2001]. Сотрудником ИКИ РАН был создан программный модуль, синхронизировавший движение купола с телескопом.

В связи с невозможностью замены устаревших к сегодняшнему дню компьютера с адаптером дискретного ввода-вывода для шины PCI снова возникла необходимость очередного этапа модернизации системы управления. Основное отличие новой системы состоит в её распределении на несколько контроллеров, решающих локальные задачи и соединённых между собой и с центральным компьютером по сети Ethernet. Это позволит управлять телескопом с любого компьютера в локальной сети, а впоследствии сделает возможным удалённое управление через интернет. Другое преимущество новой системы — это уменьшение количества реле, что повышает надёжность системы.

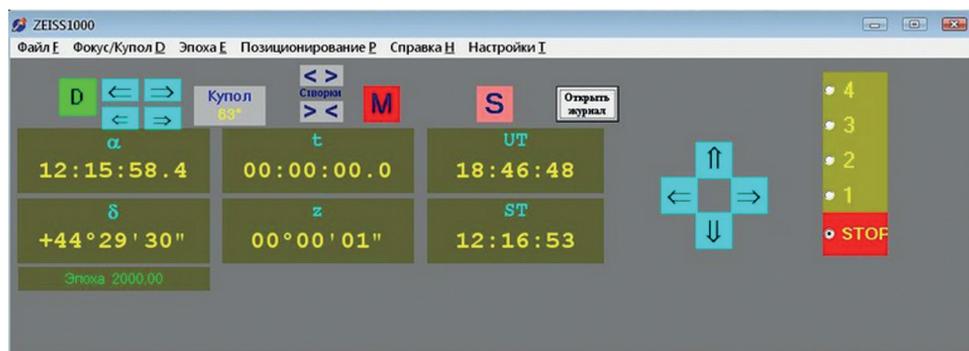


Рис 2. Пользовательский интерфейс управляющей программы

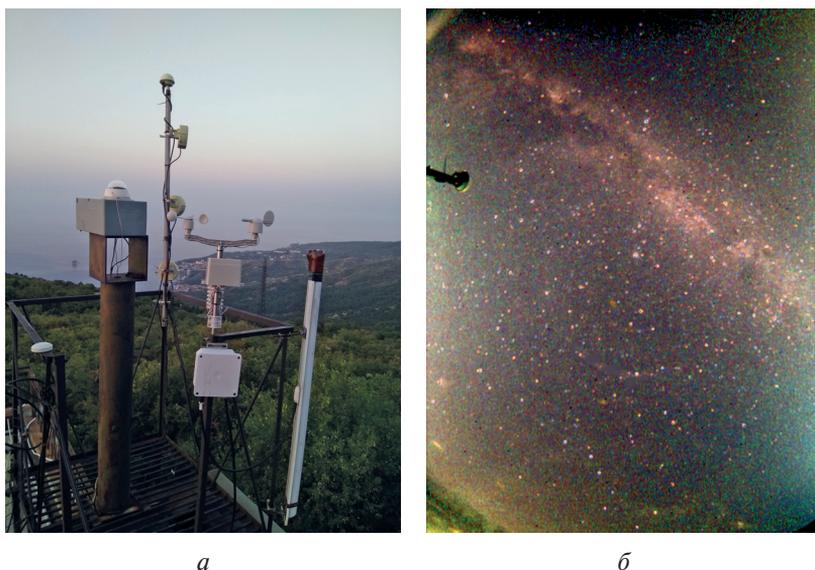


Рис 3. Широкоугольная камера (слева, вверху) и метеостанция (на переднем плане) (а); один из ночных снимков (б)

Оборудование, установленное на башне, рассматривается как периферия оптического инструмента для работы по удалённому доступу. В её состав входит метеостанция, что позволяет контролировать текущую скорость ветра и широкоугольная камера, контролирующая состояние неба. Планируется установить ещё отдельный датчик облачности. Широкоугольная камера позволяет также проводить мониторинг ярких звёзд и исследовать рассеяния света пылью в земной атмосфере. В камере используются цветной матричный фотоприёмник ST-i фирмы SBIG. Камера включена круглосуточно, ведёт съёмку в сумерки и ночью, автоматически переключая выдержку в зависимости от глубины погружения Солнца под горизонт. Информация записывается на сменные жёсткие диски, но доступна и по интернет [Угольников, Маслов, 2018; Ugolnikov et al., 2017].

Сигналы с них дают оператору принять решения о начале, продолжении или завершении наблюдений. При превышении установленного значения скорости ветра управляющая система телескопа может, после предупреждающей серии звуковых сигналов, сама сможет закрыть створки купола и остановить движение телескопа. Для визуального контроля за действиями механизмов в подкупольном пространстве имеется подвижная IP-камера.

ЗАКЛЮЧЕНИЕ

Созданный комплекс позволяет дистанционно проводить астрономические наблюдения, осуществляя контроль за всеми системами телескопа и купола. Для увеличения надёжности функционирования в экстремальных случаях планируется автоматизировать функцию закрытия купола и отключение подачи электроэнергии на телескоп.

ЛИТЕРАТУРА

- [Крючков, Николенко, 2007] *Крючков С. В., Николенко И. В.* Комплекс «Телескоп Цейс-1000 НИИ КРАО» (п. Голубой Залив, гора Кошка) и перспективы его развития // *Околосемная астрономия – 2007: материалы междунар. конф. 2008.* С. 277–281.
- [Крючков и др., 2009] *Крючков С. В., Зенькович А. Д., Николенко И. В., Неяченко Д. И.* 1-метровые телескопы на горе Кошка // *Изв. Крым. астрофиз. обсерватории* 2009. Т. 104. № 6. С. 188–192.
- [Смирнов и др., 2001] *Смирнов М. А., Барабанов С. И., Крючков С. В., Зенькович А. Д., Неяченко Д. И., Николенко И. В., Гафтонюк Н. М.* Комплекс 1-м телескопа в Симеизской обсерватории для наблюдений небесных тел в околосемном космическом пространстве // *Околосемная астрономия XXI века: сб. тр. конф. М.: ГЕОС, 2001.* С. 275–282.
- [Угольников, Маслов, 2018] *Угольников О. С., Маслов И. А.* Исследования фонового стратосферного аэрозоля с помощью многоцветных широкоугольных измерений фона сумеречного неба // *Космические исследования.* 2018. Т. 56. № 2. С. 97–104.
- [Ugolnikov et al., 2017] *Ugolnikov O. S., Galkin A. A., Pilgaev S. V., Roldugin A. V.* Noctilucent Cloud Particle Size Determination based on Multi-Wavelength All-Sky Analysis // *Planetary and Space Science.* 2017. V. 146. P. 10–19.

CONTROL SYSTEM OF ZEISS-1000 TELESCOPE MAINTAINING THE REMOTE OBSERVATIONS REGIME

S. V. Kruchkov¹, I. A. Maslov², I. V. Nikolenko^{1,3}, O. S. Ugolnikov²

¹ Institute of Astronomy of the Russian Academy of Sciences (INASAN)

² Space Research Institute of the Russian Academy of Sciences (IKI RAN)

³ Crimean astrophysical Observatory, Russian Academy of Sciences (CrAO)

Original equipment of Zeiss-1000 telescope includes optical and mechanical units controlled by special Vilati computer (Hungary). Old element base causes difficulties while maintaining and controlling the device. This system was generally updated. The work was done in several stages. Micro-controlling devices allow to run the remote controlling system of the telescope on the new element base. This includes the dome control and weather and sky quality survey.

Keywords: telescope, automation, remote control

Kryuchkov Sergey Vladimirovich — leading engineer, skruch@inasan.ru

Maslov Igor Alekseevich — head of sector, imaslov@iki.rssi.ru

Nikolenko Igor Vladimirovich — junior researcher CrAO, implementation engineer new techniques and technologies INASAN, niki63@yandex.ru

Ugolnikov Oleh Stanislavovich — Senior Reseacher, ougolnikov@gmail.com

АСТЕРОИДНО-КОМЕТНАЯ ОПАСНОСТЬ: ПОТЕНЦИАЛЬНАЯ УГРОЗА

С. А. Нароенков, Б. М. Шустов

Институт астрономии Российской академии наук (ИНАСАН)

В статье представлен краткий обзор современного состояния исследований по проблеме астероидно-кометной опасности (АКО) в России и мире. Приводятся сведения о проектах, связанных с обнаружением и мониторингом опасных астероидов, сближающихся с Землей. Кратко обсуждаются специальные космические проекты, направленные на исследование и решение проблемы АКО.

Ключевые слова: околоземные астероиды, астероидно-кометная опасность

Существование современной цивилизации подвержено ряду угроз природного характера. Часть этих угроз носит всеобщий, глобальный характер и ставит под сомнение дальнейшее развитие и само существование цивилизации в том виде, которого она достигла на протяжении последних пяти тысяч лет. Значительное место в этом списке занимает астероидно-кометная опасность (АКО) — опасность столкновения Земли с космическими телами астероидной и кометной природы. На земной поверхности, дне океанов и морей обнаружено множество кратеров ударного происхождения с размерами от нескольких сотен метров до 200–300 км. Такие кратеры образовались в результате столкновения Земли с космическими телами с размерами от десяти метров и более. Падение на Землю небольшого тела размером более 50 м может вызвать катастрофические разрушения на площади во многие тысячи квадратных километров. Даже скромные по астрономическим меркам события, подобные падению Челябинского метеорита [Емельяненко и др., 2013], способны вызвать довольно существенный материальный урон, а также иметь огромный общественный резонанс. Падения тел размером более 50 м происходят редко, но человечество должно быть готовым к реагированию не на уже свершившееся событие, а на существенную угрозу такого события.

Проблеме астероидно-кометной опасности, т.е. угрозе столкновению опасных астероидов и комет с Землей, в мире начали уделять серьёзное внимание около 30 лет назад. В 1992 г. Конгресс США поручил НАСА организовать программу обнаружения более 90 % околоземных астероидов диаметром более 1 км. С 1998 г. на постоянной основе в США заработала служба по обнаружению опасных небесных тел — Spaceguard Survey, которая к концу 2009 г. эту задачу выполнила.

По своей сути проблема астероидно-кометной опасности — комплексная. Выделяют три её основные составляющие:

- 1) обнаружение и мониторинг опасных объектов, прогноз движения и определение вероятности столкновения в будущем;
- 2) оценка последствий столкновения опасного объекта с Землей;
- 3) противодействие и уменьшение ущерба от столкновения.

Нароенков Сергей Александрович — старший научный сотрудник, кандидат физико-математических наук, snaroenkov@inasan.ru

Шустов Борис Михайлович — научный руководитель института, доктор физико-математических наук, член-кор. Российской академии наук, профессор, bshustov@inasan.ru

Задачи обнаружения, мониторинга и определения вероятности столкновения активно решаются в мире. Для обнаружения опасных астероидов, сближающихся с Землёй (АСЗ), созданы специальные широкоугольные телескопы и системы мониторинга. Больше всех продвинулись в направлении решения задачи мониторинга околоземного пространства такие государства как США и Евросоюз, поскольку в этих странах программы решения проблемы астероидно-кометной опасности системно выполняются на основе государственной поддержки.

1. ОБНАРУЖЕНИЕ ОПАСНЫХ АСТЕРОИДОВ И КОМЕТ

Общее количество открытых астероидов, сближающихся с Землёй, уже превысило 17 500 объектов (<http://minorplanetcenter.org>). Общее количество потенциально опасных астероидов с размерами от 140 м до 1 км оценивается более чем в 20 000 штук, а опасных астероидов с размерами от 50 до 140 м — в 200 000 штук. [NASA..., 2017]. Количество тел декаметрового размера (более 10 м) на орбитах, допускающих столкновение с Землёй, многие миллионы. Год от года количество известных объектов увеличивается благодаря работе специальных телескопов и сетей мониторинга околоземного пространства (рис. 1).

За обнаруженными телами ведётся постоянное слежение (мониторинг), а для изучения наиболее потенциально опасных тел проводятся специальные программы по изучению их свойств. Самые мелкие (декаметровые) тела столь многочисленны и трудно наблюдаемы, что их наблюдения возможны только в околоземном пространстве. Постоянный мониторинг таких тел невозможен. К настоящему моменту обнаружена лишь ничтожная доля (~0,001 %) всех околоземных объектов диаметром более 10 м.

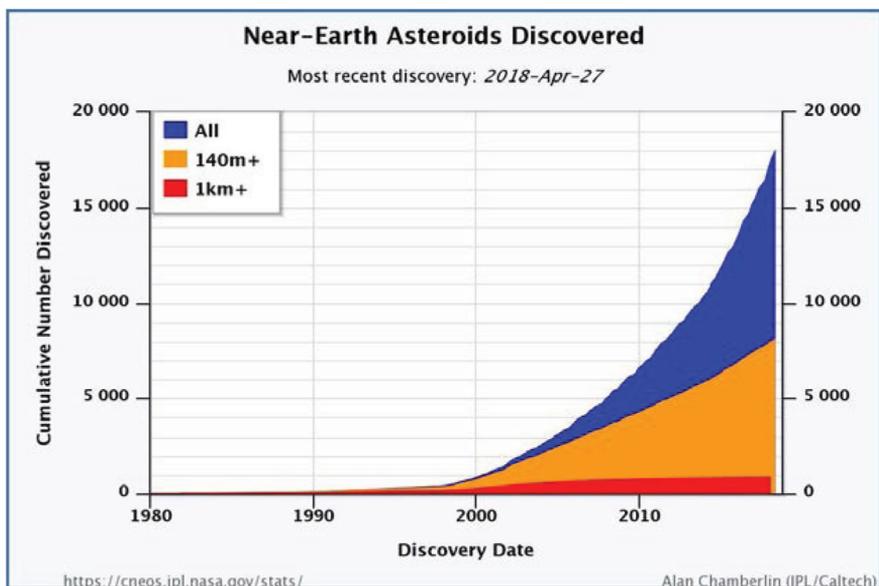


Рис. 1. Темп открытия астероидов, сближающихся с Землёй

Крупные астероиды диаметром более 140 м могут быть обнаружены заблаговременно (не менее чем за 30 дней до возможного столкновения). Для полного заблаговременного обнаружения таких тел нужны широкоугольные телескопы с диаметром главного зеркала от 1,5 м, такие как, например, широкоугольный телескоп Pan-STARRS (PS1) с диаметром главного зеркала 1,8 м и полем зрения 3° . Максимальная проникающая величина у телескопа PS1 составляет 22,5 звёздной величины за время экспозиции 30 с, что позволяет открывать астероиды размером в 140 м на расстоянии в 1 а. е.

Одной из наиболее эффективных программ по обнаружению околоземных астероидов является проект Catalina Sky Survey. Обзор проводится на обсерватории Steward Аризонского университета. Основными инструментами проекта являются 1,5-метровый обзорный телескоп с полем зрения 5 кв. град, 0,7-метровый телескоп системы Шмидта с 16-мегапиксельной ПЗС-камерой и полем зрения 19,4 кв. град, и 1-метровый узкопольный телескоп с полем зрения 0,3 кв. град для мониторинга (follow-up) вновь открытых астероидов.

В России в Институте солнечно-земной физики СО РАН для обнаружения околоземных астероидов создан телескоп АЗТ-33ВМ. Его параметры лишь немного уступают параметрам телескопов Pan-STARRS. При поле зрения около 3° и диаметре главного зеркала 1,6 м этот телескоп будет способен обнаруживать объекты 24-й звёздной величины за 2-минутную экспозицию. Телескоп был изготовлен на предприятии ЛОМО. Опытная эксплуатация первого в России широкоугольного телескопа, предназначенного для скоростного обзора неба с целью решения проблем астероидно-кометной опасности, начата в июне 2016 г.

Для обнаружения декаметровых тел в околоземном пространстве используются широкоугольные телескопы малой апертуры (диаметром от 40 см и более), собранные в мульти-апертурные системы. Характеристики таких мульти-апертурных систем позволяют осмотреть всю видимую часть небесной сферы 2–3 раза за один сеанс наблюдений. К таким системам в первую очередь стоит отнести проект ATLAS. Проект ATLAS (Asteroid Terrestrial-Impact Last Alert System) включает в себя восемь 0,5-метровых телескопов, оснащённых 110-мегапиксельными ПЗС-камерами. Главное достоинство телескопов проекта ATLAS — большое поле зрения, равное $7,4^\circ$. Максимальная видимая звёздная величина, доступная для телескопа, — 20 звёздная величина за 30-секундную экспозицию. Два телескопа способны провести обзор всей небесной сферы дважды за ночь. Предполагается, что телескоп сможет обнаруживать тела размером около 50 м не позднее, чем за 7 дней до столкновения, 140-метровые тела будут обнаружены за 3 недели до столкновения. Первые два узла системы введены в работу в 2015 и 2017 гг.

Ещё одним высокопродуктивным инструментом для открытия околоземных астероидов является проект космической обсерватории NEO-WISE. Космический аппарат WISE (Wide-Field Infrared Survey Explorer) с телескопом диаметром 40 см был запущен в космос в 2009 г. с целью проведения широкоугольных обзоров неба в ИК-диапазоне. В 2011 г. в связи с израсходованием хладагента, миссия WISE по обзору всего неба в ИК-диапазоне была прекращена, и начата новая программа исследований NEO-WISE для обнаружения и исследования околоземных астероидов. На начало 2018 г. общее количество открытых и исследованных околоземных астероидов составило 256, из них 47 потенциально-опасных и 28 комет.



Рис. 2. Графическое изображение системы мониторинга околоземного пространства INASAN NEO Finder

Доля наблюдений АСЗ, проведённых на российских обсерваториях, не так велика, как хотелось бы. Следует отметить два проекта, которые в некотором роде, проводят мониторинг околоземного пространства — ISON и MASTER. В проекте ISON [Еленин и др., 2017] для поиска астероидов работает до четырёх телескопов апертурой 40 см. К настоящему моменту в проекте ISON получено более 1 млн измерений. Открыто 1750 астероидов главного пояса, 17 АСЗ, 8 комет, 20 троянцев Юпитера, 4 объекта семейства Хильды, 4 объекта семейства Кентавра.

Для проекта MASTER основное предназначение — наблюдения быстрых оптических транзиентов, связанных, например, с космическими гамма-вспышками [Gorbovskey et al., 2013]. Открытия астероидов являются побочным продуктом обзора всего неба.

В Институте астрономии РАН разработана концепция обзорной системы мониторинга околоземного пространства INASAN NEO Finder (рис. 2) на основе кластера широкоугольных телескопов VT-78d. Параметры телескопа VT-78d: апертура — 25 см, поле зрения — 78 кв. град. Система из восьми телескопов будет способна в среднем за 20 ч до столкновения с Землёй обнаружить декаметровые тела размером от 10 м. Это позволит не позднее, чем за 4 ч до столкновения определить их орбиту и точку входа в атмосферу.

Дополнительно, перспективный кластер способен решать дополнительные научные задачи: наблюдения источников гравитационных волн, гамма вспышки и т.п.

Общепризнанно, что эффективное решение проблемы астероидно-кометной опасности возможно только в широкой международной кооперации. Недавно под эгидой ООН была создана кооперация астрономических институтов и других организаций — International Asteroid Warning Network (Международная сеть предупреждения об астероидной опасности). В 2017 г.

в рамках IAWN был проведён наблюдательный эксперимент по обнаружению и последующему изучению астероида 2012 TC₄. На протяжении 4 мес астрономы из Германии, Израиля, Италии, Канады, Колумбии, Нидерландов, России, США, Украины, ЮАР и Японии наблюдали астероид с помощью наземных и космических телескопов в оптическом и радио диапазонах с целью изучения его орбиты, формы, особенностей вращения и состава. Полученный в рамках кампании богатый наблюдательный материал позволил, в частности, точно рассчитать орбиту TC₄ и исключить возможность столкновения объекта TC₄ с Землёй при последующих сближениях. Со стороны России и Украины наблюдения астероида 2012 TC₄ были проведены в Терскольской обсерватории на 2-метровом телескопе Цейсс-2000. Наблюдательная кампания показала готовность учёных разных стран к участию в серьёзном международном сотрудничестве по предотвращению потенциальной угрозы для Земли со стороны околоземных объектов и продемонстрировала работоспособность Международной сети предупреждения об астероидной опасности (IAWN).

Челябинское событие ещё раз показало, что опасные тела могут приближаться к Земле не только с ночной стороны, но и со стороны Солнца. В 2018 г. НАСА была подготовлена карта вхождения крупных болидов в атмосферу Земли за период с 1988 по 2018 г. (рис. 3).

Количество болидов входящих в атмосферу Земли днём и ночью, примерно одинаковое. В работе [Шустов и др., 2017] рассчитано, что в среднем 600 астероидов в год могут сближаться с Землёй со стороны Солнца. Для обнаружения таких объектов Институт астрономии РАН разработал проект космической системы обнаружения и определения параметров движения опасных астероидов и комет, летящих со стороны Солнца (СОДА — Система обнаружения дневных астероидов). Главной целью космического комплекса является своевременное обнаружение опасных небесных тел и выдача предупреждения о возможном столкновении с Землёй не позднее, чем за 4 ч до возможного столкновения.

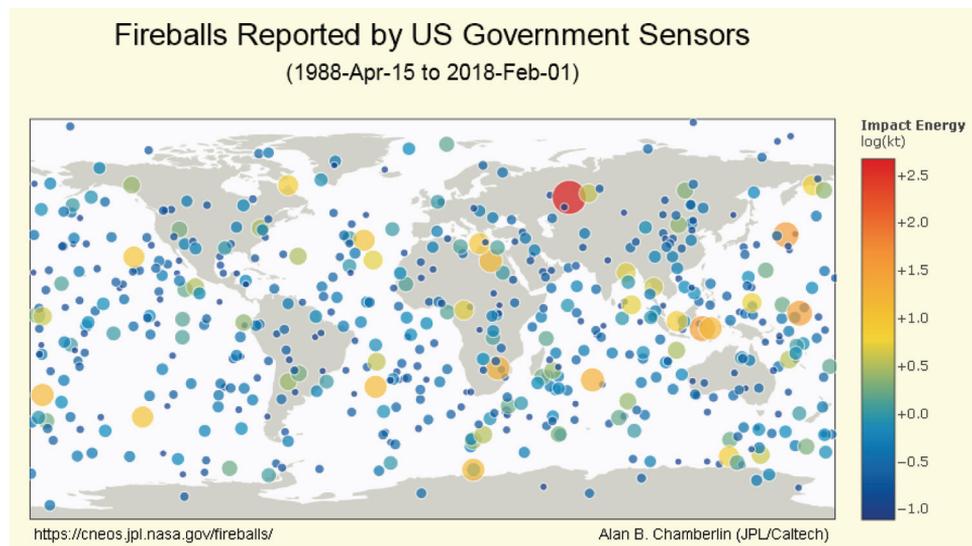


Рис. 3. Болидные события, зафиксированные в период 1988–2018 гг.

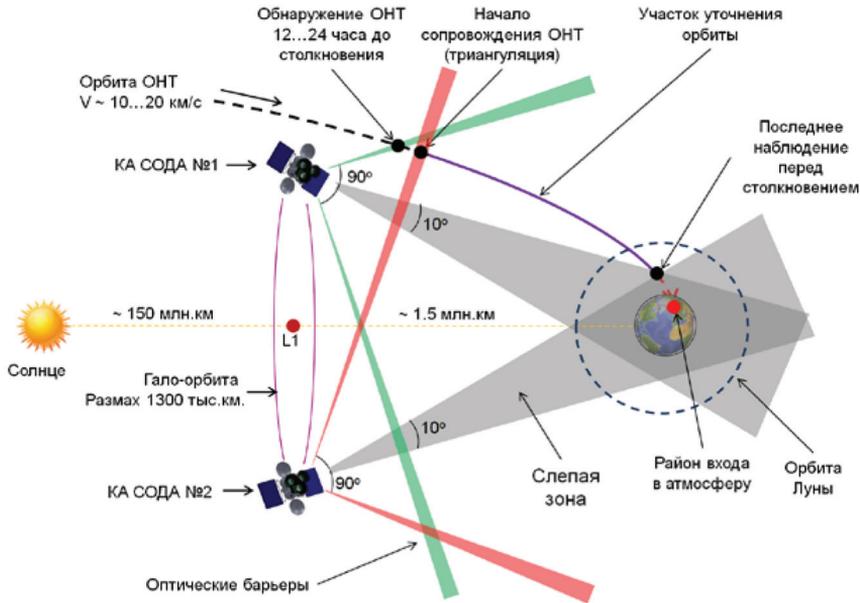


Рис. 4. Космическая система обнаружения и определения параметров движения опасных астероидов и комет (СОДА)

В проекте СОДА планируется разместить на одном (двух) космических аппаратах в окрестностях точки либрации L1 системы «Земля – Солнце» несколько широкоугольных телескопов умеренной апертуры (25–50 см) для постоянного мониторинга части околоземного пространства, недоступного для наземных телескопов (рис. 4). Предлагаемая система в настоящее время не имеет действующих аналогов.

Основными преимуществами космической системы мониторинга околоземного пространства являются: возможность круглосуточной работы, меньший фон неба, обнаружение космических объектов, недоступных для наземных телескопов. К недостаткам космических систем обнаружения можно отнести меньшую надёжность работы всей системы и относительную дороговизну всего космического комплекса.

Идеальная система мониторинга околоземного пространства для обнаружения опасных астероидов и комет должна включать в себя как наземный, так и космический сегмент для наблюдений. Сегменты должны работать параллельно и дополнять друг друга.

2. ПРОБЛЕМА ОЦЕНКИ РИСКОВ И ПОСЛЕДСТВИЙ СТОЛКНОВЕНИЯ ОПАСНОГО ОБЪЕКТА С ЗЕМЛЁЙ

Надёжная оценка степени риска конкретного события (столкновения) — главный результат, который обязана дать наука, занимающаяся проблемой АКО. Этот вид риска может потребовать принятия незамедлительных и весьма затратных решений. Риск определяется двумя основными факторами — вероят-

ностью столкновения данным телом и тяжестью последствий столкновения. Несколько упрощая, можно сказать, что за надёжную оценку этих факторов риска отвечают фундаментальные науки: астрономия, в частности, небесная механика — за оценку вероятности конкретного события (столкновения) и геофизика вкупе с науками экономического и социального направлений — за оценку последствий столкновений.

Задача оценки последствий столкновения опасного тела с Землёй является довольно сложной. Для точной оценки последствий столкновения нужно проводить сложные математические расчёты с использованием суперкомпьютеров, возможно даже в течение нескольких дней. В некоторых случаях, когда опасное тело уже подлетает к Земле, такая длительность расчётов неприемлема. Выходом из такой ситуации может являться создание банка данных последствий столкновения опасных небесных тел с Землёй. В банке данных будут содержаться результаты моделирования по основным поражающим факторам (воздушная ударная волна, образование кратера, сейсмические эффекты, тепловое излучение, ионосферные возмущения) при столкновении астероида с Землёй с учётом разнообразных условий входа тела в атмосферу. Такие расчёты могут быть проведены заранее для каждого типа последствий по проверенным методам и моделям [Шувалов и др., 2017]. В качестве прототипа такой системы можно привести проект, реализуемый Институтом динамики геосфер РАН «Калькулятор последствий столкновений космических тел с Землёй» [Naroenkov et al., 2017].

Для координации работ по обнаружению опасных небесных тел, оценке рисков столкновения с Землёй и своевременного предупреждения в НАСА создан Центр по изучению околоземных объектов (CNEOS), а также Офис по планетарной защите (Planetary Defense Office). Подобный центр также создан при Европейском космическом агентстве — SSA-NEO Coordination Centre.

3. ПРОБЛЕМА ПРОТИВОДЕЙСТВИЯ И УМЕНЬШЕНИЯ УЩЕРБА

Разработка методов и средств противодействия астероидно-кометной опасности ведётся с осознания проблемы АКО. Методы воздействия на опасное космическое тело зависят от размеров тела и времени парирования угрозы. Методы противодействия разделяются на два типа: разрушение тела и изменение орбиты тела для увода тела со столкновительной орбиты.

При выборе способа противодействия и уменьшения ущерба от падения космического тела всегда надо руководствоваться принципом здравого смысла и экономической эффективности. Нет необходимости уводить с орбиты небольшое тело, если она падает в океан, либо в малонаселённой области. В некоторых случаях, если тело небольшого размера и времени на реагирование уже не остаётся, достаточно предпринять комплекс мероприятий, направленных на защиту и эвакуацию гражданского населения, а также на подготовку инфраструктуры к возможным последствиям столкновения. Использовать дорогостоящие методы воздействия необходимо только в том случае, если ущерб в результате падения тела окажется больше, чем необходимые меры противодействия.

Увод тела со столкновительной орбиты является более предпочтительным способом, но его применение требует длительного времени на изменение орбиты. Такой способ может быть использован заблаговременно, за несколько лет до столкновения опасного объекта с Землёй. На сегодняшний день среди методов изменения орбиты опасного тела рассматривают кинетические ударники, дистанционное гравитационное воздействие (гравитационный тягач), использование реактивных двигателей малой тяги, а также методы, связанные с истечением (испарением) вещества с поверхности опасного тела.

В качестве примера изучения эффективности использования кинетического ударника можно привести миссию DART (Double Asteroid Redirection Test) [Rivkin et al., 2018]. В этом эксперименте, осуществляемом NASA в рамках совместного с ESA проекта AIDA, предполагается изменение орбиты астероида-спутника в двойной системе астероида Didymos. Размеры астероида Didymos составляют примерно 800 м, его спутника — 150 м. Предполагается, что космический зонд — ударник размером 2,4 м и массой 500 кг столкнётся со спутником астероида на скорости 6 км/с, что приведёт к изменению орбиты спутника астероида. Ожидается, что изменение скорости в спутника на 1 % заметно изменит его орбиту. Космический зонд планируют запустить с декабря 2021 г. по май 2022 г., а столкновение произойдёт в октябре 2022 г., когда расстояние между астероидом и Землёй составит 11 млн км. Учёные планируют получить точные данные об изменении орбиты с помощью радарных и оптических наблюдений двойного астероида.

Ещё один проект по изменению орбиты опасного астероида с помощью испарения вещества с поверхности называется DE-STAR (Directed Energy System for Targeting of Asteroids and exploRation). В данном проекте изменение орбиты астероида планируется с помощью испарения вещества с поверхности, посредством нагрева сконцентрированным лазерным лучом. Предполагается, что возникающие реактивные силы, при условии длительного воздействия на астероид, смогут изменить орбиту астероида (<https://arxiv.org/abs/1601.03690>). Концепция космического корабля предполагает наличие солнечных батарей большой площади для выработки электроэнергии и питающей массив лазеров киловаттного класса. Предполагается, что воздействие на астероид лазера позволит разогреть локальный участок астероида до 3000°, что вызовет испарение вещества с нагретой поверхности и создание реактивной силы. В работе (<https://arxiv.org/abs/1601.03690>) был смоделирован процесс воздействия 20-кВт лазерного луча на астероид размером 325 м. Воздействие лазером моделировалось в течение 15 лет, и на 16 год сдвиг орбиты достиг двух радиусов Земли (12 800 км).

Ядерное воздействие можно рассматривать в двух вариантах: полное разрушение тела с помощью взрыва внутри астероида и ядерное воздействие с целью испарения вещества с поверхности астероида. Взрыв на поверхности астероида, как правило, не приведёт к его разрушению, так как большая часть энергии рассеется в окружающее пространство. Для того, что бы тело было разрушено полностью необходимо произвести заглублённый взрыв, что само по себе является очень трудной задачей. Миссии полётов к астероидам показали все трудности навигации и посадки на астероид. К тому же необходимо гарантировать, что астероид будет разрушен до такой степени, что осколки астероида не причинят вреда при входе в атмосферу.

Под эгидой ООН в 2013 г. создана «Консультативная группа по планированию космических миссий (Space Missions Planning Advisory Group — SMPAG).

Основной задачей группы является подготовка международного сотрудничества для обмена информацией, разработки возможности совместных исследований и миссий для предотвращения угрозы астероидно-кометной опасности. В рамках этой группы периодически проходят встречи представителей космических агентств. Деятельность Роскосмоса в данной группе пока малозаметна.

ЗАКЛЮЧЕНИЕ

Проблема астероидно-кометной опасности требует внимания и планомерного решения. Создание высокоэффективных систем мониторинга, обнаружение и исследование потенциально-опасных астероидов дадут в конечном итоге свои результаты и плоды. Национальные космические агентства США, Евросоюза, Японии приняли программы по решению проблемы астероидно-кометной опасности. Роскосмос в этом направлении работ пока отстаёт. В России есть предпосылки для создания системы мониторинга околоземного пространства, ведутся работы по разработке и созданию как наземных, так и космических средств обнаружения опасных небесных тел. Но такие работы ведутся, в основном, в инициативном порядке. Проблема астероидно-кометной опасности - многоплановая, и ее решение, по крайней мере, в условиях России, требует четкой координации и поддержки со стороны государства

ЛИТЕРАТУРА

- [Еленин и др., 2017] *Еленин Л. В., Круглый Ю. Н., Молотов И. Е., Инасаридзе Р. Я., Румянцев В. В., Рева И. В., Шильдкнехт Т., Перец Т. Э. Г., Перец Л. Ж. Э., Шмальц С. Е., Новичонок А. О., Тунгалаг Н., Кокина Т. Н., Терешина М. А.* Роботизированный астероидный обзор сети ISON, поиск АСЗ и комет из обоих полушарий Земли // Экологический вестник научных центров Черноморского экономического сотрудничества. 2017. № 4–3. С. 32–38.
- [Емельяненко и др., 2013] *Емельяненко В. В., Попова О. П., Чугай Н. Н., Шеляков М. А., Пахомов Ю. В., Шустов Б. М., Шувалов В., Бирюков Е. Е., Рыбнов Ю. С., Маров М. Я., Рыжлова Л. В., Нароенков С. А., Карташова А. П., Харламов В. А., Трубецкая И. А.* Астрономические и физические аспекты Челябинского события 15 февраля 2013 г. // *Астрон. вестн.* 2013. Т. 47. № 4. С. 262–277.
- [Шувалов и др., 2017] *Шувалов В. В., Светцов В. В., Артемьева Н. А., Трубецкая И. А., Попова О. П., Глазачев Д. О.* Астероид Апофис — оценка опасных последствий ударов подобных тел // *Астрон. вестн.* 2017. Т. 51. № 1. С. 51–66.
- [Шустов и др., 2017] *Шустов Б. М., Нароенков С. А., Ефремова Е. В.* О населенности околоземного пространства опасными небесными телами // *Астрон. вестн.* 2017. Т. 51. № 1. С. 38–43.
- [Gorbovskey et al., 2013] *Gorbovskey E. S., Lipunov V. M., Kornilov V. G., Belinski A. A., Kuvshinov D. A., Tyurina N. V., Sankovich A. V., Krylov A. V., Shatskiy N. I., Balanutsa P. V., Chazov V. V., Kuznetsov A. S., Zimnukhov A. S., Shumkov V. P., Shurpakov S. E., Senik V. A., Gareeva D. V., Pruzhinskaya M. V., Tlatov A. G., Parkhomenko A. V., Dormidontov D. V., Krushinsky V. V., Punanova A. F., Zalozhnyh I. S., Popov A. A., Burdanov A. Yu., Yazev S. A., Budnev N. M., Ivanov K. I., Konstantinov E. N., Gress O. A., Chuvalaev O. V., Yurkov V. V., Sergienko Yu. P., Kudelina I. V., Sinyakov E. V., Karachentsev I. D., Moiseev A. V., Fatkhullin T. A.* The MASTER-II network of robotic optical telescopes. First results. // *Astronomy Reports.* 2013. V. 57. Iss. 4. P. 233–286.

- [Naroenkov et al., 2017] *Naroenkov S. A., Glazachev D. O., Kartashova A. P., Turuntaev I. S., Svetsov V. V., Shuvalov V. V., Popova O. P., Podobnaya E. D.* Impact Effects Calculator. Orbital Parameters // European Planetary Science Congress. 2017. V. 11. EPSC2017-881-1.
- [NASA..., 2017] NASA report of the Near-Earth Object Science Definition Team. Update to Determine the Feasibility of Enhancing the Search and Characterization of NEOs. NASA, 2017. 238 p.
- [Rivkin et al., 2018] *Rivkin A. S., Cheng A. F., Stickle A. M., Richardson D. C., Barnouin O., Thomas C. A., Fahnestock E.* The Double Asteroid Redirection Test (DART): Overview and Update //49th Lunar and Planetary Science Conference. 2018. id.2055.

ASTEROID–COMET HAZARD: THE POTENTIAL THREAT

S. A. Naroenkov, B. M. Shustov

Institute of Astronomy of the Russian Academy of Sciences (INASAN)

A review of the current state of investigations problem of asteroid-comet hazard (ACH) is presented. Main attention is drawn to projects for detection and monitoring of dangerous asteroids. Special space projects are briefly discussed, aimed to investigate and solve the ACH problem.

Keywords: Near Earth asteroids, asteroid-comet hazard

Naroenkov Sergey Aleksandrovich — senior researcher, PhD, snaroenkov@inasan.ru

Shustov Boris Mikhailovich — scientific director, doctor of physical and mathematical sciences, corresponding member of RAS, bshustov@inasan.ru

АСТЕРОИДНАЯ ОПАСНОСТЬ В 2018 ГОДУ: ТЕКУЩИЕ РЕАЛИИ И ПРОБЛЕМЫ

В. Г. Поль, А. В. Симонов

Акционерное общество «Научно-производственное объединение им. С. А. Лавочкина» (АО «НПО Лавочкина»)

Рассматриваются текущие результаты обнаружения астероидов, сближающихся с Землёй на весну 2018 г. Подтверждаются выводы о преувеличенной опасности астероидов для человеческой цивилизации, необходимости продолжать мониторинг опасных объектов и в первую очередь изучения следствий, вытекающих из тесного сближения астероида с Землёй.

Ключевые слова: малые тела Солнечной системы, астероиды, астероидная опасность

Проблема астероидной опасности продолжает интересовать специалистов в этой области, и как самостоятельный объект научных и технологических исследований существует уже в течение свыше 27 лет. Однако с течением времени интерес мирового сообщества, возникший на рубеже XX и XXI столетий, к этой проблеме понемногу спадает, и постепенно она начинает уступать своё место прочим проблемам глобального масштаба, эффекты которых заметны и проявляются, по-видимому, в более быстрой временной шкале. Наглядным, но не единственным примером может служить всё убыстряющийся ход глобального потепления.

По ходу осознания самого существования астероидной опасности в работах коллектива ИНАСАН [Астероидная..., 2010] было чётко сформулировано, что в целом эта проблема представляет собой комплекс взаимосвязанных задач, объединяющий в себе три составные части. Этими частями являются: обнаружение и мониторинг объектов, опасно сближающихся с Землёй, выбор способа и средства противодействия прогнозируемому столкновению, и, наконец, доставка средства противодействия к угрожающему объекту. Сегодня, фактически, как и ранее, работа по проблеме продолжается лишь в первой части, и она представлена реально проводящимся и углубляющимся мониторингом космической обстановки в околоземном пространстве. Остальные же две части практически остаются на начальной стадии, так и не выйдя за рамки поисковых чисто теоретических работ.

В подтверждение сказанного, ещё раз обратимся к оценкам частот встречи астероидов с Землёй. Эти оценки в литературе хорошо известны, и, по-видимому, в рамках текущего десятилетия вполне установились [Harris, 2009]. Они показывают, что падения дециметровых объектов типа тунгусского метеорита (характерные размеры ≤ 50 м) можно ожидать лишь единожды в среднем за многие сотни (до тысячи) лет. Далее, удары от объектов следующего увеличенного размера, имеющих масштабы свыше ~ 100 м, чреватые серьёзной региональной катастрофой, можно ожидать в среднем один раз за десяток тысяч лет и более. Эти обстоятельства и объясняют постепенно охлаждающееся отношение общества и управляющих структур к проблеме астероидной опасности.

Поль Вадим Георгиевич, инженер 2 категории, кандидат технических наук, polvad@laspace.ru

Симонов Александр Владимирович, ведущий математик, кандидат технических наук, alex.simonov@laspace.ru

Что же касается собственно первой части проблемы — обнаружение и каталогизация популяции астероидов, сближающихся с Землёй (АСЗ) то здесь реальная работа по программе, принятой в США с 1997 г., сегодня успешно выполняется в рамках ведомства NASA (ответственное подразделение — JPL, NASA's Jet Propulsion Laboratory) и продолжается с нарастающими темпами. Результатом этой программы является создание и успешное развитие специализированных аппаратных средств наблюдения, мониторинга и учёта околоземных объектов класса АСЗ. Правда, нужно отметить, что аналогичные программы вслед за США пытаются начать и в Европейском союзе. Кроме того, слабые попытки подключиться к реальному мониторингу в кооперации с другими странами имеют место и в России. К сожалению, пока вклад как Европы, так и России в общий банк данных практически незначителен, и данные, полученные в США, составляют подавляющее большинство. Поэтому состояние астероидной проблемы в первой части можно исчерпывающе охарактеризовать работами, проведёнными в США.

Начавшийся в конце XX века начальный этап создания системы мониторинга был представлен несколькими программами оптических и радиолокационных наблюдений, выполнявшихся на различной инструментальной базе. При этом наибольший вклад внесла программа MIT, проводившаяся по проекту LINEAR, которая начиналась как демонстрация возможностей использования уже имевшихся систем слежения за ИСЗ, предназначенных для обнаружения и определения характеристик АСЗ. Основой проекта LINEAR являлось использование двух телескопов с метровой апертурой. Это позволило накопить опыт наблюдений специфических объектов и постепенно разработать требования к такой системе, целевым назначением которой должно являться решение конкретной задачи обнаружения и каталогизации популяции АСЗ с повышенной надёжностью и производительностью. В результате система LINEAR послужила дальнейшей основой для разработки проекта и последующего создания специализированных систем поиска и мониторинга АСЗ второго поколения.

В настоящее время такие средства второго поколения в основном представлены проектами CATALINA и PANSTARR. Первый из них представляет собой комплекс модернизированных инструментов, среди которых наиболее мощным является телескоп с главным зеркалом диаметром 1,5 м.

Второй же проект представляет собой особый инструмент, специально спроектированный с учётом специфики обзорных задач эффективного обнаружения предельно видимых малых тел Солнечной системы (PANSTARR). В этом проекте предусматривается одновременная работа четырёх телескопов диаметром 1,8 м и полем зрения 3 кв. град (каждый на одной монтировке). Кроме того, там же предполагается установка основных компонентов получения и обработки информации, поступающей с четвёрки телескопов, которые могут работать как автономно, а также и как составная апертура. Разработка проекта PANSTARR и его ввод в строй велась поэтапно с 2008 г., и полная версия инструмента PANSTARR вошла в строй примерно с 2010–2012 гг.

Изначально нижней границей размеров тел АСЗ был несколько произвольно определён эквивалентный диаметр 140 м. Однако сегодня принято считать, что достаточно эффективно задерживаются атмосферой малые тела, имеющие характерные размеры менее 100 м. Поэтому с начала XXI в. в качестве астероидов, сближающихся с Землёй, было принято решение учитывать объекты с характерными размерами, начинающиеся уже со 100 м.

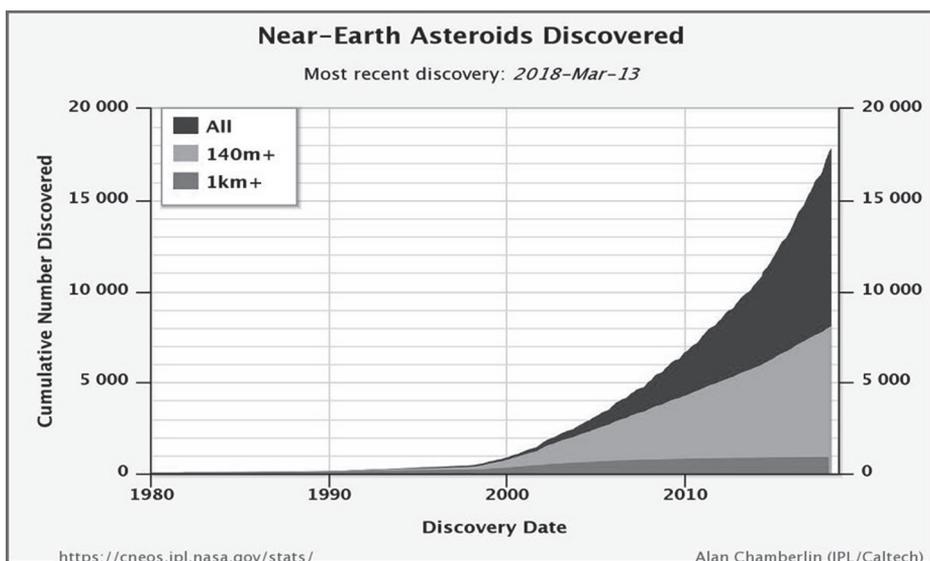


Рис. 1. Динамика роста основных категорий опасных астероидов

На рис. 1 показана интегральная динамика роста числа обнаруженных и каталогизированных астероидов, представленных тремя группами с разбивкой по характерным размерам: с эквивалентным диаметром от 1 км и выше, от 140 м до 1 км, и общее число обнаруженных объектов всех размеров. (<https://cneos.jpl.nasa.gov>).

Эти данные наглядно показывают, что число уже учтённых объектов с диаметрами километр и выше вполне приблизилось к насыщению. Также заметно, что количество найденных объектов с диаметрами от 140 м до километра всё ещё растёт, однако уже совсем не так быстро, как полное число обнаруженных объектов. Поэтому можно предположить, что объём второй группы астероидов составляет практически половину от общего количества всех обнаруженных объектов. Этот факт объясняется непрерывно повышающимся темпом регистрации весьма малых (дециметровых) астероидов, размеры которых находятся в диапазоне немногих десятков метров. Кажется, есть основания надеяться, что процесс завершения формирования второй группы астероидов фактически начался.

Для иллюстрации работы средств поиска опасных объектов приведём статистику текущей полученной информации с разносом по системам, принимавшим участие в накоплении данных. На рис. 2 (см. с. 206) приведена гистограмма АСЗ с размерами 1 км и выше, обнаруженных различными наблюдательными системами. Оно хорошо показывает, как ведущий проект LINEOS начал выполнять программу обнаружения и учёта всех астероидов километрового масштаба и, практически выполнив её на 90 % за десятилетие, передал её последовательно появляющимся проектам второго поколения — CATALINA и PANSTARR.

На следующем, рис. 3 (см. с. 206), показано распределение данных по различным системам обнаружения АСЗ с размерами, начиная с масштаба 140 м и выше. Хорошо видно, как две основные системы обнаружения АСЗ второго поколения последовательно приняли на себя нагрузку мониторинга

и начали учитывать объекты с меньшими размерами, обнаружение которых системе LINEOS было не под силу.

Этот же рисунок хорошо показывает относительные доли вклада различных наземных оптических и космических (КА NEOWISE) систем в общий объём полученной информации.

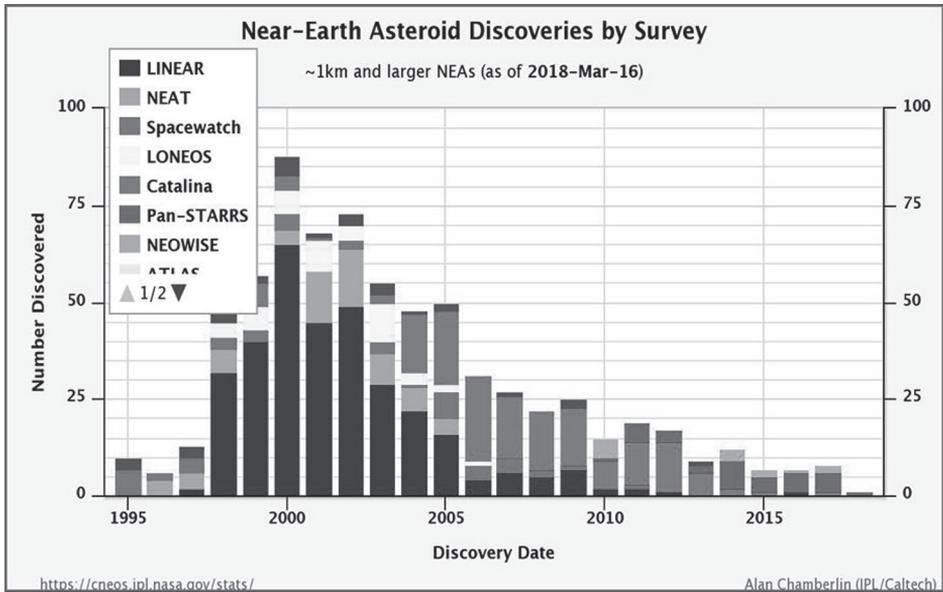


Рис. 2. Накопление данных об АСЗ размером 1 км и выше различными системами

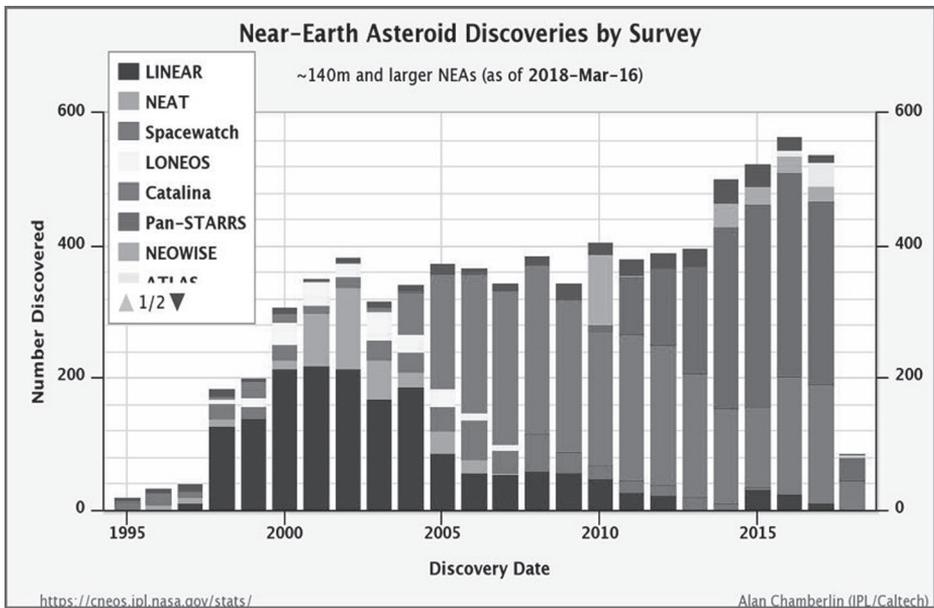


Рис. 3. Накопление данных об АСЗ масштаба 140 м и выше

Далее, на рис. 4 показан рост суммарной информации обо всех АСЗ, обнаруженных за двадцать лет выполнения программы в целом.

Здесь следует отметить результат появления средств обнаружения второго поколения — каталогизация большого количества малых АСЗ, с характерными размерами менее 140 м. Также заслуживают внимания относительные веса данных, полученных раздельно силами США и всех других международных участников, вносящих свои вклады в общее дело понимания реалий астероидной опасности.

На сайте NASA (<https://cneos.jpl.nasa.gov>) начиная с 2009 г. начато ведение статистики по обнаруженным астероидам, сближающимся с Землёй, с более детальной разбивкой на группы различных размеров. Результирующая накопившаяся статистика представлена в динамике на рис. 5 (см. с. 208).

Эти данные показывают, что темп обнаружения новых гектометровых объектов (имеющих диаметры от 100 до 1000 м) действительно проявляет тенденцию к снижению аналогично тому, как это имело место в случае крупных небесных тел километрового масштаба. Более того, наблюдается преимущественный рост объектов, имеющих размеры в районе 100 м. Что же касается сверхмалых тел (0–30 м), то они, как правило, сгорают в атмосфере и квалифицируются как болиды (fireballs), которые регистрируются национальной системой космического наблюдения тех же США. На рис. 6 (см. с. 208) приведены имеющиеся данные о болидах, наблюдавшихся за период с 1988 по 2018 г. Всего за этот срок зарегистрировано появление 719 событий с разной энергией вспышки, причём для 89,7 % из них она невелика и не превосходит уровня, эквивалентного 1 кт ТНТ (тринитротолуола). Из числа оставшихся событий только 9,3 % имеют энергию, лежащую в пределах 1–10 кт ТНТ, 1,4 % — в пределах 10–35 кт ТНТ, и только лишь одно (челябинское событие) имеет энергию вспышки, оценённую как эквивалентную ~440 кт ТНТ.

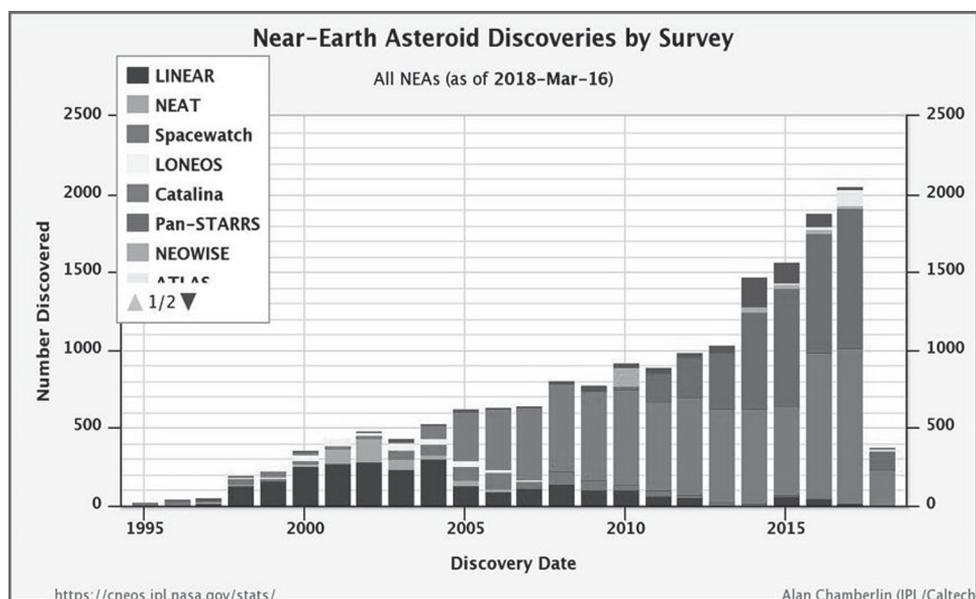


Рис. 4. Накопление данных обо всех обнаруженных АСЗ

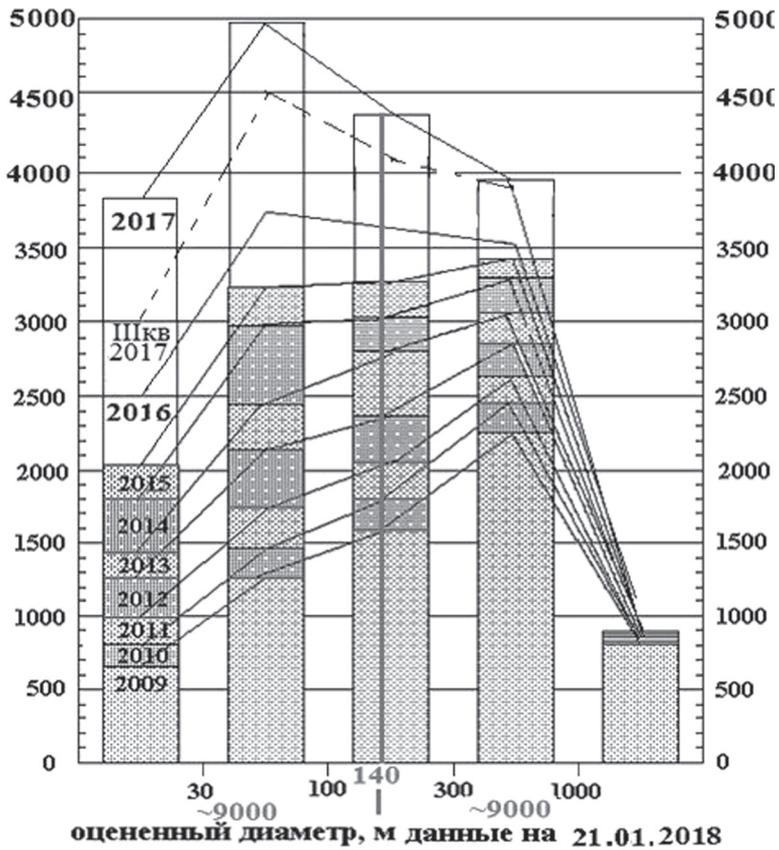
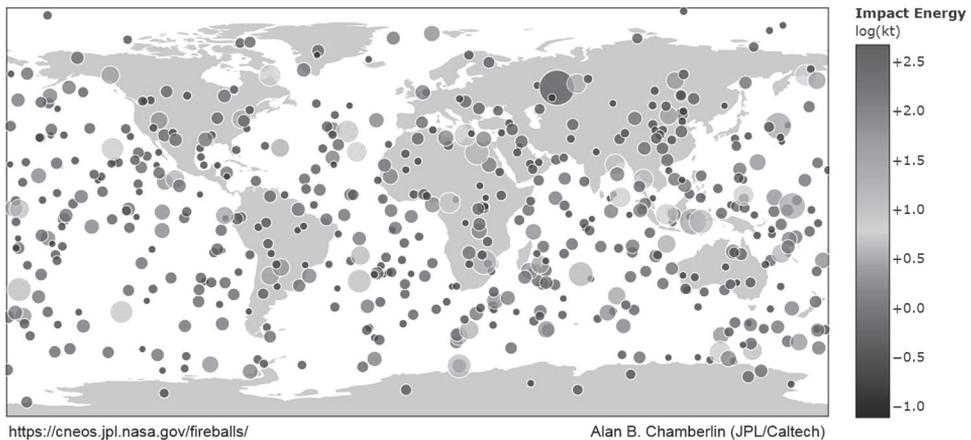


Рис. 5 Распределение АСЗ по размерам

Fireballs Reported by US Government Sensors (1988-Apr-15 to 2018-Feb-01)



<https://cneos.jpl.nasa.gov/fireballs/>

Alan B. Chamberlin (JPL/Caltech)

Рис. 6. Болиды, зарегистрированные в период с 1988 по 2018 г.

Этот рисунок, кстати, наглядно показывает «защитную роль» не только атмосферы Земли, но и той доли её поверхности (~66 %), которую занимают океаны.

Рассмотренные текущие итоги, естественно будут ещё уточняться (и уточняться), однако представляется, что уже можно сделать предварительные выводы по трём группам обнаруженных астероидов. Для первой группы наиболее массивных километровых астероидов, то можно надеяться, что она в первом (если не во втором) приближении определена по составу на 90–95 %, и, по-видимому, в будущем не должна стать источником беспокойства.

Что же касается второй группы гектометровых астероидов с диаметрами от 140 до 1000 м, то она начинает принимать определённые очертания. В дальнейшем, с течением времени, объекты этой группы постепенно будут регистрироваться в диапазоне масс 30–300 м, а её полный численный состав вряд ли превысит ~50 % от общего предполагаемого количества обнаруживаемых астероидов. Также, в рамках текущей оценки можно надеяться, что к 2023–2025 гг. численный состав популяции астероидов второй группы в целом будет завершаться аналогично тому, как группе километровых опасных астероидов. Наконец, количество остальных малых (декаметровых) астероидов, не входящих в указанные две группы, будет расти. Однако очевидно, что они не будут являться глобальными или региональными факторами беспокойства, оставаясь, в крайнем случае, предметом забот ведомства МЧС соответствующих стран.

На множестве всех АСЗ выделяется группа потенциально опасных астероидов (ПОА), орбиты которых таковы, что столкновения с Землёй становятся в принципе возможными. Условным критерием отнесения тел к группе ПОА служит чрезмерное сближение с Землёй на расстояние менее 20LD (здесь LD — среднее расстояние Луны от Земли). Группа ПОА устойчиво составляет ~15 % от численности всех АСЗ второй группы.

По-видимому, к 2025 г. вполне можно рассчитывать на практически каталогизацию и полный мониторинг объектов, входящих в группу ПОА. При этом нужно подчеркнуть, что к настоящему времени за весь период проведения поисков АСЗ (с 1995 по 2018 г.) в этой группе пока обнаружен в 2004 г. лишь один астероид Апофис, для которого уверенно прогнозируется действительно слишком близкий пролёт мимо Земли (на высотах геостационарной орбиты в 2029 г.). Все последствия этого пролёта (и в первую очередь, его повторные чрезмерные опасные сближения с Землёй после 2029 г.) на протяжении текущего столетия сегодня остаются неясными, и не поддаются полному осмыслению и надёжному прогнозу. Поэтому указанный астероид следует по-прежнему считать угрожающим объектом особого внимания. Оценка его реальной опасности должна являться типичной актуальной задачей проблематики АСЗ, подлежащей решению вслед за обнаружением любого АСЗ, чреватого чрезмерным сближением с Землёй. Что касается двух остающихся составных частей проблемы астероидной опасности, следует сказать следующее.

Прежде всего, в преобладающем большинстве предложений о способах противодействия угрозе, до сих пор отсутствует ясное понимание и анализ несоизмеримости масштабов технологически возможных ресурсов, располагаемых человечеством, и необходимого энергетического воздействия на небесное тело (даже на малое), имеющее скорости движения, измеряемые десятками километров в секунду. Указанное необходимое воздействие подразумевает либо изменение траектории тела, либо его гарантированную дезинтеграцию до уровня

компонент обычных метеорных потоков, от которых нас защитит земная атмосфера. Без учёта отмеченного обстоятельства почти все проекты противодействия угрозе будут оставаться на уровне ненаучной фантастики, и будут выражать лишь частное мнение авторов таких предложений.

Что же касается проблемы АСЗ в целом, то логично принять вывод о том, что приоритетные и первоочередные реальные действия в такой части должны определяться наличием уже обнаруженной конкретной угрозы, причём соразмерно её оцениваемой степени. Поэтому, обнаружение фактически единственного астероида Апофис за 20 лет поисков нужно рассматривать как предупреждение о редкой, возможной, но количественно и надёжно до сих пор системно не оценённой реальной угрозе Земле.

Обнаружение Апофиса проявило новую грань астероидной опасности, которая заключается в *неопределённости* прогнозированной орбиты астероида, возникающей *после* чрезмерного сближения с Землёй. Периодичность движения небесных тел приводит к тому, что в солнечной системе конфигурация тесного сближения двух тел, через некоторое время, в идеале, должно повториться точно. Однако реально, за счёт возмущений, движение небесных тел является *почти-периодическим*, и обстоятельства сближений повторяются лишь приближённо. При этом параметр сближения становится случайной величиной, и апостериорная орбита астероида в момент сближения подвергается резким непредсказуемым изменениям. Как следствие, могут возникать новые опасные сближения с Землёй вплоть до удара по ней [Chesley, 2006]. Теории и понимание этого эффекта пока недостаёт глубины и практических выводов.

Заканчивая рассмотрение реалий астероидной опасности на начало 2018 г. можно сделать выводы, что текущие первоочередные задачи астероидной опасности заключаются в следующем:

- 1) продолжение обнаружения и каталогизации АСЗ и ПОА силами существующих и действующих в настоящее время наземных зарубежных и отечественных средств наблюдения;
- 2) дальнейшая разработка предложений по реализации отечественных средств наблюдения за АСЗ космического базирования;
- 3) проведение исследований в части анализа опасных последствий тесного сближения астероида с Землёй на примере астероида Апофис;
- 4) разработка методических основ значительного повышения точности определения орбиты и прогнозирования текущих координат АСЗ;
- 5) разработка предложений по реализации миссии посещения мало-го астероида и оснащения его радиомаяком для уточнения его орбиты с использованием средств VLBI для траекторных измерений.

ЛИТЕРАТУРА

- [Астероидная..., 2010] Астероидная опасность: вчера, сегодня, завтра / под ред. Шустова Б. М., Рыхловой Л. В. М.: Физматлит, 2010. 384 с.
- [Chesley, 2006] Chesley S. R. Potencial impact detection for near-Earth asteroids, the case of 99942 Apophis (2004 MN4) // Proc. IAU Symposium 29th. 2005, Cambridge University press, Cambridge. 2006. P. 215–228.
- [Harris, 2009] Harris A. W. Estimating the NEO population and impact risk: past, present and future // 1st IAA Planetary Defense Conf. Protecting Earth from Asteroids. Granada, Spain, 27–30 Apr. 2009.

ASTEROID THREAT IN 2018: CURRENT STATUS AND PROBLEMS

V. G. Pol, A. V. Simonov

Lavochkin Research and Production Association, NPO Lavochkin

In article is considered the results of the discovery of Near-Earth Asteroids at the spring of 2018. It is confirmed that exaggerated the danger of asteroids for the existence of human civilization, the necessary to continue to discover the dangerous objects and first of all to study consequences arising from the asteroid's close encounter with Earth.

Keywords: small bodies of the solar system, asteroids, asteroid threat

Pol Vadim Georgievich — engineer 2nd category, candidate of technical sciences, polvad@laspace.ru

Simonov Alexander Vladimirovich — lead mathematician, candidate of technical sciences, alex.simonov@laspace.ru

ОЦЕНКА ТОЧНОСТИ ОПРЕДЕЛЕНИЯ СТОЛКНОВИТЕЛЬНОЙ С ЗЕМЛЁЙ ОРБИТЫ ОПАСНОГО АСТЕРОИДА АПОФИС ПО ОПТИЧЕСКИМ ИЗМЕРЕНИЯМ КОМПЛЕКСА «НЕБОСВОД»

П. Гуо¹, В. В. Ивашкин^{1,2}

¹ Институт прикладной математики имени М. В. Келдыша
Российской академии наук (ИПМ им. М. В. Келдыша РАН)

² Московский государственный технический университет
имени Н. Э. Баумана (МГТУ им. Н. Э. Баумана)

Для модельного случая возможного столкновения астероида Апофис с Землёй в 2036 г. выполнен анализ задачи определения параметров орбиты астероида по результатам измерений проектируемого сейчас космического комплекса «Небосвод». Разработаны алгоритмы определения орбиты астероида по оптическим измерениям прямого восхождения и склонения линии визирования астероида с борта КА данного комплекса. В основе этих алгоритмов — метод Гаусса и метод наименьших квадратов. Для случая наблюдения астероида после его возможного сближения с Землёй в 2029 г. и до вероятного столкновения в 2036 г. получены оценки точностей определения параметров орбиты астероида Апофис, в том числе — географических координат (долготы и широты) точки падения астероида на Землю, времени его столкновения с Землёй, а также прицельной дальности, перигейного расстояния.

Ключевые слова: комплекс «Небосвод», точность навигации, астероид Апофис, столкновение с Землёй

ВВЕДЕНИЕ

Для обнаружения опасных для Земли небесных тел (ОНТ), в частности, астероидов и комет, Российской корпорацией «Комета» разрабатывается космическая система «Небосвод» с мощными телескопами апертурой 1,5 м [Кулешов и др., 2013]. В состав этой системы входят два комплекса. Первый комплекс с космическими аппаратами, установленными на геосинхронных орбитах, регулярно осматривает всю небесную сферу, кроме околосолнечной области, которую невозможно наблюдать из-за засветки Солнцем аппаратуры наблюдения. Второй космический комплекс с космическими аппаратами, установленными на гелиоцентрической орбите Земли на расстоянии от неё 40–80 млн км, регулярно осматривает «сбоку» пространство между Солнцем и Землёй, недоступное для наблюдения с Земли [Алыбин и др., 2017; Емельянов, Меркушев, 2005; Кулешов и др., 2013].

Гуо Пэн — аспирант, 869792831@qq.com

Ивашкин Вячеслав Васильевич — главный научный сотрудник, доктор физико-математических наук, профессор Ivashkin@Keldysh.ru

Целью работы является разработка алгоритмов определения параметров орбиты ОНТ по измерениям комплекса «Небосвод», а также оценка точности навигации при их наблюдении. Для этого принят вариант, что комплекс «Небосвод» включает в себя четыре космического аппарата (КА). При этом КА-1 и КА-2 установлены на геосинхронных орбитах, для них принято: период 86 164 с, большая полуось $a = 42\,164$ км; высота перигея $h = 34\,000$ км, эксцентриситет $e \approx 0,042$; наклонение i их орбит соответствует 11° и 12° , географическая долгота восходящего узла орбит $\Omega_G = 35^\circ$ и 135° (КА-1 находится над Подмосковьем, КА-2 — над Дальним востоком), аргумент перигея $\omega = 0^\circ$ [Кулешов и др., 2013]. Момент прохождения перигея принят для обеих орбит КА-1 и КА-2: $JD(t_p) = 2\,462\,196,5$ (1 марта 2029 г.). КА-3 и КА-4 установлены на гелиоцентрической орбите Земли на расстоянии около 40 млн км (0,267 а. е.) от Земли, КА-3 — впереди Земли и КА-4 — позади Земли [Алыбин и др., 2017; Емельянов, Меркушев, 2005; Кулешов и др., 2013].

На каждом космическом аппарате комплекса устанавливаются высокоточные телескопы, позволяющие сканировать небо и визировать звёзды, а также астероиды и кометы. Полагаем, что обработка изображений звёзд и астероида, формируемых на матрице ПЗС, позволяет опознать звёзды, построить на борту КА инерциальную систему координат и привязать к ней линию визирования астероида с борта КА определением её прямого восхождения α и склонения δ , соответствующих моменту измерения t . Вектор измеряемых параметров обозначается через $\psi(t) = (\alpha(t), \delta(t))$. Точности этих измерений определяются характеристиками оптической системы визирования, в частности, матрицы ПЗС. В качестве модели ошибок измерений полагаем, что суммарная ошибка угловых измерений состоит из трёх частей: $\Delta\psi_{\text{и}} = \Delta\psi_{\text{и1}} + \Delta\psi_{\text{и2}} + \Delta\psi_{\text{сист}}$, где $\Delta\psi_{\text{и1}}$ — случайная ошибка измерения углового положения точечного объекта; $\Delta\psi_{\text{и2}}$ — случайная ошибка измерения за счёт неточечного характера объекта; $\Delta\psi_{\text{сист}}$ — систематическая ошибка измерений. Случайные ошибки $\Delta\psi_{\text{и1}}$ и $\Delta\psi_{\text{и2}}$ угловых измерений распределены по нормальному закону с нулевым математическим ожиданием и среднеквадратичными отклонениями (СКО) $\sigma_{\text{и1}}$ и $\sigma_{\text{и2}}$. Так как система ещё проектируется и нет окончательных данных, то по рекомендации авторов проекта системы «Небосвод» [Кулешов и др., 2013] анализ выполнен для вариантов СКО $\sigma_{\text{и1}}$ измерения углового положения точечного объекта: $0,2''$; $0,02''$. Последняя угловая ошибка (10^{-7} рад) соответствует линейной ошибке ~ 15 км на расстоянии 1 а. е. Принято, что $\text{СКО } \sigma_{\text{и2}}(t) = \sigma_R/3\rho$, где $\sigma_R = R_A$ есть СКО линейной ошибки, равное среднему радиусу объекта R_A ; $\rho(t)$ — расстояние от КА до визируемого объекта в данный момент измерений. При этом случайные ошибки измерения прямого восхождения и склонения распределены по нормальному закону с нулевыми математическими ожиданиями и СКО:

$$\sigma_\alpha = \frac{\sigma_{\text{и}}}{\cos\delta}, \quad \sigma_\delta = \sigma_{\text{и}}, \quad \sigma_{\text{и}} = \left(\sigma_{\text{и1}}^2 + \sigma_{\text{и2}}^2 \right)^{1/2}. \quad (1)$$

Систематические ошибки измерений этих углов $\Delta\psi_{\text{сист}}$ приняты равными $\sigma_{\text{и1}}$.

В качестве предмета исследования взята траектория астероида (99942) Апофис в модельном случае его столкновения с Землёй в 2036 г. Разработаны алгоритмы определения орбиты астероида по оптическим измерениям комплекса «Небосвод» на основе метода Гаусса и метода наименьших квадратов. Получены оценки точностей определения параметров орбиты астероида Апофис.

1. АЛГОРИТМЫ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ОРБИТАЛЬНОГО ДВИЖЕНИЯ АСТЕРОИДА

Для данной задачи оценки точности навигации траектория астероида Апофис рассчитывается по начальным данным численным интегрированием дифференциальных уравнений (ДУ) движения астероида. Использовалась система ДУ движения в геоцентрической геоэкваatorialной невращающейся прямоугольной системе координат на эпоху J2000. Учитывались возмущения от Луны, Солнца, планет Солнечной системы (Меркурий, Венера, Марс, Юпитер, Сатурн), а также возмущения от сжатия Земли и давления солнечного света. Радиус-векторы Луны, Солнца и других планет относительно Земли и их гравитационные параметры берутся из JPL-Эфемериды DE421. Для учёта влияния солнечного света в качестве формы астероида Апофис принят однородный шар, средний радиус R_A которого равен 160 м, при этом коэффициент, характеризующий отражение, $C_A = 1,44$, средняя плотность астероида принята $2,5 \text{ г/см}^3$, т. е. масса астероида составляет $\sim 4,3 \cdot 10^{10} \text{ кг}$.

В соответствии с исследованием возможного столкновения астероида Апофис с Землёй в работах [Ивашкин, Стихно, 2009; Ивашкин и др., 2017; Ivashkin, Stikhno, 2007a, b] в качестве номинальных начальных данных (на момент $t_0 = 1$ июня 2028 г. UTC, когда расстояние от Земли до астероида Апофис около 100 млн км) для уравнений движения приняты такие радиус-вектор и вектор скорости Апофиса, что при дальнейшем движении астероид Апофис пролетит в 2029 г. на минимальном расстоянии $\sim 36,8$ тыс. км от центра Земли и далее, в 2036 г. столкнётся с Землёй в момент времени $t_c = 8 \text{ ч } 46 \text{ мин } 3 \text{ с } 13$ апреля 2036 г. UTC ($JD(t_c) = 2\,464\,796,8653125$). При этом географические координаты точки соударения с Землёй составляют $\varphi_c = 23,6040^\circ$, $\lambda_c = -126,3174^\circ$, эта точка находится в Тихом океане, примерно в 1000 км от западного побережья Мексики [Ивашкин и др., 2017]. На данном этапе для приближённого прогнозирования движения КА применён аналитический метод с учётом вековых возмущений от второй зональной гармоники J_2 [Абалакин и др., 1976]. Получая координаты астероида и КА, определяем по ним возмущённые значения измеряемых углов.

На каждый момент измерений проверяются два условия видимости. Это, во-первых, отсутствие вхождения астероида в тень от Земли и Луны, т. е. радиус-вектор КА – астероид не должен пересекать Землю и Луну, проходящие между КА и астероидом. При этом приближённо полагаем Землю и Луну шарами среднего радиуса R_E и R_L ($R_E = 6378,137 \text{ км}$, $R_L = 1737,4 \text{ км}$). Во-вторых, угол β между направлениями с КА на астероид и на центр Солнца не должен быть меньше 30° , что определяется защитной блендой телескопа [Кулешов и др., 2013]. Если хотя одно из этих условий нарушается, то полагаем, что в данный момент измерение не может производиться.

Разработанный алгоритм определения параметров орбиты астероида состоит из двух частей. Сначала на основе нескольких бортовых измерений прямого восхождения α и склонения δ определяется начальное приближение для элементов орбиты астероида с помощью классического метода Гаусса — перехода от угловых координат к радиус-векторам [Ноздрин, Рудаков, 1970; Vallado, 1997]. Затем, после набора некоторого достаточного числа измерений, выполняется статистическая обработка измерений и производится уточнение элементов орбиты. За основу алгоритма уточнения элементов орбиты астеро-

ида взята универсальная методика метода наименьших квадратов (МНК) определения траектории небесного тела по результатам измерений [Аким, Энеев, 1963; Tapley et al., 2004; Лысенко и др., 2014; Eneev et al., 2010]. При этом предполагаем, что измерения производятся с борта КА системы «Небосвод» в моменты времени t_j , $j = 1, 2, \dots, N$. В каждый момент измерений выполняются наблюдения астероида с борта КА и измерения двух углов — его прямого восхождения α и склонения δ . Предполагается наличие в измерениях некоторых случайных и систематических ошибок, описанных выше. При этом уточнение параметров орбиты астероида производится итерационно.

Кратко опишем алгоритм определения параметров движения астероида с помощью МНК. В качестве оцениваемых параметров \hat{x} , задающих орбиту астероида, на данном этапе работы берётся 6-мерный вектор состояния астероида (например, $\mathbf{r}(t_0)$, $\mathbf{v}(t_0)$), отнесённые к фиксированному моменту времени оценивания $t_{\text{оц}}$:

$$\hat{x} = x(t_{\text{оц}}), \quad \hat{x} \in \mathbf{R}^6. \quad (2)$$

Согласно МНК, формула оценки имеет вид:

$$\delta\hat{x} = A^{-1} B \Delta\psi, \quad (3)$$

$$A = \left(\frac{\partial\psi}{\partial\hat{x}} \right)^T \mathbf{P} \left(\frac{\partial\psi}{\partial\hat{x}} \right), \quad B = \left(\frac{\partial\psi}{\partial\hat{x}} \right)^T \mathbf{P}, \quad \Delta\psi = \psi - \tilde{\psi}, \quad (4)$$

где $\delta\hat{x}$ — поправка к оцениваемому вектору; ψ — вектор измерений, размером $2N$; N — общее количество моментов измерений; $\tilde{\psi}$ — вектор расчётных значений измеряемых функций; $\Delta\psi$ — вектор невязки измерений (О–С); $\partial\psi/\partial\hat{x}$ — матрица частных производных от измеряемых параметров по оцениваемым параметрам астероида; индекс T — знак транспонирования; \mathbf{P} — весовая матрица измерений:

$$\mathbf{P} = \text{diag} \left[\sigma_{\alpha}^{-2}(t_1), \sigma_{\delta}^{-2}(t_1), \dots, \sigma_{\alpha}^{-2}(t_N), \sigma_{\delta}^{-2}(t_N) \right]. \quad (5)$$

Оценка вектора \hat{x} получается итерационно. Вычисления завершаются приближением с номером m , для которого выполняется неравенство

$$\left| \delta\hat{x}_i^{(m)} \right| < \varepsilon_i, \quad i = 1, \dots, 6, \quad (6)$$

где $\varepsilon \in \mathbf{R}^6$ — набор заданных точностей.

Конечную оценку параметров движения найдём как сумму ряда

$$\hat{x} = \hat{x}^{(0)} + \sum_{s=1}^m \delta\hat{x}^{(s)}, \quad (7)$$

где $\hat{x}^{(0)}$ — начальное приближение оцениваемых параметров по методу Гаусса.

2. АЛГОРИТМЫ ОЦЕНКИ ТОЧНОСТЕЙ НАВИГАЦИИ И ПРОГНОЗИРОВАНИЯ ДВИЖЕНИЯ АСТЕРОИДА

Оценки точности определения орбиты астероида сделаны с помощью двух методов — аналитического метода и метода статистических испытаний Монте-Карло [Степнов, 1985]. В методе Монте-Карло делается некоторое большое количество ($\sim 10^3 - 10^4$) независимых испытаний измерений с ошибками, затем выполняется обработка этих измерений, определяется орбита астероида, сравнивается эта расчётная орбита с известной исходной точной орбитой и определяется точность навигации.

В аналитическом методе делаем расчёт ковариационной матрицы K_W для ошибок определения вектора W оцениваемых параметров орбиты вследствие случайных ошибок измерений и отклонения $\delta W_{\text{сист}}$ из-за систематических ошибок измерений $\Delta\psi_{\text{сист}}$

$$K_{\hat{x}} = A^{-1}, \quad K_W = \frac{\partial W}{\partial \hat{x}} K_{\hat{x}} \left(\frac{\partial W}{\partial \hat{x}} \right)^T, \quad (8)$$

$$\delta W_{\text{сист}} = \frac{\partial W}{\partial \hat{x}} A^{-1} B \Delta\psi_{\text{сист}}. \quad (9)$$

при этом для A и B полагаем, что весовая матрица P (5) соответствует реальным ошибкам измерений. Отметим, что здесь в нашей задаче матрица частных производных $\partial W / \partial \hat{x}$ определяется численно, методом конечных разностей, и аналитически.

Главная диагональ ковариационной матрицы K_W даёт выражения для СКО $\sigma_{W_i} = \sqrt{(K_W)_{ii}}$, $i = 1, 2, \dots, n$ (n — размерность вектора W). Предполагаем, что максимальное отклонение W_i из-за случайных ошибок измерений составляет $3\sigma_{W_i}$, поэтому максимальное отклонение $\delta W_{i \text{ max}}$ от обеих ошибок измерений для компонента W_i определяем следующим образом:

$$\delta W_{i \text{ max}} = 3\sigma_{W_i} + \left| \delta W_{\text{сист}i} \right|, \quad i = 1, 2, \dots, n, \quad (10)$$

где $\delta W_{\text{сист}i}$ — i -компонента вектора $\delta W_{\text{сист}}$ (9).

Оценки точностей определения параметров орбиты астероида Апофис получены, кроме вектора \hat{x} : по географическим координатам (долготе λ_c и широте φ_c) точки падения на Землю, времени столкновения t_c , а также по перигейному расстоянию r_π и времени прохождения перигея t_π , по компонентам вектора прицельной дальности b в картинной плоскости у Земли (ξ, η). Здесь ось ξ направлена по вектору прицельной дальности b_0 номинальной траектории астероида, ось η — по вектору $C = b_0 \times V_{\infty 0}$, где $V_{\infty 0}$ — номинальная скорость на бесконечности. Набор этих параметров составляет вектор W :

$$W = (\lambda_c, \varphi_c, t_c, r_\pi, t_\pi, b)^T. \quad (11)$$

Также определяем ошибку $\delta b = (\delta \xi^2 + \delta \eta^2)^{1/2}$, которая характеризует разброс вектора прицельной дальности на картинной плоскости.

На данном этапе анализа точность оценок вычисляется при известных характеристиках распределения ошибок измерений (5), при этом не учитывается влияние ошибок знания положения КА, ошибок модели, наличие априорной информации и т.д. [Бахшиян, 2012; Бахшиян и др. 1980; Лидов, 1984, 2010;

Эльясберг, 1976; Tapley et al., 2004]. Влияние этих факторов предполагается исследовать на следующем этапе анализа.

3. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ АНАЛИЗА

Анализ данной задачи включает два этапа. На первом выполнен анализ точности навигации при измерениях только с борта КА-1 и КА-2 комплекса «Небосвод» для пролётной орбиты астероида 99942 Апофис до его тесного сближения с Землёй в 2029 г. При этом получены оценки точностей определения параметров орбиты астероида Апофис — по прицельной дальности в картинной плоскости у Земли, перигейному расстоянию, а также по времени прохождения перигея Земли в 2029 г. Результаты анализа этого этапа представлены в работе [Ивашкин, Гуо, 2016]. Настоящая работа является её развитием. Здесь, на втором этапе, рассмотрена задача оценки точности навигации для модельного случая столкновения астероида Апофис с Землёй в 2036 г. на участке его полёта в течение 7 лет после его тесного сближения с Землёй в 2029 г. и до столкновения в 2036 г. при наблюдениях с помощью всех четырёх КА комплекса «Небосвод».

3.1. Анализ видимости при наблюдении астероида Апофис

Видимость при измерениях с борта КА-1 и КА-2. Анализ показывает, что влияние тени Земли и Луны на видимость Апофиса мало. В течение 7 лет (2029–2036) существуют несколько десятков небольших зон невидимости (ЗНВ) из-за тени Земли при измерениях с бортов КА-1 и КА-2. Длительность этих зон $\Delta t_{\text{ЗНВ}}$ не более 70 мин, например, для КА-1 на интервалах 0–1 ч с 22 июля по 11 августа 2030 г., 21–22 ч с 7 августа по 6 сентября 2031 г., в 22–23 ч с 21 января по 26 февраля 2032 г. и др. Также существуют несколько небольших (до 10 мин) зон невидимости из-за тени Луны, например, для КА-2 имеются ЗНВ в 20 ч 2 мин по 20 ч 12 мин 3 августа 2029 г., в 5 ч 2–12 мин 30 сентября 2029 г., и др.

Засветка от Солнца оказывает сильное влияние на возможность наблюдения астероида Апофис. Так, для интервала с 13 апреля 2029 г. до 13 апреля 2036 г. для КА-1 и КА-2 существуют две близкие длительные зоны невидимости из-за засветки от Солнца. Первая ЗНВ находится с начала 14 апреля до начала 4 мая 2029 г., длительность этого интервала $\Delta t_{\text{ЗНВ1}} \approx 20$ сут. Вторая зона ЗНВ расположена с 19 ч 13 апреля 2032 г. (3 года после тесного сближения астероида Апофис с Землёй в 2029 г.) по 23 ч 13 июля 2034 г., $\Delta t_{\text{ЗНВ2}} = 821$ сут $\approx 2,25$ г. (таблица), она разъединяет интервал времени наблюдения на два относительно больших сеанса наблюдения астероида Апофис. Первый сеанс находится в интервале с 4 мая 2029 г. по 13 апреля 2032 г., $\Delta t_{\text{сеанс1}} = 1076$ сут $\approx 2,95$ г. Второй сеанс продолжается с 14 июля 2034 г. по 13 апреля 2036 г., $\Delta t_{\text{сеанс2}} = 640$ сут $\approx 1,75$ г. На рис. 1 во вращающейся вместе с Землёй гелиоцентрической системе координат более наглядно показана орбита Апофиса относительно Солнца (в точке (0, 0)) и Земли (в точке (1, 0)) после его сближения с Землёй в 2029 г. до столкновения с Землёй в 2036 г. Местоположения КА-3 и КА-4 отмечены треугольниками. Угол Земля –

Солнце – КА-3 (или КА-4) составляет примерно 15° . Там показаны также выходящие из КА линии визирования астероида для предельных случаев $\beta = 30^\circ$, они выделяют зоны невидимости из-за засветки от Солнца. На орбите также звёздами отмечены местоположения астероида Апофис в соответствующие моменты времени.

Зоны невидимости из-за засветки от Солнца при измерениях с борта КА

№ КА	Начало ЗНВ	Конец ЗНВ	$\Delta t_{\text{зв}}$, сут
КА-1	1 ч 14.04.2029	3 ч 04.05.2029	20,08
	19 ч 13.04.2032	22 ч 13.07.2034	821,13
КА-2	2 ч 14.04.2029	13 ч 04.05.2029	20,45
	19 ч 13.04.2032	23 ч 13.07.2034	821,17
КА-3	7 ч 08.06.2029	10 ч 02.07.2029	24,13
	8 ч 11.06.2031	16 ч 27.10.2031	138,33
	15 ч 27.02.2032	21 ч 10.09.2033	561,25
	20 ч 12.02.2034	20 ч 26.05.2034	103
КА-4	15 ч 23.05.2032	8 ч 23.08.2034	821,70
КА-1–4	15 ч 23.05.2032	21 ч 10.09.2033	475,25
	20 ч 12.02.2034	20 ч 26.05.2034	103
КА-1(2) + КА-5(6)	19 ч 13.04.2032	15 ч 23.05.2032	39,8 (0)

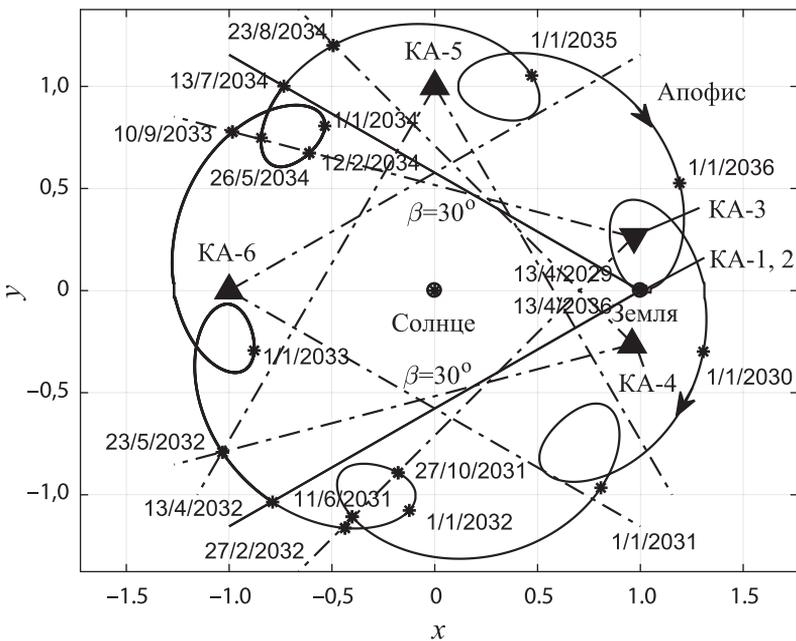


Рис. 1. Орбита Апофиса относительно Земли и Солнца после его сближения с Землей в 2029 г. до столкновения с Землей в 2036 г.

Угол склонения для КА-1 периодически меняется в диапазоне $\delta \in [-23,55^\circ, 39,6^\circ]$. Так как КА-1 и КА-2 обычно довольно близки друг к другу относительно астероида Апофис, то значения измеряемых углов для КА-2 отличаются от углов для КА-1 незначительно.

Видимость при измерениях с борта КА-3 и КА-4. При наблюдении астероида Апофис с бортов КА-3 и КА-4 видимость заметно не улучшается. Для КА-3 имеются четыре длительных зоны невидимости, продолжительностью всего 826,7 сут $\approx 2,265$ г. (см. таблицу и рис. 1). Для КА-4 есть одна зона невидимости из-за засветки от Солнца. Эта зона получается с 15 ч 23.05.2032 по 8 ч 23.08.2034, длительность её составляет 821 сут $\approx 2,25$ г. При наблюдении астероида Апофис с борта всех КА имеются общие длительные зоны невидимости с 23.05.2032 по 10.09.2033 и с 12.02.2034 по 26.05.2034, всего 578,25 сут $\approx 1,58$ г. При этом для существенного улучшения видимости место размещения КА на гелиоцентрической орбите Земли следует взять так, чтобы угол Земля – Солнце – КА составил 90° (КА-5) – 180° (КА-6) (см. таблицу и рис. 1).

3.2. Оценка точностей навигации

С учётом проведённого анализа видимости при наблюдении астероида с борта КА-1 – КА-4, когда существуют длительные зоны невидимости в 2029–2036 гг. (см. таблицу), разделяем траекторию астероида Апофис на четыре участка его наблюдения: 1) дальний — от 4 мая 2029 г. (примерно 20 сут после его тесного сближения с Землёй, начальное расстояние от Земли до астероида $\sim 10,6$ млн км) и до столкновения астероида Апофис с Землёй в апреле 2036 г.; 2) средний — от начала наблюдения 14 июля 2034 г. (это начало второго длительного сеанса наблюдения астероида, здесь астероид находится на 304 млн км (~ 2 а. е.) от Земли) до столкновения астероида Апофис с Землёй в 2036 г.; 3) 1-й ближний (за месяц до столкновения астероида Апофис) — начало наблюдений 14 марта 2036 г. (на расстоянии $\sim 16,6$ млн км до Земли); 4) 2-й ближний (за неделю до столкновения астероида Апофис) — начало наблюдений 7 апреля 2036 г. (на расстоянии 3 млн км до Земли).

Сделано сравнение статистических ошибок для двух способов определения орбиты астероида: аналитически и методом Монте-Карло — на четырёх участках траекторий полёта Апофис. Анализ выполнен для двух вариантов случайной ошибки $\sigma_{ил}$: $0,2''$; $0,02''$, систематические ошибки $\Delta\psi_{сист}$ равны $\sigma_{ил}$. Для метода Монте-Карло проводилось 5000 испытаний. Для сравнения обоих методов оценки точности навигации приведём оценку области рассеивания точек пересечения оцениваемых траекторий с картиной плоскостью, а также точек вероятного падения на поверхности Земли в 2036 г. из-за ошибок измерений.

Рассмотрим сначала случай измерений с бортов КА-1 и КА-2 на 2-м ближнем от Земли участке полёта астероида (начало за неделю до столкновения), интервал времени между измерениями $dt = 6$ ч, получается всего 26 моментов измерений для каждого КА. Получены эллипсы рассеивания на картинной плоскости, а также на поверхности Земли для данной программы измерений при $\sigma_{ил} = 0,2''$ и $0,02''$. На рис. 2 (см. с. 220) приведены эти эллипсы рассеивания на картинной плоскости для двух способов анализа: аналитического (штриховая кривая), и численного метода Монте-Карло (сплошная) при $\sigma_{ил} = 0,02''$.

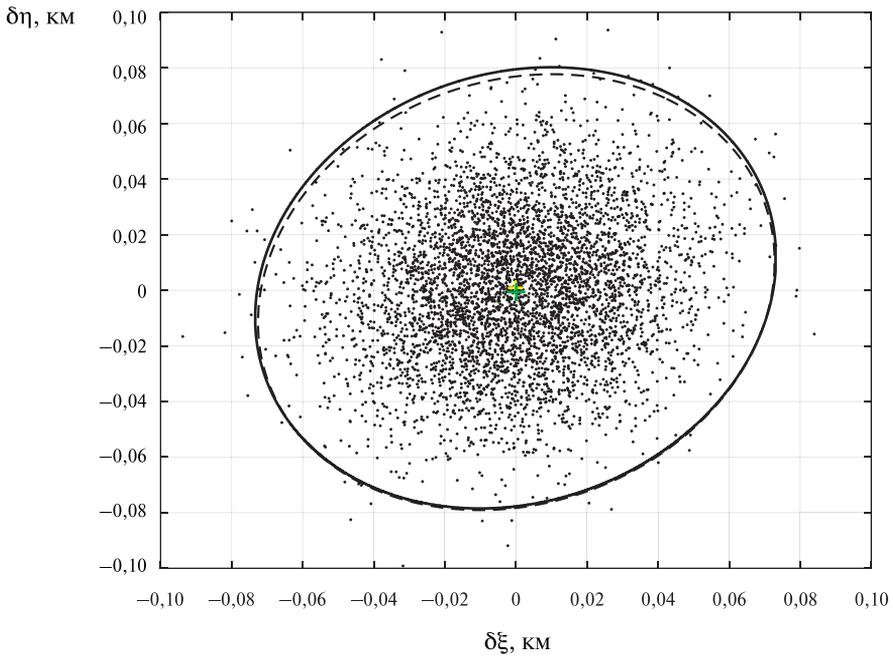


Рис. 2. Эллипсы рассеивания в картинной плоскости для двух способов расчёта: аналитического (штриховая линия) и численного (сплошная), для случая $\sigma_{и1} = 0,02''$

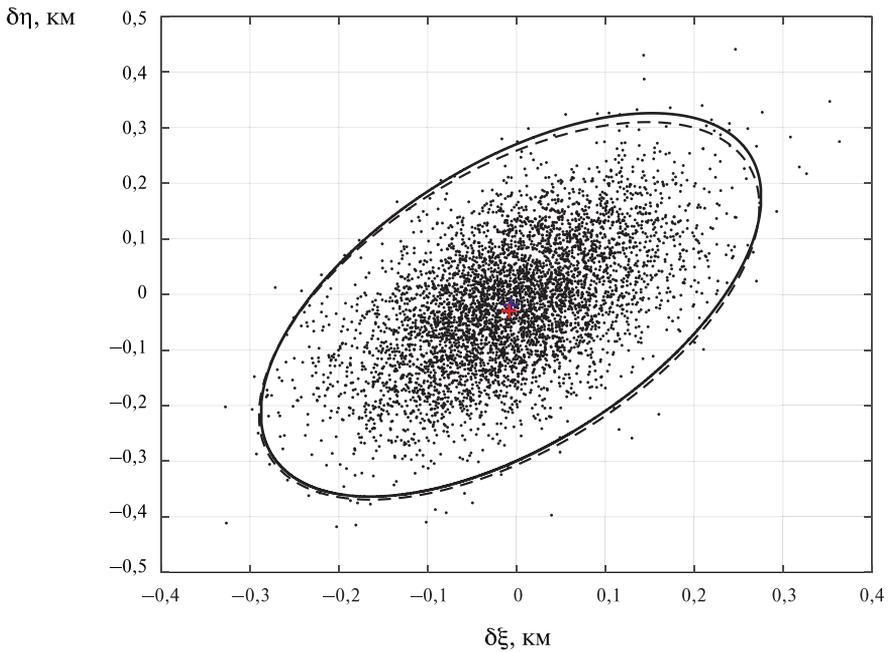


Рис. 3. Эллипсы рассеивания в картинной плоскости для двух способов расчёта: аналитического (штриховая линия) и численного (сплошная), для случая $\sigma_{и1} = 0,2''$

Соответственно, на рис. 3 (см. с. 220) показаны эллипсы рассеивания на картинной плоскости для случая $\sigma_{и1} = 0,2''$. Анализ показывает, что оценки точности определения орбиты астероида по аналитическому методу и методу Монте-Карло близки друг к другу. Точности определения орбиты астероида по компонентам вектора прицельной дальности орбиты ($\delta\xi$, $\delta\eta$), $\delta b = (\delta\xi^2 + \delta\eta^2)^{1/2}$ и по координатам точки падения на Земле ($\delta\lambda_c$, $\delta\varphi_c$) довольно высокие. Для $\sigma_{и1} = 0,02''$ максимальная ошибка $\delta b \approx 80$ м, $\delta\lambda_c \approx \delta\varphi_c \approx 4,7 \cdot 10^{-4}$ град (~ 50 м). Для $\sigma_{и1} = 0,2''$, $\delta b \approx 369$ м, $\delta\lambda_c \approx 2,5 \cdot 10^{-3}$ град, $\delta\varphi_c \approx 1,5 \cdot 10^{-3}$ град (250 м).

Для других участков навигации оба метода также дают близкие результаты. При этом для дальнего участка точности навигации будут, в отличие от ближнего участка, примерно пропорциональны ошибке $\sigma_{и1}$. Здесь отметим, что точность навигации пропорциональна суммарной точности углового измерения $\sigma_{и1}$ (1). На дальнем участке расстояние ρ от КА до астероида достаточно велико, т.е., $\sigma_{и2}$ намного меньше $\sigma_{и1}$, точность навигации оказалась примерно линейной в зависимости от $\sigma_{и1} \approx \sigma_{и2}$, а на ближнем участке эта линейность нарушается из-за большего СКО $\sigma_{и2}$.

Анализ точности навигации на четырёх участках полёта астероида. Для варианта ошибки $\sigma_{и1} = 0,02''$, получена оценка точности навигации. Полагаем, что измерения проводятся одновременно с бортов КА-1, -2 или КА-1, -3 или КА-3, -4.

1. Дальний участок. Полагаем, что измерения проводятся на моменты времени с 0 ч 4 мая 2029 г. до столкновения астероида Апофис с Землёй в 2036 г. равномерно через $\Delta t = 5$ сут, и исключим моменты времени в ЗНВ, в частности, из-за засветки от Солнца. Получена зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала или от времени до столкновения для варианта $\sigma_{и1} = 0,02''$ (рис. 4, см. с. 222). Видно, что в начале мерной базы измерения с бортов КА-3, -4 или с КА-1, -3 лучше, чем измерения с бортов КА-1, -2. Через полгода измерений получается точность навигации $\delta r_{\pi} \approx 2$ км (см. рис. 4) и $\delta b \approx 130$ км. Оценки точностей навигации по географическим координатам точки падения на Землю дают: $\delta\lambda_c \approx 1,13$ град, $\delta\varphi_c \approx 0,33$ град. При этом точности времени прохождения перигея δt_{π} и времени столкновения δt_c примерно равны, $\delta t_{\pi} \approx \delta t_c \approx 7,3$ с.

2. Средний участок. На этом участке полагаем, что наблюдения делаются с 14 июля 2034 г. через $\Delta t = 5$ сут. На рис. 5 (см. с. 222) приведена зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала для варианта $\sigma_{и1} = 0,02''$. Через полгода измерений получаются точности навигации $\delta r_{\pi} \approx 10$ км. Соответственно $\delta b \approx 800$ км, $\delta\lambda_c \approx 4,5$ град, $\delta\varphi_c \approx 2,3$ град, $\delta t_{\pi} \approx \delta t_c \approx 60$ с. После обработки всех измерений на данном участке точность навигации к моменту столкновения астероида Апофис может достигаться $\delta r_{\pi} \approx 1,6$ км (рис. 5), $\delta b \approx 9,6$ км, $\delta\lambda_c \approx 0,08$ град, $\delta\varphi_c \approx 0,033$ град, $\delta t_{\pi} \approx \delta t_c \approx 1,4$ с.

3. Ближние участки (за месяц – неделю до столкновения). Рассмотрим два ближних участка наблюдения. Это, во-первых, первый ближний (за месяц до столкновения) участок. Полагаем, что наблюдения выполняются с 14 марта 2036 г., при этом начальное расстояние от астероида до Земли составляет 16,6 млн км. На втором ближнем участке (за неделю до столкновения) начало измерений — 7 апреля 2036 г., на расстоянии 3 млн км до Земли.

На рис. 6 (см. с. 223) приведена зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала на 1-м ближнем (за месяц до столкновения) участке навигации для варианта $\sigma_{и1} = 0,02''$, $\Delta t = 6$ ч.

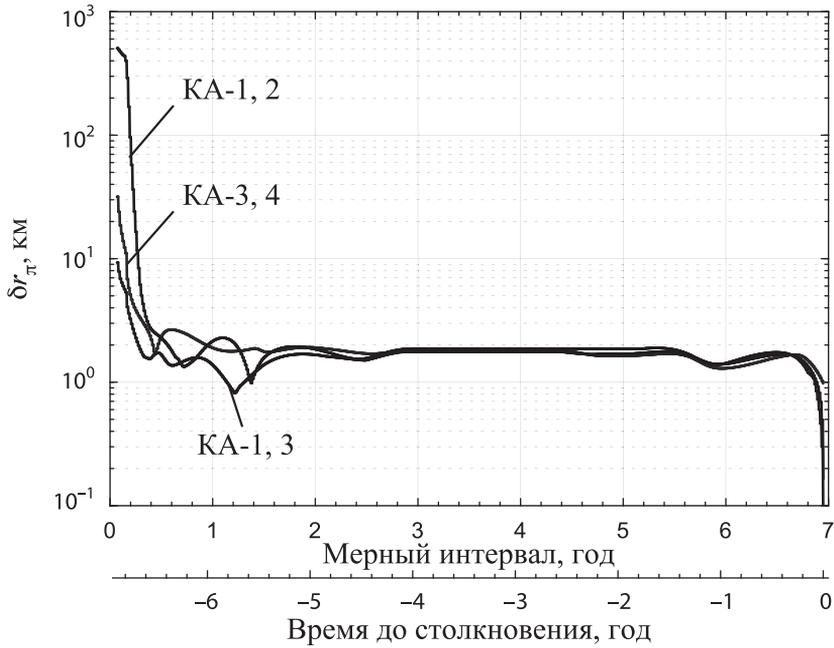


Рис. 4. Зависимости точности навигации по перигейному расстоянию δr_{π} от мерного интервала на дальнем участке навигации ($\sigma_{и1} = 0,02''$, $\Delta t = 5$ сут)

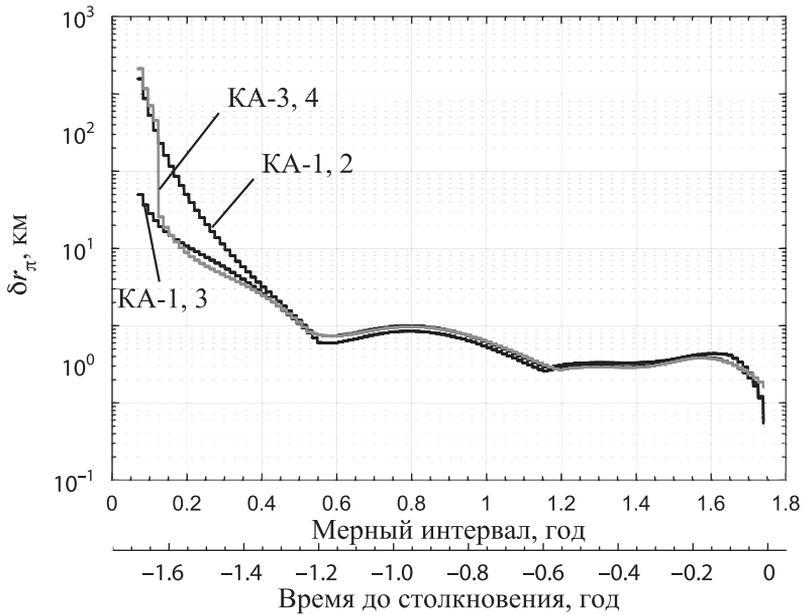


Рис. 5. Зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала на среднем участке навигации ($\sigma_{и1} = 0,02''$, $\Delta t = 5$ сут)

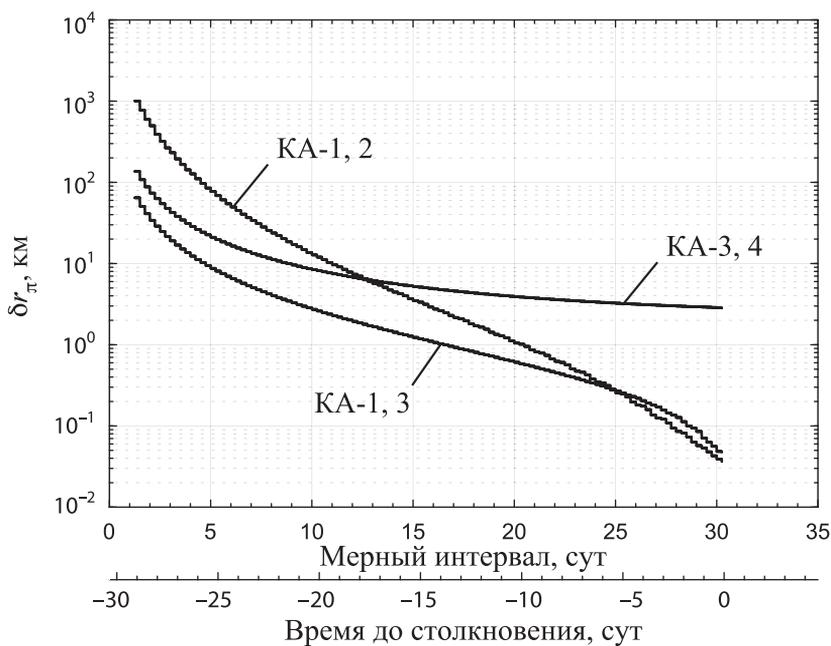


Рис. 6. Зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала на 1-м ближнем участке (начало за месяц до столкновения) ($\sigma_{\text{ил}} = 0,02''$, $\Delta t = 6$ ч)

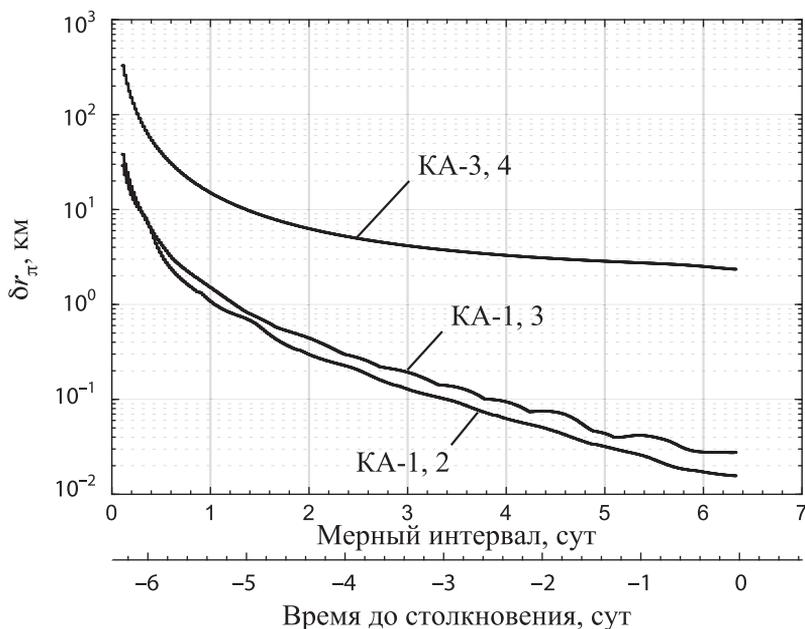


Рис. 7. Зависимость точности навигации по перигейному расстоянию δr_{π} от мерного интервала на 2-м ближнем участке (начало за неделю до столкновения) ($\sigma_{\text{ил}} = 0,02''$, $\Delta t = 30$ мин)

Точность навигации перед столкновением составляет $\delta r_{\pi} \approx 50$ м (соответственно $\delta b \approx 120$ м, $\delta \lambda_c \approx 8 \cdot 10^{-4}$ град, $\delta \varphi_c \approx 6 \cdot 10^{-4}$ град, $\delta t_{\pi} \approx \delta t_c \approx 0,08$ с) для вариантов измерений с бортов КА-1, -2 и КА-1, -3. А для варианта измерений с бортов КА-3, -4, точность навигации намного хуже, $\delta r_{\pi} \approx 3$ км ($\delta b \approx 23$ км, $\delta \lambda_c \approx 0,13$ град, $\delta \varphi_c \approx 0,033$ град, $\delta t_{\pi} \approx \delta t_c \approx 0,5$ с).

Аналогично получена точность навигации для 2-го ближнего (за неделю до столкновения) участка навигации для варианта $\sigma_{и1} = 0,02''$, $\Delta t = 30$ мин. Для вариантов измерений с бортов КА-1, -2 и КА-1, -3 точность навигации по перигейному расстоянию δr_{π} перед столкновением не хуже 30 м (рис. 7, см. с. 223), $\delta b \approx 60$ м, точности по географическим координатам точки падения на Землю составляют $\delta \lambda_c \approx 4 \cdot 10^{-4}$ град, $\delta \varphi_c \approx 3 \cdot 10^{-4}$ град, по времени $\delta t_{\pi} \approx \delta t_c \approx 0,04$ с. А при измерениях с борта КА-3, -4, точности навигации хуже, они составляют $\delta r_{\pi} \approx 2,4$ км, $\delta b \approx 30$ км, $\delta \lambda_c \approx 0,17$ град, $\delta \varphi_c \approx 0,05$ град, $\delta t_{\pi} \approx \delta t_c \approx 0,65$ с.

Таким образом, проведённый сравнительный анализ точности навигации при измерениях с борта КА-1, -2 или КА-1, -3 или КА-3, -4 на дальнем, среднем и ближних участках показывает, что измерения с КА-3 и КА-4 на дальнем участке лучше, чем с КА-1 и КА-2, а для ближних участков измерения с КА-1 и КА-2 лучше, чем с КА-3 и КА-4.

ВЫВОДЫ

Разработаны алгоритмы определения орбиты астероида и получены оценки точности при космических измерениях с помощью системы «Небосвод» как для пролётного, так и для попадающего в 2036 г. в Землю вариантов орбиты астероида. Сделан анализ видимости астероида, необходимой для выполнения измерений. Получены оценки точностей навигации — по прицельной дальности, перигейному расстоянию и времени прохождения перигея, а также по географическим координатам (долготе и широте) точки падения на Землю, времени столкновения для угловых точностей $0,2''$ и $0,02''$. Анализ показывает, что оценки точности определения орбиты астероида по аналитическому методу и методу Монте-Карло близки друг к другу. Сделан сравнительный анализ точности навигации при разных наборах измерений: с бортов КА-1, -2 или КА-1, -3, или КА-3, -4; Измерения с КА-3 и КА-4 на дальнем участке лучше, чем с КА-1 и КА-2. Для ближних участков измерения с КА-1 и КА-2 лучше, чем с КА-3 и КА-4.

ЛИТЕРАТУРА

- [Абалакин и др., 1976] *Абалакин В. К., Аксенов Е. П., Гребенников Е. А.* и др. Справочное руководство по небесной механике и астродинамике / под ред. Г. Н. Дубошина. М.: Наука, 1976. 864 с.
- [Аким, Энеев, 1963] *Аким Э. Л., Энеев Т. М.* Определение параметров движения космического летательного аппарата по данным траекторных измерений // *Космические исследования*. 1963. Т. 1. Вып. 1. С. 5–50.
- [Алыбин и др., 2017] *Алыбин В. Г., Белый А. М., Берёзкин В. В.* и др. Космическая система обзора небесной сферы для наблюдения небесных объектов и обнаружения опасных для Земли небесных тел - астероидов и комет: Патент РФ № 2610066. 2017.

- [Бахшиян, 2012] *Бахшиян Б. Ц.* Оценивание и коррекция параметров движущихся систем: курс лекций. М.: ИКИ РАН, 2012. 72 с.
- [Бахшиян и др., 1980] *Бахшиян Б. Ц., Назиров Р. Р., Эльясберг П. Е.* Определение и коррекция движения. М.: Наука, 1980, 360 с.
- [Емельянов, Меркушев, 2005] *Емельянов В. А., Меркушев Ю. К.* Точность определения орбит малых ОНТ с помощью двух космических телескопов, размещенных на орбите Земли // *Околосветная астрономия — 2005: сб. тр. конф. 19–24 сентября 2005, Казань. 2005. С. 102–108.*
- [Ивашкин, Гуо, 2016] *Ивашкин В. В., Гуо П.* Оценка точности определения параметров орбиты опасного астероида по оптическим измерениям комплекса «Небосвод» // *Изв. вузов. Физика. 2016. Т. 59. № 10/2. С. 51–56.*
- [Ивашкин, Стихно, 2009] *Ивашкин В. В., Стихно К. А.* О предотвращении возможного столкновения астероида Арофис с Землей // *Астрономический вестник. 2009. Т. 43. № 6. С. 502–516.*
- [Ивашкин и др., 2017] *Ивашкин В. В., Стихно К. А., Гуо П.* О структуре множества вероятных траекторий соударения астероида Апофис с Землей в 2036 г. // *ДАН. 2017. Т. 475. № 4. С. 389–394.*
- [Кулешов и др., 2013] *Кулешов Ю. П., Егоров В. Л., Мисник В. П.* и др. Принципы и основные технические решения создания астрономического космического комплекса обнаружения и определения параметров движения опасных для Земли астероидов и комет (комплекс «Небосвод») // *Экологический вестник научных центров ЧЭС. 2013. № 4. Т. 2. С. 89–97.*
- [Лидов, 1984] *Лидов М. Л.* Минимаксная задача оценивания параметров траектории в непрерывной постановке // *Космические исследования. 1984. Т. 22. № 4. С. 483.*
- [Лидов, 2010] *Лидов М. Л.* Минимаксные методы оценивания: препринт ИПМ им. М. В. Келдыша. 2010. № 71. 87 с.
- [Лысенко и др., 2014] *Лысенко Л. Н., Бетанов В. В., Звягин Ф. В.* Теоретические основы баллистико-навигационного обеспечения космических полетов. М.: МГТУ им. Н. Э. Баумана, 2014. 518 с.
- [Ноздрин, Рудаков, 1970] *Ноздрин В. И., Рудаков В. М.* Методы определения орбит: пер. с англ. / под ред. В. Г. Демина. М.: Мир, 1970. 472 с.
- [Степнов, 1985] *Степнов М. Н.* Статистические методы обработки результатов механических испытаний: справочник. М.: Машиностроение, 1985. 232 с.
- [Эльясберг, 1976] *Эльясберг П. Е.* Определение движения по результатам измерений. М.: Наука, 1976. 416 с.
- [Eneev et al., 2010] *Eneev T. M., Ivashkin V. V., Sharov V. A., Bagdasaryan J. V.* Space autonomous navigation system of Soviet project for manned fly by Moon // *Acta Astronautica. 2010. V. 66. P. 341–347.*
- [Ivashkin, Stikhno, 2007a] *Ivashkin V. V., Stikhno C. A.* A problem of the orbit correction for the near-Earth Asteroid Apophis // *58th Intern. Astronautical Congress: Proc. Hyderabad, India, September 24–28, 2007. 2007. Paper IAC-07-C1.7.08.pdf. 11 p.*
- [Ivashkin, Stikhno, 2007b] *Ivashkin V. V., Stikhno C. A.* An analysis of the correction problem for the near-Earth asteroid (99942) Apophis = 2004 MN4 // *The 2007 Planetary Defense Conference. March 5–8, 2007, Washington: Washington Univ., 2007. 7 p.*
- [Tapley et al., 2004] *Tapley B., Schutz B., Born G.* Statistical orbit determination. Elsevier Academic Press, 2004, 563 p.
- [Vallado, 1997] *Vallado D. A.* Fundamentals of Astrodynamics and Applications. N. Y.: Space Technology Series, 1997. 944 p.

AN ACCURACY ESTIMATION OF DETERMINING THE COLLISIONAL ORBIT OF THE DANGEROUS ASTEROID APOPHIS BY THE OPTICAL MEASUREMENTS FROM THE “NEBOSVOD” COMPLEX

P. Guo¹, V. V. Ivashkin^{1,2}

¹ Keldysh Institute of Applied Mathematics Russian Academy of Sciences

² Bauman Moscow State Technical University

For the model case of the asteroid Apophis collision with the Earth in 2036, an analysis has been made for the problem of determining the parameters of the asteroid's orbit according to the measurements by the now projected “Nebosvod” space complex. Algorithms for determining the asteroid's orbit based on the results of optical measurements of right ascension and declination of the line-of-sight of the asteroid from the spacecraft are developed. These algorithms are based on the Gauss' method and least squares method. For the case when the asteroid is observed after its approach to the Earth in 2029 and before the collision in 2036, accuracies estimations of the parameters determination for the asteroid Apophis orbit are obtained. There are estimated, in particular, geographic coordinates (longitude and latitude) of the impact point of the asteroid on the Earth, the collision time, as well as the b-vector of the orbit, perigee distance.

Keywords: “Nebosvod” system, navigation accuracy, asteroid Apophis, collision with Earth

Guo Peng — postgraduate student, 869792831@qq.com

Ivashkin Vyacheslav Vasilievich — chief scientist, doctor of physical and mathematical sciences, professor, Ivashkin@Keldysh.ru

ИССЛЕДОВАНИЕ ТОЧНОСТИ ОПРЕДЕЛЕНИЯ ОРБИТЫ И МЕСТА ВХОДА В АТМОСФЕРУ ОПАСНОГО НЕБЕСНОГО ТЕЛА ПО НАБЛЮДЕНИЯМ КА СОДА

С. А. Нароенков, А. С. Шугаров

Институт астрономии Российской академии наук (ИНАСАН)

В работе проведены исследования точности определения орбиты и эллипса рассеивания возможного места входа в атмосферу опасного небесного тела при наблюдении с помощью разрабатываемого КА СОДА (Система обнаружения дневных астероидов). В качестве примера была исследована орбита Челябинского метеороида. Оценены возможности точного определения орбиты опасного небесного тела при наблюдении с одного и двух космических аппаратов СОДА, расположенных на орбите вокруг точки либрации L_1 системы Солнце – Земля.

Ключевые слова: космический мониторинг, околоземные астероиды, астероидно-кометная опасность

ВВЕДЕНИЕ

Наша планета постоянно сталкивается с космическими телами. Большая часть таких столкновений происходит с небольшими объектами. Метеороиды размером от песчинки до метра эффектно сгорают в атмосфере, и тогда мы наблюдаем явление метеора — «падающие звёзды». Более крупные тела, размером от 20 м и более, сталкиваются с Землёй значительно реже, примерно один раз в 130–240 лет [Емельяненко, Нароенков, 2015].

Челябинское событие 2013 г. показало, какими могут быть последствия от столкновения 20-метрового тела с атмосферой Земли [Емельяненко и др., 2013]. Воздушный взрыв Челябинского болида вызвал массовые повреждения в жилых домах и общественных зданиях. За медицинской помощью обратилось более 1000 человек, а финансовые затраты на восстановление повреждённой инфраструктуры были оценены в 1 млрд руб. На поверхности Земли можно найти множество свидетельств столкновения космических объектов — астероидов и комет в прошлом. Например, Тунгусское событие (1908), Сихотелинский метеоритный дождь (1947), Аризонский кратер (50 тыс. лет назад) дают наглядное представление о последствиях столкновения опасного астероида размером несколько десятков метров с Землёй. Челябинское событие внесло изменения в наши представления о приоритетах в решении задачи противодействия астероидно-кометной опасности:

- Тела размером более 10 м, сталкивающиеся с Землёй, несут серьёзную опасность. Ранее к падениям подобных дециметровых тел относились главным образом с «академическим» интересом.
- Челябинское тело, вошедшее в атмосферу Земли со стороны Солнца, до входа в атмосферу не было (и не могло быть) обнаружено ни одной из наземных систем обнаружения.

Нароенков Сергей Александрович — старший научный сотрудник, кандидат физико-математических наук, snaroenkov@inasan.ru

Шугаров Андрей Сергеевич — младший научный сотрудник, shugarov@inasan.ru

- Появилось понимание, что необходимо создание средств для обнаружения тел декаметрового размера, приближающихся со стороны Солнца.

СИСТЕМА ОБНАРУЖЕНИЯ ДНЕВНЫХ АСТЕРОИДОВ

В рамках работ по разработке средств мониторинга околоземного пространства в Институте астрономии РАН был разработан аванпроект космической системы обнаружения опасных небесных тел «Система обнаружения дневных астероидов (СОДА)» [Шустов и др., 2015].

В основе проекта СОДА лежит несколько основополагающих технических решений.

Первое решение состоит в том, что космический телескоп необходимо разместить в окрестности точки либрации L1 в системе Солнце—Земля (рис. 1). Это позволит наблюдать опасные небесные тела (ОНТ), летящие к Земле от Солнца, «в хвост» при оптимальном фазовом угле, для чего достаточно небольших телескопов с апертурой 30 см.

Второе решение заключается в использовании барьерного способа обнаружения. Основное преимущество барьера — существенное сокращение необходимой площади обзора, что позволяет использовать небольшое количество телескопов с умеренным полем зрения (около 3°) и умеренными по размеру детекторами.

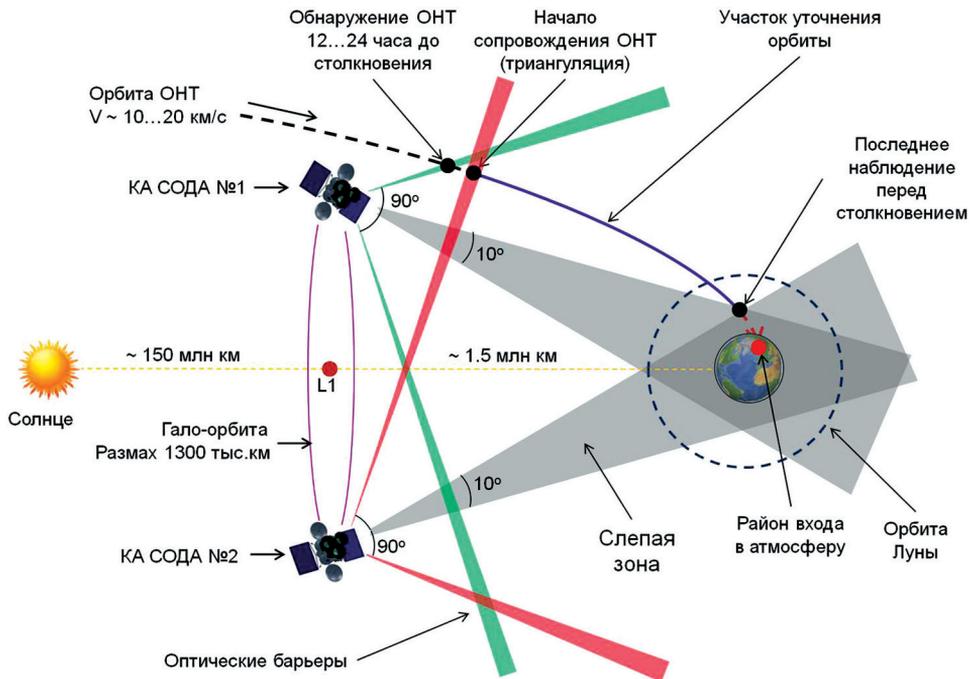


Рис. 1. Схема работы проекта СОДА

Третье решение состоит в том, что после обнаружения ОНТ с помощью барьера, если тело будет классифицировано как угрожающее после первичного грубого определения орбиты, оно будет наблюдаться в индивидуальном порядке по целеуказанию на протяжении всей орбиты, вплоть до возможного столкновения с Землёй.

Четвёртое решение — необходимость оснащения КА специальной маской, экранирующей входные зрачки телескопов от яркого диска Земли. Данное техническое решение позволит избежать появления рассеянного света, и, следовательно, падения чувствительности телескопа, а также позволит наблюдать тела на угловых расстояниях от Земли вплоть до 5° . Это продлит дугу наблюдений и увеличит точность предсказания места входа тела в атмосферу.

Пятое решение заключается в том, что на гало-орбите вокруг точки либрации L1 желательно разместить два аппарата на максимально возможном удалении друг от друга для проведения квазисинхронных наблюдений опасных тел с целью увеличения точности определения орбиты, увеличения полноты обнаружения, резервирование системы.

В качестве космической платформы предполагается использоваться мало-размерную платформу (рис. 2). На КА СОДА возможна установка дополнительной научной аппаратуры для проведения наблюдений Солнца и Земли. Установка четвёртого телескопа, направленного строго на Землю, позволит получать снимки с разрешением до 10 км, пропускная способность радиоканала позволит передавать несколько десятков снимков в разных спектральных полосах в час.

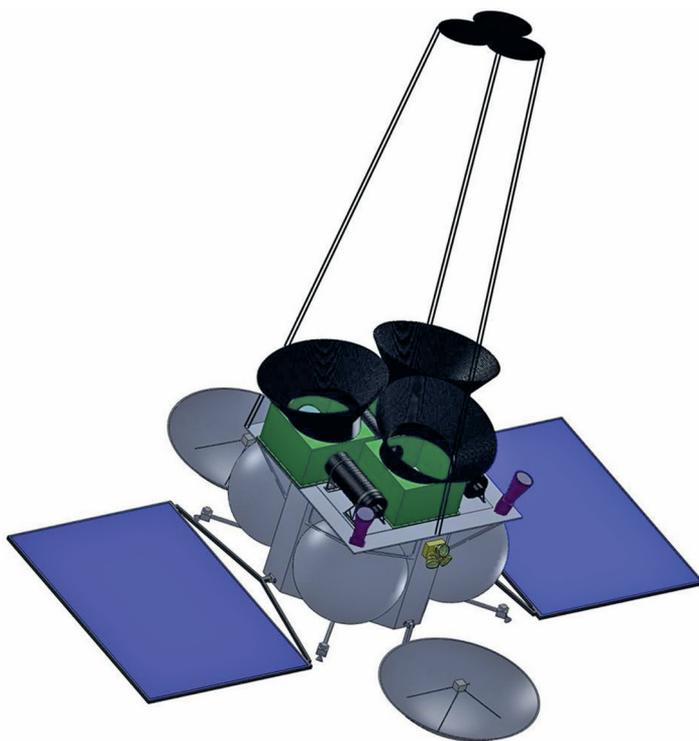


Рис. 2. Компоновка КА с тремя широкоугольными телескопами и маской

Предлагаемая система в настоящее время не имеет действующих аналогов. Расположение специализированного широкоугольного телескопа в окрестности точки Лагранжа L1 системы Солнце–Земля позволит обнаруживать мелкие (от 10 м) опасные тела, приближающиеся к Земле со стороны дневного неба. Данная область принципиально недоступна для действующих наземных обзорных систем.

МОДЕЛИРОВАНИЕ

Результаты моделирования показывают, что проект СОДА позволит обнаружить в течение 5 лет не менее 3000 астероидов размером более 10 м, сближающихся с Землёй со стороны Солнца. Эти данные будут иметь большое значение для фундаментальной науки (исследований динамики малых тел Солнечной системы). Ожидается, что за 5–10 лет работы проекта будет обнаружено по крайней мере одно-два опасных тела на столкновительной орбите, для которых будет заблаговременно (не менее 4 ч) и с достаточной точностью определена точка входа в атмосферу Земли. Событий, требующих внимания (пролёт тел вблизи Земли) будет существенно больше.

При оценке возможностей космической системы СОДА для корректного определения орбиты опасных небесных объектов было проведено моделирование процесса обнаружения космического тела, движущегося по траектории Челябинского болида. Мы провели исследование возможностей системы, состоящей из одного аппарата, и возможностей двух аппаратов, движущихся по гало-орбите в противофазе. На основе орбиты Челябинского болида были созданы фиктивные наблюдения, как будто получаемые с борта космического аппарата. На основе фиктивных наблюдений поэтапно уточнялась орбита опасного небесного тела и проводилась оценка области вероятного входа космического тела в атмосферу Земли.

Процесс моделирования определения орбиты был разбит на несколько этапов. На первом было проведено моделирование движения опасного объекта в момент пересечения участка небесной сферы, попадающей в поле зрения телескопа одного из космических аппаратов. Момент пересечения оптического барьера был выбран в качестве начальной точки отсчёта. Каждый последующий этап моделирования соответствовал одному часу наблюдений объекта. На каждом из них формировался специальный набор наблюдений (эфемерид) опасного объекта с интервалом в 3 мин для каждого КА СОДА. В течение одного часа предполагалось получение 21 наблюдения с каждого космического аппарата. Ошибка единичного наблюдения при моделировании была задана в пределах $\pm 1,4$ угл. с. Предположительно такая ошибка будет получена при определении астрометрических координат объекта на «настоящем» кадре в соответствии с предполагаемым масштабом изображения и размером пикселя ПЗС-камеры. Процесс моделирования заканчивался за 4 ч до столкновения опасного тела с Землёй.

Моделирование движения космического тела до столкновения с Землёй осуществляется путём решения дифференциальных уравнений движения с учётом всех возмущающих сил, действующих на тело до входа в атмосферу Земли. Для решения дифференциальных уравнений используется интегратор Эверхарта 15-го порядка и учитываются возмущения от Солнца, всех планет,

Луны и Плутона [Авдюшев, 2010], координаты планет задаются эфемеридой DE405. Определение орбиты опасного небесного тела по наблюдениям было выполнено дифференциальным методом Гаусса – Ньютона. Моделирование виртуальных орбит на основе ошибок определения орбитальных элементов было осуществлено с помощью ковариационной матрицы Холецкого. В качестве орбиты опасного небесного тела была использована орбита Челябинского болида из работы [Emel’yanenko et al., 2013]. В работе мы рассмотрели два варианта использования космических аппаратов. В первом были использованы наблюдения с двух аппаратов, во втором варианте процесс обнаружения объекта и определения орбиты моделировался только с использованием наблюдений с одного космического аппарата.

Первый этап моделирования процесса обнаружения космического тела показал, что объект попадает в поле зрения телескопа за 26 ч до столкновения с Землёй. Оптический барьер, создаваемый полем зрения телескопа (3°), объект проходит за 50 мин, что даёт нам 15 единичных измерений для определения предварительной орбиты опасного тела. Предварительная орбита объекта, определённая по наблюдения с одного аппарата, имеет значительные ошибки. Неопределённость в долготе восходящего узла достигает $\pm 49^\circ$, в аргументе перигелия — $\pm 121^\circ$. Моделирование движения виртуальных объектов, созданных в соответствии с ошибками определения первоначальной орбиты, показало, что объект пройдёт от Земли в будущем двое суток на расстоянии $1,2 \pm 0,037$ млн км. Ошибки Кеплеровых элементов не позволяют определить опасность такого космического объекта для Земли, поэтому требуются дополнительные наблюдения.

На втором этапе моделирования обнаружения космического тела к первоначальным 15 наблюдениям были добавлены ещё 42 наблюдения, полученные с двух космических аппаратов. Орбита опасного объекта, определённая на основании 57 наблюдений, уже имеет существенно меньшие ошибки, чем на первом этапе моделирования (табл. 1). Моделирование движения виртуальных объектов, созданных в соответствии с ошибками определения орбиты, показало, что объект с вероятностью 100 % столкнётся с Землёй. Место предполагаемого входа в атмосферу опасного небесного тела составляет 1600×90 км (рис. 3, см. с. 232). Такая точность определения орбиты и предполагаемого места входа опасного тела в атмосферу позволяет оценить угрозу столкновения, но область неопределённости точки входа в атмосферу опасного тела слишком велика. Поэтому необходимы дальнейшие наблюдения для увеличения точности определения орбиты.

Таблица 1. Орбитальные параметры опасного небесного тела после двух часов наблюдений

Орбитальные параметры	Значение
Большая полуось a	$1,84774 \pm 0,00044$ а. е.
Эксцентриситет e	$0,600507 \pm 0,000084$
Аргумент перицентра ω	$109,0338 \pm 0,0069^\circ$
Долгота восходящего угла Ω	$326,44733 \pm 0,00023^\circ$
Наклонение i	$5,59640 \pm 0,00152^\circ$
Средняя аномалия M_0	$17,39425 \pm 0,0081^\circ$
T_0	15.02.2013

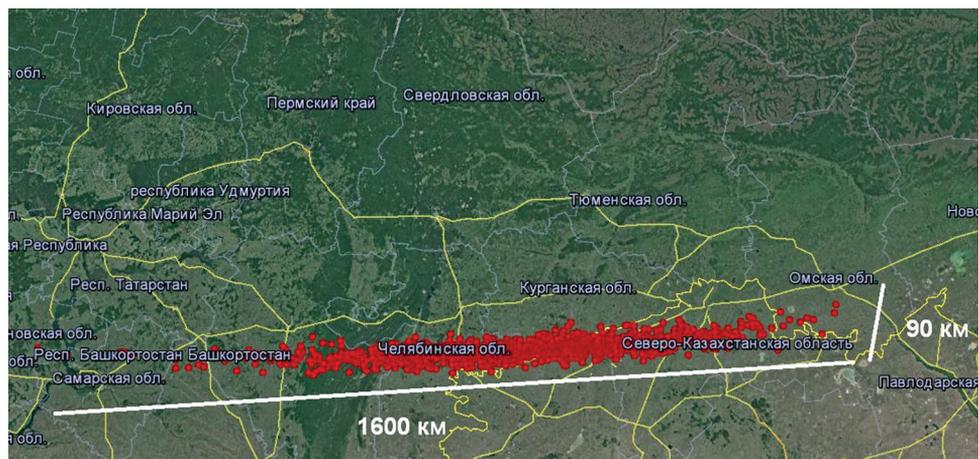


Рис. 3. Возможная область падения космического тела при использовании наблюдений двух космических аппаратов длительностью 2 ч

Последующие наблюдения, полученные с двух космических аппаратов, позволяют более точно определить орбиту и место входа в атмосферу опасного небесного тела. В табл. 2 представлены результаты моделирования, позволяющие оценить размеры вероятной области входа в атмосферу опасного тела.

Таблица 2. Размеры области вероятного входа опасного небесного тела в атмосферу Земли в зависимости от продолжительности наблюдения ОНТ

Время столкновения (ч)	Размеры области вероятного входа опасного небесного тела в атмосферу Земли (км)	Статус прогноза. Информирование государственных органов
25	—	Первое обнаружение одним КА
24	1600×90	Наблюдение опасного объекта с двух аппаратов, подтверждение возможности столкновения объекта с Землёй
23	490×35	Информирование МЧС о столкновении Земли с астероидом. Страна падения: РФ или Казахстан
22	235×25	РФ: Курганская, Челябинская, Тюменская области Казахстан: Костанайская область
21	135×18	Информирование МЧС о районе падения астероида
20	96×15	Информирование МЧС о районе падения астероида
19	65×13	Информирование МЧС о районе падения астероида
18	55×10	Информирование МЧС о точных координатах падения
17	43×10	Последнее наблюдение. Выдача окончательного прогноза о месте входа в атмосферу, траектории на атмосферном участке, энергии взрыва
4	21×4	Последнее наблюдение. Выдача окончательного прогноза о месте входа в атмосферу, траектории на атмосферном участке, энергии взрыва

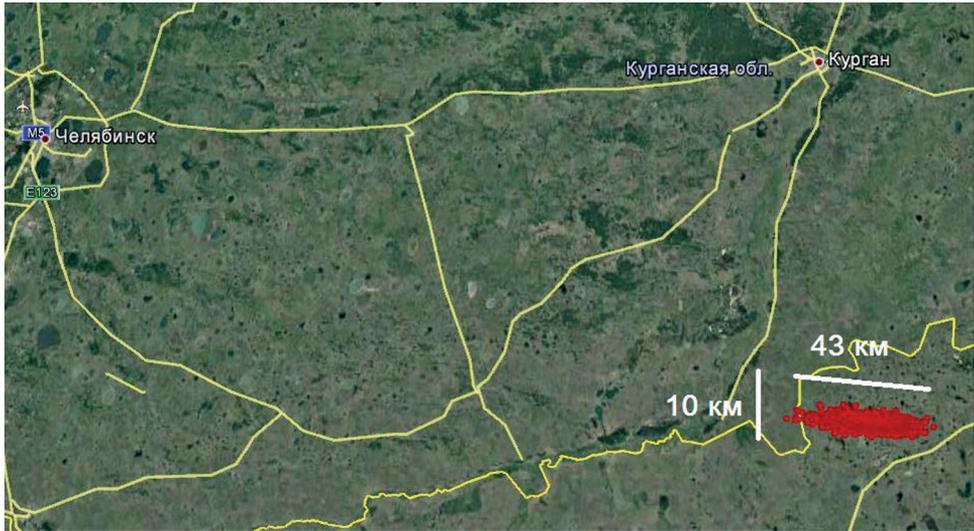


Рис. 4. Возможная область падения космического тела при использовании наблюдений двух космических аппаратов длительностью 9 ч

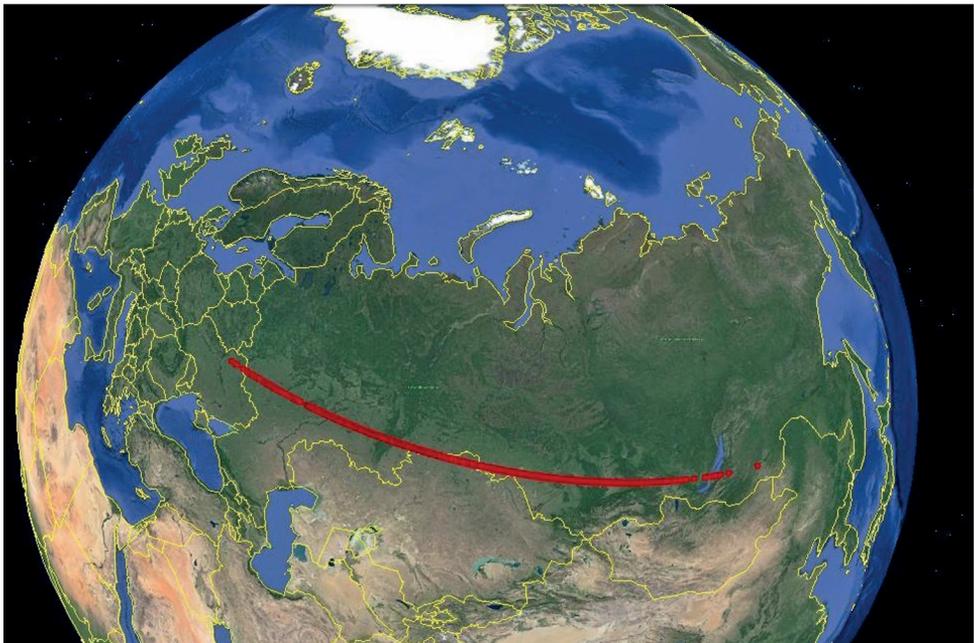


Рис. 5. Возможная область падения космического тела при использовании наблюдений одного космического аппарата длительностью 22 ч

Проведённое моделирование показывает, что наблюдение опасного небесного тела с двух КА в течение 9 ч достаточно, чтобы спрогнозировать область падения в пределах нескольких десятков километров (рис. 4, см. с. 233). Точность определения орбиты и неопределённость точки входа в атмосферу может быть повышена путём уменьшения ошибки единичного наблюдения.

Использование наблюдений только одного космического аппарата не позволяет спрогнозировать область входа опасного тела в атмосферу (рис. 5, см. с. 233). Даже за 4 ч до столкновения с Землёй возможная область падения космического тела представляет собой линию длиной более 6000 км.

ВЫВОДЫ

Челябинское событие изменило наши взгляды на проблему астероидно-кометной опасности. Теперь мы понимаем, что необходимо создавать специальные системы для обнаружения опасных космических объектов размером от 10 м, сближающихся с Землёй со стороны Солнца. В данной работе мы провели моделирование возможностей системы определения орбиты опасных небесных тел, угрожающих Земле. Результаты моделирования точности определения орбит опасных небесных тел на примере Челябинского болида показали, что система, состоящая из двух космических телескопов, способна предсказать падение опасного астероида с точностью в несколько десятков километров за 17 ч до события. Использование только одного космического аппарата позволяет подтвердить факт столкновения, но определение точного места падения затруднительно. Точность определения места падения космического тела может быть увеличена при уменьшении ошибки единичного наблюдения.

ЛИТЕРАТУРА

- [Авдюшев, 2010] *Авдюшев В. А.* Интегратор Гаусса-Эверхарта // Вычислительные технологии. 2010. Т. 15. № 4. С. 31–47.
- [Емельяненко, Нароенков, 2015] *Емельяненко В. В., Нароенков С. А.* Динамические особенности опасных околоземных объектов // Астрофиз. бюл. 2015. Т. 70. № 3. С. 361–368.
- [Емельяненко и др., 2013] *Емельяненко В. В., Попова О. П., Чугай Н. Н., Шеляков М. А., Пахомов Ю. В., Шустов Б. М., Шувалов В., Бирюков Е. Е., Рыбнов Ю. С., Маров М. Я., Рыклова Л. В., Нароенков С. А., Карташова А. П., Харламов В. А., Трубецкая И. А.* Астрономические и физические аспекты Челябинского события 15 февраля 2013 г. // Астрон. вестн. 2013. Т. 47. № 4. С. 262–277. (= *Emel'yanenko V. V., Popova O. P., Chugai N. N., Shelyakov M. A., Pakhomov Yu. V., Shustov B. M., Shuvalov V. V., Biryukov E. E., Rybnov Yu. S., Marov M. Ya., Rykhlava L. V., Naroenkov S. A., Kartashova A. P., Kharlamov V. A., Trubetskaya I. A.* // *Astronomical and Physical Aspects of the Chelyabinsk Event (Feb. 15, 2013)*. Sol. Syst. Res. 2013. V. 47. Iss. 4. P. 240–254.)
- [Шустов и др., 2015] *Шустов Б. М., Шугаров А. С., Нароенков С. А., Прохоров М. Е.* Астрономические аспекты космических угроз: новые задачи и подходы к проблеме астероидно-кометной опасности после Челябинского события 15 февраля 2013 г. // Астроном. журн. 2015. Т. 92. № 10. С. 867–880.
- [Emel'yanenko et al., 2014] *Emel'yanenko V. V., Naroenkov S. A., Jenniskens P., Popova O. P.* The orbit and dynamical evolution of the Chelyabinsk object // *Meteoritics and Planetary Science*. 2014. V. 49. Iss. 11. P. 2169–2174.

INVESTIGATION OF THE ACCURACY OF DETERMINING THE ORBIT AND THE PLACE OF ENTRY INTO THE ATMOSPHERE OF HAZARDOUS ASTEROIDS BY THE SODA SC OBSERVATIONS

S. A. Naroenkov, A. S. Shugarov

Institute of Astronomy of the Russian Academy of Sciences (INASAN)

In this work were conducted the accuracy of the orbit and ellipse scattering of the possible entry point into the atmosphere of a hazardous asteroids was monitored with the aid of the developing SODA space system (Daytime Asteroids Detection System). As an example, the orbit of the Chelyabinsk meteoroid was investigated. The possibility of an accurate determination of the orbit of a hazardous celestial body by observation from one and two SODA spacecraft located in orbit around the libration point L1 of the Sun-Earth system is estimated.

Keywords: space monitoring, near Earth asteroids, asteroid-comet hazard

Naroenkov Sergey Aleksandrovich — senior researcher, PhD, snaroenkov@inasan.ru

Shugarov Andrey Sergeevich — minor scientific researcher, shugarov@inasan.ru

ПРИМЕНЕНИЕ ВЫСОКОЭФФЕКТИВНОГО МЕТОДА ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ ДВИЖЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ ДИНАМИКИ ПОЛЁТА АСТЕРОИДОВ, СБЛИЖАЮЩИХСЯ С ЗЕМЛЁЙ, И НАПРАВЛЯЕМЫХ К НИМ КОСМИЧЕСКИХ АППАРАТОВ

Ю. Ф. Колюка, Т. И. Афанасьева

Центральный научно-исследовательский институт машиностроения
(ЦНИИмаш)

Рассматриваются вопросы, связанные с динамикой полёта астероидов, сближающихся с Землёй (АСЗ) до опасных расстояний, и направляемых к ним КА. Изложены основные принципы оригинального численного метода, позволяющего с высокой точностью и быстродействием осуществлять интегрирование дифференциальных уравнений движения естественных и искусственных космических объектов, который может быть рекомендован для решения сложных прецизионных задач динамики АСЗ. Приводятся результаты применения данного метода для прогнозирования на длительных интервалах времени сближений астероидов Апофис и 2001 JV1 с Землёй, а также для исследования динамики относительного движения КА и астероида 2001 JV1. Оценена точность прогнозирования движения указанных АСЗ на интервалах времени до 200 лет.

Ключевые слова: астероид, сближающийся с Землёй, минимальное расстояние между орбитами, динамика полёта, метод численного интегрирования, прогнозирование сближений с Землёй, точность прогнозирования

ВВЕДЕНИЕ

Данная работа относится к области исследований, связанных с динамикой полёта астероидов, сближающихся с Землёй (АСЗ), и поиском путей противодействия угрозам, создаваемым этими небесными телами.

Для решения прецизионных задач динамики полёта подобных тел, таких, например, как высокоточный прогноз движения АСЗ по орбите и расчёт параметров их сближений с Землёй с максимально возможной точностью, требуется высокоэффективный программно-математический аппарат, позволяющий с высокой точностью и быстродействием осуществлять интегрирование соответствующих дифференциальных уравнений движения, в том числе, и на длительных интервалах времени.

Точный расчёт движения астероида, а также совершающего вблизи него полёт КА, может потребоваться также при решении научно-исследовательских и прикладных задач, связанных с данным небесным телом, с помощью космической техники. В этом случае также будет необходим эффективный аппарат для высокоточного интегрирования уравнений движения астероида и направляемого к нему КА.

Колюка Юрий Фёдорович — начальник лаборатории, кандидат технических наук, старший научный сотрудник, yfk@mcc.rsa.ru

Афанасьева Татьяна Иосифовна — ведущий научный сотрудник, кандидат технических наук, yfk@mcc.rsa.ru

В работе даётся описание сравнительно малоизвестного оригинального численного метода, позволяющего с высокой точностью и быстродействием осуществлять интегрирование дифференциальных уравнений движения естественных и искусственных космических объектов, совершающих полёты по орбитам произвольного типа, при использовании моделей движения (правых частей уравнений) любой сложности.

Особенностью метода является представление решения на каждом шаге в виде интерполяционных полиномов высокого порядка, строящихся по предлагаемой технологии на специальной системе неравно отстоящих узловых точек, близких к узлам многочленов Чебышева. Данный подход получения решения позволяет непосредственно находить параметры движения объекта с заданной точностью в любой точке внутри шага. Метод позволяет надёжно контролировать методическую ошибку на шаге и легко без дополнительных перевычислений менять его длину. Применительно к уравнениям небесной механики метод является самоначинающимся, т.е. не требует привлечения к началу процесса интегрирования каких-либо других более простых методов.

Приводятся результаты применения данного метода при исследовании динамики движения по орбите и сближениях с Землёй на длительных интервалах времени двух астероидов *Апофис* (SPK-ID: 2099942, класс Атонов) и *2001 JV1* (SPK-ID: 3076812, класс Аполлонов), а также динамики относительного движения КА и астероида *2001 JV1*.

ТРЕБОВАНИЯ К СРЕДСТВАМ ВЫСОКОТОЧНОГО ПРОГНОЗИРОВАНИЯ ДВИЖЕНИЯ АСТЕРОИДОВ, СБЛИЖАЮЩИХСЯ С ЗЕМЛЁЙ, И НАПРАВЛЯЕМЫХ К НИМ КОМИЧЕСКИХ АППАРАТОВ

Для исследования динамики полёта астероидов сближающихся с Землёй и решения прецизионных задач этой динамики, связанных, например, с тесными сближениями таких объектов с массивными небесными телами (большими планетами), в результате которых могут происходить существенные изменения параметров их орбит, или (что особенно важно) связанных с высокоточным прогнозированием и тщательным расчётом условий и параметров опасных сближений подобных АСЗ с Землёй, включая попадания в так называемые зоны резонансных возвратов, приводящие к последующим столкновениям АСЗ с Землёй, необходимо использовать максимально точные данные о параметрах движения этих небесных тел, адекватные модели действующих на них сил и соответствующий программно-математический аппарат, позволяющий эффективно осуществлять высокоточный прогноз их орбит, в том числе, и на длительных интервалах времени. Под аппаратом прогнозирования орбит небесных тел обычно понимаются программно-математические средства, реализующие определённые методы решения дифференциальных уравнений, описывающих движение таких тел, по заданным начальным условиям (НУ).

Точный расчёт параметров движения некоторого выбранного астероида, а также подлетающего к нему или уже совершающего вблизи него полёт КА, может потребоваться также при решении научно-исследовательских и прикладных

задач, связанных с данным небесным телом, с помощью космической техники. Например, это может быть задача отклонения данного АСЗ от Земли путём определённого воздействия на него со стороны КА. В этом случае также будет необходим эффективный аппарат для высокоточного интегрирования уравнений движения астероида и направляемого к нему КА.

Тем самым в обоих указанных случаях для осуществления точного расчёта параметров движения естественных и искусственных небесных тел на задаваемых интервалах прогнозирования актуальным является вопрос о наличии методов и программных средств, позволяющих в удобной форме, с высокой точностью и приемлемым быстродействием находить решения дифференциальных уравнений вида:

$$\ddot{\vec{r}} = \vec{F}(\vec{r}, \dot{\vec{r}}, \sum_{\alpha} a_{\alpha}, t). \quad (1)$$

Здесь предполагается, что движение небесных тел (как естественных, так и искусственных) описывается в декартовой инерциальной системе координат XYZ и характеризуется их векторами состояния $\{\vec{r}(t), \vec{V}(t)\}$ в каждый момент времени. (В случае необходимости векторы состояния могут преобразовываться в определённые наборы элементов орбит.) Векторная функция \vec{F} правой части выражения (1) реализует заданную модель движения космического объекта (КО), представляя собой результирующее ускорение от действующих на КО сил, рассматриваемых в данной модели. В общем случае \vec{F} зависит от текущих параметров $\vec{r}, \dot{\vec{r}} = \vec{V}$ на момент t , а также от набора дополнительных параметров $\sum_{\alpha} a_{\alpha}$, характерных для выбранной модели движения. Конкретная траектория движения рассматриваемого КО при заданной модели действующих на него сил, характеризуемая начальными условиями $\{\vec{r}_0 = \vec{r}(t_0), \vec{V}_0 = \vec{V}(t_0)\}$, получается в результате решения уравнения (1) по этим НУ. Начальные условия движения КО могут задаваться и в виде определённого набора элементов $\{\vec{q}\}$ на момент t_0 , взаимно однозначно связанных с начальным вектором состояния $\{\vec{r}_0, \vec{V}_0\}$.

Заметим, что уравнения движения КО могут быть записаны не только в виде формулы (1) в декартовой системе координат, но и в виде с использованием элементов орбиты. Однако в последнем случае они будут гораздо сложнее.

В настоящее время, с учётом достигнутого прогресса в развитии вычислительной техники, наиболее широко используемым способом решения уравнений движения вида (1) является их численное интегрирование с помощью того или иного метода. Данный способ является наиболее универсальным с точки зрения применимости к орбитам любого типа и возможности учёта в модели движения в ясной и удобной форме влияния различных возмущающих сил произвольной природы.

В настоящее время существует множество различных методов, позволяющих осуществлять численное интегрирование уравнений движения вида (1), которые с разными уровнями требований к точности и быстродействию выполняемых расчётов применяются при решении разнообразных задач динамики полёта разных КО.

В то же время, решение ряда задач, связанных с полётом небесных тел, в том числе, астероидов сближающихся с Землёй, а также КА, совершающих полёт вблизи АСЗ, предъявляют особенно высокие требования к точности про-

гнозирования их движения с учётом особых обстоятельств полётов этих космических объектов. В случае АСЗ такими обстоятельствами могут являться, например, тесные сближения этих объектов с большими планетами, включая резонансные возвраты к таким телам (в частности, к Земле при пролёте АСЗ через так называемые «замочные скважины» вблизи Земли), в результате которых орбиты АСЗ резко меняются, а точность прогноза падает.

В целом требования к численному интегрированию уравнений движения (1) при решении современных и перспективных задач динамики полёта астероидов, сближающихся с Землёй, и направляемых к ним КА, включая задачи долгосрочного прогноза движения АСЗ по орбите и их опасных сближений с Землёй, могут быть сформулированы следующим образом:

- 1) обеспечение высокой точности расчёта параметров движения КО (требуемой по условиям решаемой задачи) при прогнозировании на длительных интервалах времени;
- 2) высокое быстродействие численного интегрирования;
- 3) универсальность метода, в плане возможности его одинаковой по эффективности применимости для интегрирования уравнений движения КО на орбитах любого типа — от околокруговых до высокоэллиптических орбит;
- 4) учёт с высокой точностью влияния возмущающих ускорений любой сложности от возмущений любого типа, в том числе, — достижение максимально возможной точности интегрирования при прохождении КО участков тесных сближений с другими небесными телами;
- 5) осуществление эффективного контроля методической погрешности на шаге и возможность достижения заданного уровня этой погрешности;
- 6) возможность прогнозирования движения в прямом и обратном направлениях течения времени;
- 7) простота и удобство расчёта большого количества выходных данных о параметрах движения КО в близко отстоящих точках.

ОСНОВНЫЕ ПОЛОЖЕНИЯ ВЫСОКОЭФФЕКТИВНОГО УНИВЕРСАЛЬНОГО МЕТОДА ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ ДВИЖЕНИЯ НЕБЕСНЫХ ТЕЛ

Одним из методов, удовлетворяющих всем перечисленным выше требованиям, является сравнительно недавно разработанный и ещё недостаточно широко известный оригинальный высокоэффективный универсальный численный метод интегрирования уравнений движения небесных тел, совершающих полёт по орбитам любого типа, как в околоземном пространстве, так и в пределах солнечной системы. Основные положения данного метода изложены ниже.

1. Решение уравнений (1) на шаге длины H представляется в виде интерполяционных многочленов высокого порядка в форме Лагранжа для $\bar{r}(t), \dot{\bar{r}}(t)$, а также для функции правых частей $\bar{F}(t)$, так что

$$\bar{r}(t) = \sum_{j=-m}^m \bar{r}(t_j) L_j(\tau), \quad (2)$$

$$\dot{\bar{r}}(t) = \sum_{j=-m}^m \dot{\bar{r}}(t_j) L_j(\tau), \quad (2')$$

$$\bar{F}(t) = \sum_{j=-m}^m \bar{F}(t_j) L_j(\tau). \quad (2'')$$

Здесь $\{t_j\}$ — узлы интерполяции; $L_j(\tau)$ — так называемые коэффициенты Лагранжа, $\tau = (t - t_0)/H$. Такая форма представления решения позволяет удобно и просто рассчитывать значения $\bar{r}, \dot{\bar{r}}$ и \bar{F} в любой точке в пределах интервала $t_0 - H \leq t \leq t_0 + H$, где t_0 — начало шага.

В качестве узлов интерполяции для полиномов (2-2'') используется множество неравноотстоящих точек, симметрично расположенных относительно начальной точки t_0 . Данное множество точек после приведения к нормированному виду представляет специальную расширяющуюся систему базовых точек $\{\tau_j\}$, симметрично распределённых на интервале $[-1, 1]$ по отношению к центральной точке $\tau_0 = 0$. Количество используемых базовых точек $n + 1 = 2m + 1$ определяет степень интерполирующих полиномов n . Система базовых точек построена таким образом, что при каждом рассматриваемом порядке интерполяции эти точки являются близкими к узлам классических полиномов Чебышева такого же порядка на интервале $[-1, 1]$. Так что характер поведения и точность интерполяции полиномов, построенных на предложенном множестве базовых точек, будет мало отличаться от подобных характеристик, свойственных полиномам Чебышева.

2. Для реализации текущего шага интегрирования используется история решения на предыдущем шаге, представляемая полиномиальным приближением функции \bar{F} на том шаге, и процедура «предсказания-коррекции» (предиктор-корректор схема) для нахождения значений \bar{F}_j в узловых точках нового шага, расположенных за пределами предшествующего шага. При реализации текущего шага процедура коррекции производится с использованием заранее рассчитанных с высокой точностью «матриц интегрирования» $B1 = \|b1_{jk}\|$, $B2 = \|b2_{jk}\|$, элементы которых определяются следующим образом:

$$b1_{jk} = \int_0^{\tau_k} L_j(\tau) d\tau,$$

$$b2_{jk} = \int_0^{\tau_k} \int_0^u L_j(u) du d\tau,$$

где

$$L_j(\tau) = L_j(t) = \tilde{c}_j \prod_{\substack{k=-m, \\ k \neq j}}^m (\tau - \tau_k) \text{ и } \tilde{c}_j = \frac{1}{\prod_{\substack{k=-m, \\ k \neq j}}^m (\tau_j - \tau_k)}.$$

Зная $\bar{F}(t_j)$ ($j = -m, \dots, 0, \dots, m$), значения $\bar{r}(t_k)$, $\dot{\bar{r}}(t_k)$ ($k = -m, \dots, 0, \dots, m$) определяются следующим образом:

$$\dot{\bar{r}}(t_k) = \bar{r}_1 + H \sum_{j=-m}^m \bar{F}(t_j) b1_{jk}, \quad (3)$$

$$\bar{r}(t_k) = \bar{r}_0 + \bar{r}_1 H \tau_k + H^2 \sum_{j=-m}^m \bar{F}(t_j) b_{2jk}, \quad (3')$$

где $\bar{r}_0 = \bar{r}(t_0)$, $\bar{r}_1 = \dot{\bar{r}}(t_0)$.

При $t_j < t_0$ функции $\bar{F}(t_j)$ попадают в «зону действия» предыдущего шага и рассчитываются в указанных точках по интерполяционным формулам (2''), действительным для предыдущего шага.

Поскольку в общем случае $\bar{F}(t_j) = \bar{F}(\bar{r}(t_j), \dot{\bar{r}}(t_j), t_j)$, то при $t_j > t_0$ значения $\bar{F}(t_j)$ вместе со значениями $\bar{r}(t_j)$, $\dot{\bar{r}}(t_j)$ находятся из (6×m) неявных уравнений вида (3), (3'), соответствующих значениям индекса $k = 1, \dots, m$.

Решение данной системы осуществляется итерационно, с применением предиктор-корректор схемы. В этом случае первоначально путём экстраполяции на интервал $[t_0, t_0 + H]$ интерполяционного полинома для \bar{F} , построенного на предыдущем шаге, находятся начальные приближения $\bar{F}^{(0)}(t_j)$ значений функции \bar{F} в точках $t_j > t_0$ ($j = 1, \dots, m$) текущего шага. После этого методом последовательных приближений осуществляется поиск значений \bar{r} , $\dot{\bar{r}}$ в точках $t_k > t_0$ ($k = 1, \dots, m$), а также значений $\bar{F}(t_j)$ в точках $t_j > t_0$ ($j = 1, \dots, m$). Для этого на каждой итерации s ($s = 1, 2, \dots$) последовательно рассчитываются значения

$$\bar{r}^{(s)}(t_k) = \bar{r}_1 + H \sum_{j=-m}^0 \bar{F}(t_j) b_{1jk} + H \sum_{j=1}^m \bar{F}^{(s-1)}(t_j) b_{1jk},$$

$$\bar{r}^{(s)}(t_k) = \bar{r}_0 + \bar{r}_1 H \tau_k + H^2 \sum_{j=-m}^0 \bar{F}(t_j) b_{2jk} + H^2 \sum_{j=1}^m \bar{F}^{(s-1)}(t_j) b_{2jk}$$

при значениях индекса $k = 1, 2, \dots, m$. При этом в данном процессе значения \bar{F} в точках $t_j \leq t_0$, остаются без изменений, т.е. при $j = -m, \dots, 0$ имеет место $\bar{F}^{(s-1)}(t_j) = \bar{F}(t_j)$, где $\bar{F}(t_j)$ рассчитываются по интерполяционным формулам для $\bar{F}(t)$ с предыдущего шага.

Далее, полагая $j = k$ ($j, k > 0$), рассчитываются значения

$$\bar{F}^{(s)}(t_j) = \bar{F}(\bar{r}^{(s)}(t_k), \dot{\bar{r}}^{(s)}(t_k), t_k).$$

Итерационный процесс для каждой точки $t_k > t_0$ ($k = 1, \dots, m$) продолжается до выполнения условий:

$$\left. \begin{aligned} \left| \bar{r}^{(s+1)}(t_{i_k}) - \bar{r}^{(s)}(t_{i_k}) \right| < \varepsilon, \\ \left| \dot{\bar{r}}^{(s+1)}(t_{i_k}) - \dot{\bar{r}}^{(s)}(t_{i_k}) \right| H < \varepsilon, \end{aligned} \right\} \quad (4)$$

где ε — малая величина, определяющая заданный уровень точности.

Количество итераций, требуемых для обеспечения выполнения условий сходимости (4), зависит от величины шага H и удалённости точки t_k от начала шага. При оптимальной длине H обычно требуется не более 3-4 итераций для сходимости в наиболее удаленной интерполяционной точке t_m реализуемого шага.

3. Для обеспечения требуемой точности интегрирования в методе используется специальная процедура контроля методической погрешности на шаге

и уменьшения этой погрешности, в случае необходимости, за счет соответствующего изменения длины шага.

Если принять, что функции $\bar{F} = \bar{F}\left[\bar{r}, \dot{\bar{r}}, \sum_{\alpha} a_{\alpha}, t\right]$ являются аналитическими относительно своих параметров в некоторой окрестности начала текущего шага интегрирования t_0 и величина этого шага H меньше радиуса этой окрестности, то можно показать, что методические ошибки данного метода порядка n на шаге для компонент вектора $\bar{r}(\delta_n x)$, составляющих вектора $\dot{\bar{r}}(\delta_n \dot{x})$ и составляющих вектор-функции $\bar{F}(\delta_n f)$ будут иметь следующие оценки:

$$\delta_n x \leq H^{n+1} \max |\tilde{\omega}_{n+1}(\tau)|,$$

$$\delta_n \dot{x} \leq (n+2)MH^n \max |\tilde{\omega}_{n+1}(\tau)|,$$

$$\delta_n f \leq \max_{t \in [t_0-H, t_0+H]} \left| \bar{F}^{(n+1)}(t) \right| \frac{H^{n+1}}{(n+1)!} |\omega_{n+1}(\tau)|.$$

Здесь

$$M = \max_{t \in [t_0-H, t_0+H]} \frac{|\bar{r}^{(n+1)}(t)|}{(n+1)!},$$

$$\tilde{\omega}_{n+1}(\tau) = \prod_{j=-m}^m (\tau - \tau_j) = \tau(\tau^2 - \tau_1^2) \dots (\tau^2 - \tau_m^2).$$

Для выбранной в методе системы базовых точек $\{\tau_j\}$ всегда будет иметь место

$$\max |\tilde{\omega}(\tau)| \ll 1 \text{ при } \tau \in [-1, 1].$$

Так, например,

$$\max |\tilde{\omega}_{11}(\tau)| < 0,002, \quad \max |\tilde{\omega}_{13}(\tau)| < 0,001, \quad \max |\tilde{\omega}_{15}(\tau)| < 0,0005.$$

Методическая погрешность порядка заданного уровня малости ε (в линейной мере) будет обеспечена на шаге при длине шага $H \leq H_o$, где H_o — оптимальная длина шага, определяемая из условия:

$$\max_{t \in [t_0-H, t_0+H]} \left| \bar{F}^{(n)}(t) \right| \frac{H_o^{n+2}}{n!} = \varepsilon.$$

Если текущий шаг интегрирования реализован с длиной шага H , то приближенное оптимальное значение длины шага \tilde{H}_o на этом шаге интегрирования может быть определено на основании зависимости

$$\tilde{H}_o = H^{n+2} \sqrt{\frac{\varepsilon/H^2}{|\tilde{\Delta}^{(n)} \bar{F}|}},$$

где $\tilde{\Delta}^{(n)} \bar{F}$ — разделённые разности, рассчитываемые согласно формулам

$$\tilde{\Delta}^{(n)} \bar{F} = \sum_{j=-m}^m \tilde{c}_j \bar{F}(t_j), \text{ при значениях } \tilde{c}_j = \left(\prod_{\substack{k=-m, \\ k \neq j}}^m (\tau_j - \tau_k) \right)^{-1}.$$

При реализации метода практически нет необходимости уменьшать шаг, пока имеет место $\tilde{H}_o \geq \frac{1}{2}H$. Также не рекомендуется изменять длину шага в случае, когда оптимальный шаг $\tilde{H}_o > H$, если текущий шаг уже реализован. Однако, если определённое на шаге оптимальное значение \tilde{H}_o меньше половины используемой длины текущего шага, т. е. $\tilde{H}_o < H/2$, то для сохранения методической точности интегрирования следует повторить реализацию текущего шага с использованием его оптимальной длины \tilde{H}_o .

При переходе к выполнению следующего шага его длина H_{next} определяется следующим образом:

$$H_{next} = \beta \alpha H,$$

где H — используемая длина текущего шага; $\alpha = \tilde{H}_o/H$, \tilde{H}_o — оптимальная длина шага, определённая на текущем шаге; β — некоторый эмпирический коэффициент, который можно, например, положить значением $\beta = 1,2$.

4. Метод предусматривает выполнение начала процесса интегрирования уравнений (1) без привлечения каких-либо других методов. В этом случае на начальном (нулевом) шаге интегрирования используется следующее приближение к решению задачи Коши для системы уравнений (1) при заданных начальных условиях $\bar{r}_0 = \bar{r}(t_0)$, $\bar{r}_1 = \dot{\bar{r}}(t_0)$:

$$\bar{r}(t) = \bar{r}_0 + \bar{r}_1(t - t_0) + \bar{r}_2(t - t_0)^2 + O_1(|\bar{r}_3|(t - t_0)^3), \quad (5)$$

$$\dot{\bar{r}}(t) = \bar{r}_1 + 2\bar{r}_2(t - t_0) + O_2(3|\bar{r}_3|(t - t_0)^2). \quad (5')$$

Значения \bar{r}_2 в разложениях (5), (5') определяются согласно формуле

$$\bar{r}_2 = \frac{1}{2} \bar{F}(\bar{r}(t_0), \dot{\bar{r}}(t_0), t_0),$$

и могут быть легко вычислены в результате однократного обращения к расчёту правых частей уравнений (1).

Подобное приближение вполне достаточно для начала итерационного процесса по нахождению значений $\dot{\bar{r}}(t_{0_k}), \bar{r}(t_{0_k})$ в узловых точках на нулевом шаге, при относительно небольшой его длине, согласно описанной выше процедуре. Длину начального шага H_o рекомендуется выбирать как

$$H_o = \gamma \cdot \min \left\{ \frac{|\bar{r}_0|}{|\bar{r}_1|}, \frac{|\bar{r}_1|}{|2\bar{r}_2|} \right\},$$

где γ — некоторая константа, удовлетворяющая условию $\gamma \leq 1/8$.

В отличие от процедуры выполнения текущего шага, процедура выполнения нулевого шага предусматривает применение схемы «предсказание — коррекция» ко всем узловым точкам, расположенным как слева, так и справа от начала шага t_0 .

Таким образом, представленный метод можно отнести к неявным методам численного интегрирования высокого порядка. Применительно к рассматриваемым уравнениям движения небесных тел метод является самоначинающимся. Без каких-либо сложностей метод может применяться для интегрирования уравнений движения в прямом и обратном направлениях течения времени. Характерными особенностями данного метода являются:

- представление решения уравнений движения в удобной форме и с высокой точностью в любой точке окрестности начала текущего шага $[t_0 - H, t_0 + H]$ при достаточно большой длине шага H . Простота получения большого количества выходных параметров движения в близкостоящих по времени точках без привлечения дополнительных процедур интерполяции;
- достижение высокого порядка аппроксимации правых частей уравнений при сокращении количества сложных вычислений за счёт использования «истории» решения с предыдущего шага;
- наличие эффективной и несложной процедуры контроля методической ошибкой на шаге и изменения длины шага для обеспечения требуемого уровня точности интегрирования;
- возможность изменения порядка метода в процессе интегрирования;
- вычислительная устойчивость метода при интегрировании уравнений движения со сложными правыми частями на длительных интервалах времени.

Данный метод программно реализован, прошёл всестороннюю основательную отработку на орбитах разного типа, и уже в течение ряда лет успешно применяется в Центре управления полётами ЦНИИмаш при решении широкого спектра баллистико-навигационных и эфемеридных задач для различных космических объектов искусственного и естественного происхождения, совершающих полёты в ближнем и дальнем космосе.

Более подробно и детально представленный метод описан в работе [Колюка, 2018].

РЕШЕНИЕ ПРЕЦИЗИОННЫХ ЗАДАЧ ДИНАМИКИ ПОЛЁТА АСЗ С ПОМОЩЬЮ ВЫСОКОЭФФЕКТИВНОГО МЕТОДА ЧИСЛЕННОГО ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ ДВИЖЕНИЯ

Представленный выше метод численного интегрирования был применён для решения прецизионных задач динамики движения астероидов, сближающихся с Землёй до близких расстояний. В качестве исследуемых АСЗ были рассмотрены: а) относящийся к классу *Атонов* околоземный астероид *Апофис* (№ 99942), открытый 19.06.2004 г., и б) открытый 14.05.2001 АСЗ с обозначением 2001 JV1 из класса *Аполлонов*.

Для описания движения данных небесных тел использовалась декартова гелиоцентрическая эллиптическая инерциальная система координат (ИСК) стандартной эпохи J2000. Уравнения движения астероида в данной системе координат в общем случае могут быть представлены в виде:

$$\ddot{\vec{r}}_{as} = -\mu_c \frac{\vec{r}_{as}}{r_{as}^3} + \sum_i \mu_i \left[\frac{\vec{r}_i - \vec{r}_{as}}{|\vec{r}_i - \vec{r}_{as}|^3} - \frac{\vec{r}_i}{r_i^3} \right] + \sum_j \vec{F}_{add}^j, \quad (6)$$

где $\vec{r}_{as} = \vec{r}_{as}(t)$ — вектор положения астероида в указанной ИСК, $r_{as} = |\vec{r}_{as}|$; μ_c , μ_i — гравитационные постоянные Солнца и i -го возмущающего тела;

$\vec{r}_i = \vec{r}_i(t)$ — вектор положения i -го возмущающего тела в той же ИСК, $r_i = |\vec{r}_i|$;
 \vec{F}_{add}^j — ускорения, создаваемые дополнительными возмущениями.

Второе слагаемое в правой части уравнений (6) в первую очередь представляет ускорения за счет действия на астероид гравитационного притяжения со стороны больших планет и Луны, но в той же форме могут быть представлены и возмущающие ускорения, создаваемые притяжением наиболее крупных малых планет. В число дополнительных возмущающих членов $\{\vec{F}_{add}^j\}$ могут быть включены ускорения от релятивистских эффектов, нецентральности гравитационного поля Земли и других тел, с которыми может тесно сближаться астероид, от влияния эффекта Ярковского, давления солнечного света на поверхность астероида и др.

После выбора конкретной модели движения астероида, т. е. учёта в правых частях выражения (6) определённых конкретных возмущающих факторов, прогнозирование движения астероида в рамках этой модели по заданным начальным условиям $\{\vec{r}_{as}(t_0), \dot{\vec{r}}_{as}(t_0)\}$ осуществляется посредством интегрирования уравнений (6) с помощью выбранного метода.

В нашем случае для оценки возможностей предложенного метода при решении прецизионных задач динамики полёта астероидов, сближающихся с Землёй, рассматривалась задача численного прогнозирования движения АСЗ на заданные моменты и интервалы времени в рамках модели движения, предусматривающей учёт возмущений от всех больших планет солнечной системы. Для описания движения Земли и других планет в этом случае использовались разработанные в Лаборатории реактивного движения (JPL) высокоточные эфемериды DE 405 (URL: <ftp://ssd.jpl.nasa.gov/pub/eph/exsport/DE405>).

Высокоточное прогнозирование движения астероида *Апофис*

Исходные параметры движения астероида Апофис

Для анализа характеристик численного прогнозирования движения астероида *Апофис* с помощью рассматриваемого метода в качестве исходных параметров движения этого тела были взяты уточнённые элементы его орбиты, опубликованные в статье [Brozović et al., 2018]. Эти элементы были получены в результате определения орбиты астероида по данным оптических и радиолокационных наблюдений данного объекта, выполненным в период с 15 марта 2004 г. по 3 января 2015 г., т. е. на мерном интервале $\sim 10,8$ года. Всего в состав измерительной информации, использованной при определении орбиты, вошли:

- а) *оптические наблюдения*: 4435 измерений прямого восхождения α и 4435 измерений склонения δ , выполненных в период 2004–2015 гг.;
- б) *радиолокационные наблюдения*: два измерения времени запаздывания отражённого сигнала τ и пять измерений доплеровского смещения частоты $\Delta f(2\tau + 5\Delta f)$, выполненные в 2005–2006 гг.; $15\tau + 24\Delta f$, выполненные в 2012–2013 гг.

При обработке указанного состава измерений по методу наименьших квадратов была достигнута мера согласования $\sigma_0 = 0,276$. Полученные в результате

согласования измерений уточнённые элементы орбиты *Апофиса* и их среднеквадратические ошибки приведены в табл. 1. Элементы привязаны к гелиоцентрической эклиптической инерциальной системе координат эпохи J2000 и отнесены к начальному моменту времени $t_0 = 2008.09.24, 0^h$ TDB.

Таблица 1. Уточнённые элементы орбиты астероида *Апофис* в гелиоцентрической эклиптической ИСК эпохи J2000 на момент $t_0 = 2008.09.24, 0^h$ TDB

Элемент	Уточнённое значение	Среднеквадратическая ошибка (1σ)
Большая полуось a (а. е.)	0,9224383009046	$0,1256 \cdot 10^{-9}$
Эксцентриситет e	0,19119529425	$0,3360 \cdot 10^{-8}$
Наклонение i (град)	3,33136922846	$0,33093 \cdot 10^{-6}$
Аргумент перигелия ω (град)	126,4018808361	$0,206361 \cdot 10^{-4}$
Долгота восходящего узла Ω (град)	204,4460284242	$0,210642 \cdot 10^{-4}$
Средняя аномалия M_0 (град)	180,429385930	$0,2038 \cdot 10^{-5}$
Период T (сут)	323,59694852066	$0,6607^{-7}$

В соответствии с уточнёнными данными об орбите астероида *Апофис* минимальное расстояние между его этой орбитой и орбитой Земли (MOID) составило значение MOID = 0,000315683 а. е.

Для осуществления процесса прогнозирования движения *Апофиса* с помощью рассматриваемого метода, т. е. численного интегрирования этим методом уравнений движения вида (6) в декартовой системе координат XYZ, уточнённые элементы орбиты этого астероида с помощью соответствующего преобразования были переведены в начальные условия его движения — координаты и скорости на момент t_0 — в прямоугольной гелиоцентрической эклиптической ИСК эпохи J2000. Эти начальные условия приведены в табл. 2.

Таблица 2. Начальные условия движения астероида *Апофис* в прямоугольной гелиоцентрической эклиптической ИСК эпохи J2000 на момент времени $t_0 = 2008.09.24, 0^h$ TDB (JD = 2454733.5)

Координаты (км)		Скорость (км/с)	
X	-0,1438773995301221D+09	V_x	-0,1231544540281771D+02
Y	0,7912549429570936D+08	V_y	-0,2237273832593018D+02
Z	-0,7658767194222013D+07	V_z	0,8888791473042577D+00

Характер движения астероида *Апофис* по орбите и особенности его сближения с Землёй

В результате численного прогнозирования движения астероида Апофис по приведённым выше начальным условиям было установлено, что двигаясь по орбите, астероид периодически приближается к Солнцу до минимальных расстояний в перигелии $\sim 0,75$ а. е. и удаляется от него, достигая в афелии расстояний $\sim 1,10$ а. е.

Практически синхронно с изменением расстояния от астероида до Солнца меняется и расстояние между орбитами астероида и Земли, обозначаемое как

OID (orbit intersection distance). Математически OID в каждый момент времени t определяется как кратчайшее расстояние между положением астероида в пространстве в этот момент времени и кривой, представляющей орбиту Земли. Минимальное значение, достигаемое OID при движении астероида по орбите, обозначается как MOID (minimal OID).

На рис. 1 и 2 показана динамика движения астероида *Апофис* по орбите и относительно Земли в периоды с 2022 по 2060 г. и с 2057 по 2108 г. соответственно. Здесь кривой синего цвета показано изменение расстояния от Солнца до астероида, кривой зелёного цвета — изменение расстояния от Земли до астероида (для краткости изложения обозначим это расстояние как ρ_{A-Z}) и кривой красного цвета — изменение расстояния между орбитами астероида и Земли (OID).

На рис. 1 и 2 виден периодический характер изменения исследуемых расстояний.

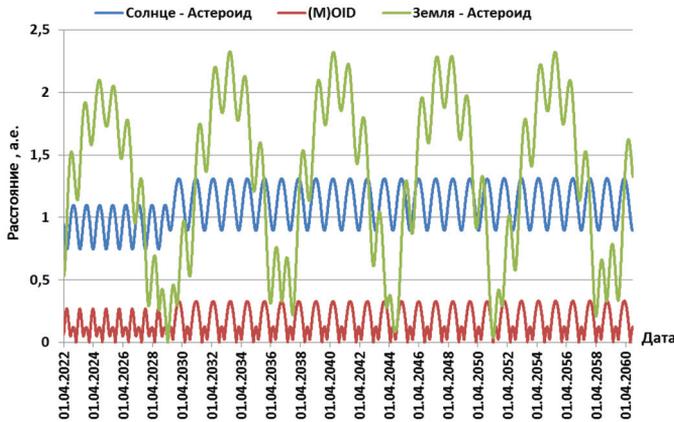


Рис. 1. Динамика движения астероида *Апофис* по орбите и относительно Земли за период 2022–2060 гг.

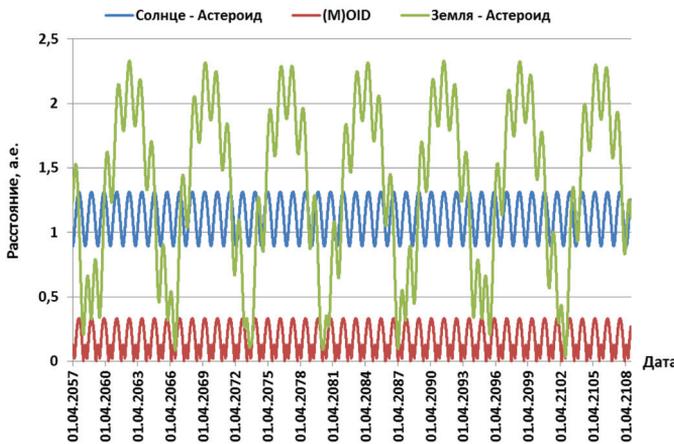


Рис. 2. Динамика движения астероида *Апофис* по орбите и относительно Земли за период 2057–2108 гг.

В районах приближения астероида к Солнцу OID 'ы достигают своих минимальных значений. Кроме того, на рис. 1 отчётливо заметно, что после довольно тесного сближения *Анофиса* с Землёй в 2029 г. график изменения расстояния от Солнца до *Анофиса* поднялся вверх, что свидетельствует, в частности, об изменении большой полуоси орбиты астероида в результате совершения им гравитационного манёвра в сфере действия притяжения Земли. Расстояния в перигелии и афелии орбиты астероида в этом случае также изменились: минимальное расстояние до Солнца увеличилось со значения $q \sim 0,75$ а.е. до $q \sim 0,92$ а.е., максимальное расстояние до Солнца — со значения $Q \sim 1,10$ а.е. до $Q \sim 1,27$ а.е.

Сравнивая рис. 1 и 2 видно, что характер периодического движения астероида по орбите после 2060 г. находится в полном соответствии с характером, который был установлен в период после 2029 г.

Из приведённых на рисунках графиков следует, что минимальные значения ρ_{A-3} достигаются в районах участков орбит, когда OID стремится к своим минимальным значениям, хотя времена достижения минимальных значений $\min \rho_{A-3}$ не полностью совпадают со временами, когда расстояния между орбитами достигают своих минимальных значений $MOID$. При этом далеко не в каждом периоде, когда астероид проходит на минимальном расстоянии от орбиты Земли расстояние ρ_{A-3} также становится малым.

Анализируя рис. 1 и 2 можно установить, что наиболее тесные сближения астероида *Анофиса* с Землёй в рассматриваемый период времени могут произойти в 2029, 2051 и 2102 гг.

Более детально характер изменения расстояний Земля – астероид (ρ_{A-3}) и расстояний между орбитами астероида и Земли (OID) при сближениях в апреле 2029 г., апреле 2051 г. и сентябре 2102 г. показан на рис. 3, 4 и 5 соответственно. Напомним, что потенциально опасными для Земли астероидами, сближающимися с Землёй, считаются те, для которых минимальное расстояние между орбитами астероида и Земли удовлетворяет условию: $MOID \leq 0,05$ а.е.

Геометрия сближения астероида *Анофиса* с Землёй в апреле 2029 г., показанная на рис. 3, свидетельствует в пользу вывода о довольно высокой опасности этого события. В период рассматриваемого сближения в момент времени $t_{MOID} = 2029.04.13, 23^h$ TDB прогнозируется достижение минимального расстояния между орбитами астероида *Анофиса* и Земли со значением $MOID = 0.000121$ а.е. В рассматриваемом событии минимальное расстояние от Земли до *Анофиса* прогнозируется величиной $\min \rho_{A-3} = 0,000257$ а.е. $\approx 38,5$ тыс. км, достигаемой в момент времени $t_{ap} = 2029.04.13, 22^h$ TDB.

Значение минимального расстояния между орбитами *Анофиса* и Земли в апреле 2051 г. прогнозируется величиной $MOID = 0,000186$ а.е., достижение которого должно произойти в момент $t_{MOID} = 2051.04.16, 10^h$ TDB. В рассматриваемом сближении минимальное расстояние от Земли до *Анофиса* прогнозируется на эпоху $t_{ap} = 2051.04.19, 10^h$ TDB со значением $\min \rho_{A-3} = 0,038267$ а.е. $\approx 5,725$ млн км. Несмотря на то, что величины $MOID$ для сближений в апреле 2029 г. и апреле 2051 г. соизмеримы, геометрия сближения астероида *Анофиса* с Землёй в апреле 2051 г., показанная на рис. 4, свидетельствует, что данное сближение не представляет опасности для Земли, так как Земля в момент t_{MOID} будет находиться на значительном удалении от астероида.

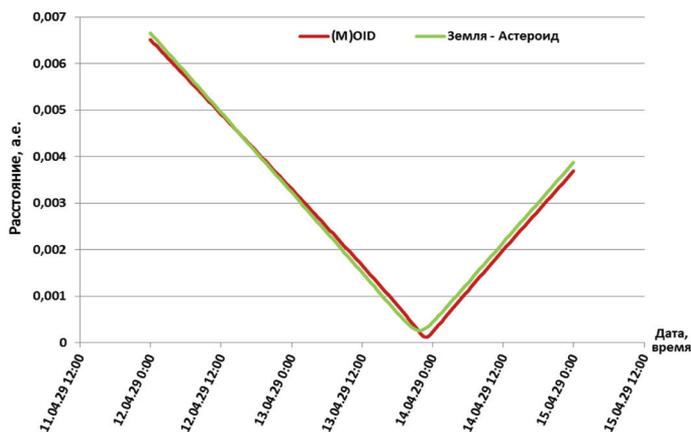


Рис. 3. Геометрия сближения астероида *Апофис* с Землёй в апреле 2029 г.

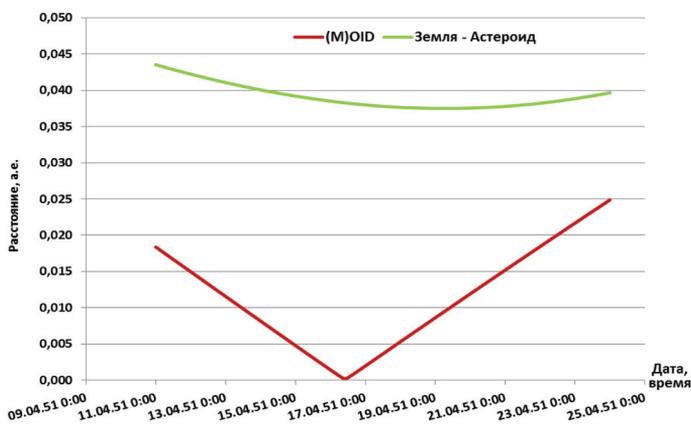


Рис. 4. Геометрия сближения астероида *Апофис* с Землёй в апреле 2051 г.

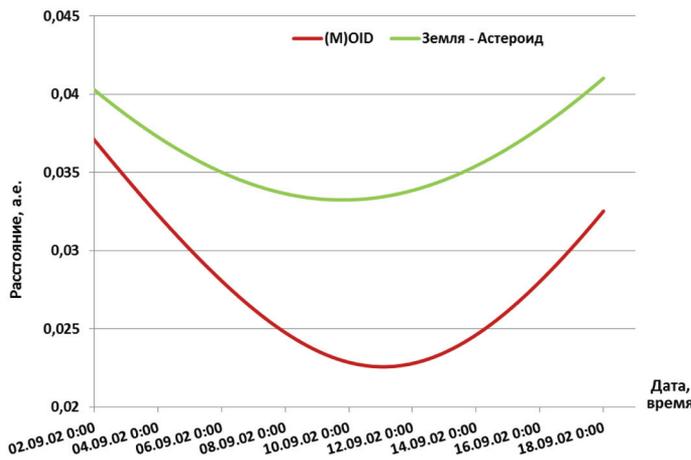


Рис. 5. Геометрия сближения астероида *Апофис* с Землёй в сентябре 2102 г.

При сближении астероида *Апофис* с Землёй в сентябре 2102 г. достижение минимального расстояния между орбитами этих небесных тел прогнозируется на время $t_{\text{MOID}} = 2102.09.11, 2^{\text{h}}$ TDB, когда величина MOID составит 0,022576 а.е. Прогнозируемое значение минимального расстояния между *Апофисом* и Землёй в этом сближении оценивается величиной $\min \rho_{\text{A-3}} = 0,033245$ а.е. $\approx 4,973$ млн км, и это расстояние будет достигнуто в момент $t_{\text{ap}} = 2102.09.10, 19^{\text{h}}$ TDB. Таким образом, сентябрьское сближение 2102 г. также не будет представлять опасности для Земли.

Высокоточное прогнозирование движения астероида 2001 JV1

Исходные данные об орбите астероида 2001 JV1

Определение орбиты астероида было выполнено в Лаборатории реактивного движения (JPL' Solar System Dynamics Group, 05.08.2013. URL: <http://ssd.jpl.nasa.gov/>) по результатам его наблюдений на мерном интервале, охватывающем $\sim 9,5$ лет (с 14 мая 2001 г. по 12 ноября 2010 г.), включающем следующие измерения:

- а) *оптические наблюдения*: 129 измерений α + 129 измерений δ , выполненные в период с мая по сентябрь 2001 г.; 1 измерение α + 1 измерение δ , выполненные 12.09.2010 г.;
- б) *радиолокационные наблюдения*: 1 измерение τ + 2 измерения Δf , выполненные в мае 2001 г.

При определении параметров орбиты астероида была получена мера согласования измерений $\sigma_0 = 0,423$.

Результатом определения орбиты астероида явились значения её параметров в гелиоцентрической эклиптической инерциальной системе координат эпохи J2000, отнесённые к начальному моменту времени $t_0 = 2013.11.04, 0^{\text{h}}$ TDB. Полученные параметры орбиты астероида 2001 JV1 вместе с их среднеквадратическими ошибками приведены в табл. 3.

Таблица 3. Параметры орбиты астероида 2001 JV1 в гелиоцентрической эклиптической ИСК эпохи J2000 на момент времени $t_0 = 2013.11.04, 0^{\text{h}}$ TDB

Элемент	Уточненное значение	Среднеквадратическая ошибка (1 σ)
Большая полуось a (а.е.)	1,7049595281	$1,0071 \cdot 10^{-7}$
Эксцентриситет e	0,4354004982	$1,441 \cdot 10^{-7}$
Наклонение i (град)	3,33136922846	$1,9153 \cdot 10^{-5}$
Аргумент перигелия ω (град)	200,9607668545	$8,5299 \cdot 10^{-5}$
Долгота восходящего узла Ω (град)	92,3242370595	$7,9819 \cdot 10^{-5}$
Средняя аномалия M_0 (град)	194,2926945884	$0,18334 \cdot 10^{-3}$
Время прохождения перицентра t_{π}	2014, 11, 13, 29040	0,000447JD
Период T (годы)	2,226277585	$1,9725 \cdot 10^{-7}$
Расстояние в перигелии q (а.е.)	0,962619300126	$2,14110^{-7}$
Расстояние в афелии Q (а.е.)	2,447299756069	$1,445610^{-7}$

Соответствующие приведённым элементам орбиты координаты и скорости астероида *2001 JV1* на начальный момент времени $t_0 = 2013.11.04$, 0^h TDB в прямоугольной инерциальной гелиоцентрической эклиптической системе координат эпохи J2000 даны в табл. 4.

Таблица 4. Начальные условия движения астероида *2001 JV1* в прямоугольной гелиоцентрической эклиптической ИСК эпохи J2000 на момент времени $t_0 = 2013.11.04$, 0^h TDB (JD = 2456600,5)

Координаты (км)		Скорости (км/с)	
X	-0,1786124548953805D+09	V_x	-0,1182722956857645D+02
Y	0,3170777628770292D+09	V_y	-0,8131638593531846D+01
Z	0,1925202940222267D+08	V_z	0,1412139847937905D+01

Динамика движения астероида *2001 JV1* по орбите и особенности его сближения с Землёй в период 2018–2118 гг.

Орбита астероида *2001 JV1* существенно отличается по форме и размерам от орбиты астероида *Апофис*: у орбиты *2001 JV1* почти в два раза больше большая полуось и более чем в два раза больше эксцентриситет по сравнению с орбитой *Апофиса*. Двигаясь по своей достаточно вытянутой орбите астероид *2001 JV1* удаляется от Солнца до расстояний $\sim 2,45$ а.е. в афелии и приближается к нему на расстояния $\sim 0,96$ а.е. в перигелии. Данный астероид может сближаться с Землёй до расстояний $\sim 0,02$ а.е.

Характер движения астероида *2001 JV1* по орбите вокруг Солнца и относительно Земли на столетнем интервале — в период с 2018 по 2118 г. — показан на рис. 6.

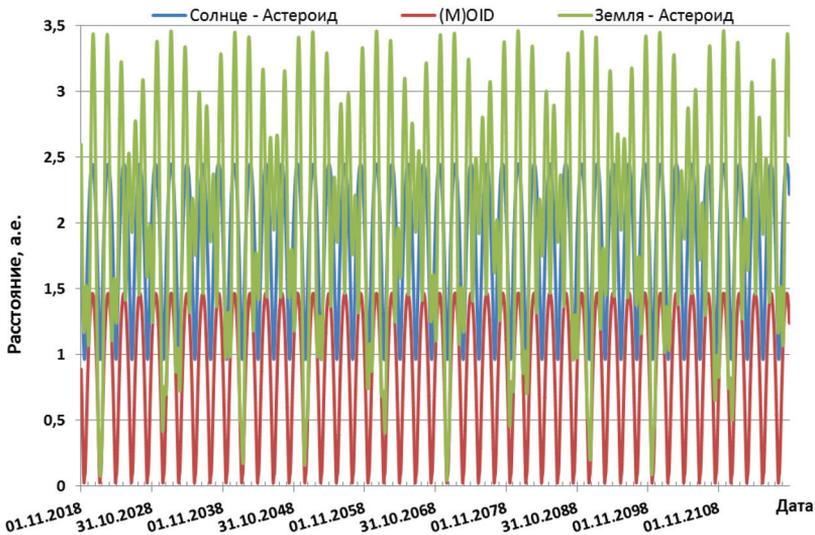


Рис. 6. Динамика движения астероида *2001 JV1* по орбите и относительно Земли в период 2018–2118 гг.

Здесь также как и на рис. 1, кривая синего цвета отображает характер изменения гелиоцентрического расстояния до астероида, кривая зелёного цвета — изменение расстояния от Земли до астероида и кривая красного цвета — изменение расстояния между орбитами астероида и Земли (OID). На рисунке виден периодический характер изменения всех этих величин. В районах приближения астероида к Солнцу OID'ы достигают своих минимальных значений, характеризуемых значениями MOID $\sim 0,02$ а. е. Тем самым в рассматриваемый период времени астероид 2001 JV1 не сможет приблизиться к Земле на расстоянии меньшие, чем $\min \rho_{A-3} \sim 3$ млн км.

В табл. 5 приведены результаты прогноза наиболее тесных сближений астероида 2001 JV1 с Землёй, в которых астероид приблизится к Земле на расстояние меньшее, чем 1 а. е., в период с 2018 по 2118 г.

Таблица 5 Основные характеристики наиболее тесных сближений астероида 2001 JV1 с Землёй в период 2018–2118 гг.

Год	MOID		$\min \rho_{A-3}$	
	Дата	Величина (а. е.)	Дата	Величина (а. е.)
2021	23.06	0,02217	22.07	0,0744
2030	20.05	0,02186	08.05	0,4162
2041	07.07	0,02137	03.09	0,1690
2050	05.06	0,02177	25.05	0,1573
2059	01.05	0,02107	22.04–23.04	0,7380
2061	22.07	0,02192	21.09	0,4019
2070	18.06	0,02098	27.06	0,0412
2079	18.05	0,02115	07.05	0,4528
2081	10.08	0,02089	07.10	0,7006
2108	07.05	0,02096	27.04	0,6494
2110	30.07	0,02036	26.09	0,5000

Данные, приведённые в табл. 5, свидетельствуют, что наиболее опасным событием для Земли является сближение 2070 г., когда $\min \rho_{A-3} = 0,0412$ а. е. $\approx 6,584$ млн км. В 2021 г. астероид пролетит на минимальном расстоянии от Земли примерно через месяц после того, как он пройдёт точку MOID.

РЕШЕНИЕ ЗАДАЧИ ВЫСОКОТОЧНОГО ПРОГНОЗИРОВАНИЯ ОТНОСИТЕЛЬНОГО ДВИЖЕНИЯ СОВЕРШАЮЩЕГО ПОЛЁТ ВБЛИЗИ АСЗ КОСМИЧЕСКОГО АППАРАТА ПУТЁМ СОВМЕСТНОГО ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ ДВИЖЕНИЯ ЭТИХ ТЕЛ

В перспективных космических программах могут возникнуть задачи, когда потребуется одновременно рассчитывать параметры движения АСЗ и направляемых к ним с определёнными целями КА. Одним из способов описания дви-

жения КА относительно астероида в этом случае может являться совместное интегрирование уравнений движения этих тел, записанных в декартовой гелиоцентрической ИСК.

Так, если представить уравнения движения АСЗ в форме (6), а уравнения совершающего полёт вблизи него КА в форме

$$\ddot{\bar{r}}_{КА} = -\mu_c \frac{\bar{r}_{КА}}{r_{КА}^3} + \sum_i \mu_i \left[\frac{\bar{r}_i - \bar{r}_{КА}}{|\bar{r}_i - \bar{r}_{КА}|^3} - \frac{\bar{r}_i}{r_i^3} \right] + \mu_{as} \left[\frac{\bar{r}_{as} - \bar{r}_{КА}}{|\bar{r}_{as} - \bar{r}_{КА}|^3} - \frac{\bar{r}_{as}}{r_{as}^3} \right] + \sum_k \bar{F}_{add}^k, \quad (7)$$

где $\bar{r}_{КА} = \bar{r}_{КА}(t)$ — гелиоцентрический вектор положения КА в указанной ИСК; второе слагаемое в правых частях (7) представляет суммарное ускорение от притяжения крупных небесных тел (планет, Луны и др.); третье слагаемое — ускорение от притяжения рассматриваемого астероида (μ_{as} — гравитационная постоянная этого астероида), то положение $\bar{r}_{КА}(t)$ и скорость $\dot{\bar{r}}_{КА}(t)$ КА относительно астероида могут быть определены согласно зависимостям

$$\bar{\rho}_{КА}(t) = \bar{r}_{КА}(t) - \bar{r}_{as}(t), \quad \bar{v}_{КА}(t) = \dot{\bar{r}}_{КА}(t) - \dot{\bar{r}}_{as}(t), \quad (8)$$

в которых $\bar{r}_{КА}(t)$ и $\dot{\bar{r}}_{КА}(t)$ находятся в результате интегрирования уравнений (7) и (6) соответственно.

Для корректного описания движения КА относительно астероида-цели в состав действующих на него сил, строго говоря, должны быть включены все те же силы, действие которых учитывается в модели движения астероида, т. е. в уравнениях (6). При этом в состав дополнительных возмущающих ускорений $\{\bar{F}_{add}^k\}$ в уравнениях (7) могут, в случае необходимости, включаться ускорения, вызываемые влиянием нецентральной части гравитационного поля астероида.

В плане оценки возможностей предлагаемого высокоэффективного метода численного интегрирования для обеспечения решения задачи описания относительного движения КА с высокой точностью в рамках указанной постановки, было осуществлено совместное интегрирование уравнений движения астероида 2001 JV1 и совершающего вокруг него полёт гипотетического КА. В работе [Колюка, Афанасьева, 2015] было установлено, что у астероида 2001 JV1 может существовать спутник, совершающий полёт в пределах гравитационной сферы действия этого астероида, которая, вообще говоря, зависит от гелиоцентрического расстояния до астероида и определяется как

$$\rho_{сф} = r_{as} \times 0,45388 \cdot 10^{-8} \text{ а.е.}$$

На расстояниях порядка 2 а.е. от Солнца радиус этой сферы будет $\sim 1,36$ км.

Если предположить, что спутник астероида 2001 JV1 движется вокруг него по круговой орбите с радиусом $\rho \approx 500$ м, то его орбитальная скорость будет иметь величину $v \approx 0,019$ м/с, а период обращения составит $\sim 45,5$ ч.

В случае же нахождения спутника на орбите радиуса $\rho \approx 300$ м, аналогичные параметры его движения будут определяться как $v \approx 0,025$ м/с, $T \sim 21,1$ ч.

Предположим далее, что направленный к астероиду 2001 JV1 космический аппарат имеет в момент времени t_0 астероидоцентрические координаты и скорости $\bar{\rho}_{КА}(t_0)$, $\bar{v}_{КА}(t_0)$ в рассматриваемой ИСК.

Тогда координаты и скорости КА в гелиоцентрической ИСК на тот же момент времени определяются согласно следующим зависимостям:

$$\bar{V}_{KA}(t_0) = \bar{V}_{as}(t_0) + \bar{v}_{KA}(t_0) \frac{\bar{V}_{as}}{|\bar{V}_{as}|}(t_0), \quad \bar{r}_{KA}(t_0) = \bar{r}_{as}(t_0) + \frac{[\bar{V}_{as} \times \bar{b}]}{[\bar{V}_{as} \times \bar{b}]} \rho_{KA}(t_0),$$

где $\bar{b} = [\bar{r}_{as} \times \bar{V}_{as}]$.

В табл. 6 приведены начальные условия движения астероида 2001 JV1 и совершающего вблизи него полёт КА в прямоугольной гелиоцентрической геоэкваториальной системе координат эпохи J2000 на начальный момент времени $t_0 = 2022.05.08, 0^h$ TDB. Начальные условия движения КА соответствуют круговым орбитам спутника астероида радиуса $\rho = 300$ м (обозначены как КА_300) и радиуса $\rho = 500$ м (КА_500). Приведенные НУ движения астероида и КА использовались при оценке точности расчёта параметров относительного движения КА на основе результатов совместного интегрирования уравнений движения (6) и (7) с помощью предлагаемого численного метода.

Таблица 6. Начальные условия движения астероида 2001 JV1 и совершающего вблизи него полёт КА в прямоугольной гелиоцентрической геоэкваториальной системе координат эпохи J2000

Параметр	Начальный момент времени $t_0 = 2022.05.08, 0$ TDB		
	2001 JV1	КА_300	КА_500
X (км)	-0,4121717334265118D+07	-0,4121717257380990D+07	-0,4121717206124900D+07
Y (км)	0,3175584181849552D+09	0,3175584184549090D+09	0,3175584186348780D+09
Z (км)	0,1365029349297056D+09	0,1365029350356020D+09	0,1365029351061990D+09
V_x (км/с)	-0,1509847750857243D+02	-0,1509850122300760D+02	-0,1509849587742260D+02
V_y (км/с)	0,3041800380717994D+01	0,3041805158324060D+01	0,3041804081380890D+01
V_z (км/с)	0,3207758474355037D+01	0,3207763512623310D+01	0,3207762376923020D+01

ОЦЕНКА ТОЧНОСТИ ПРОГНОЗИРОВАНИЯ ДВИЖЕНИЯ АСЗ И ДВИЖЕНИЯ КА ОТНОСИТЕЛЬНО АСТЕРОИДА-ЦЕЛИ НА ДЛИТЕЛЬНЫХ ИНТЕРВАЛАХ ВРЕМЕНИ С ПОМОЩЬЮ ПРЕДЛАГАЕМОГО ЧИСЛЕННОГО МЕТОДА

Оценка точности прогнозирования движения рассматриваемых АСЗ и КА, совершающего полёт вблизи одного из этих астероидов, с помощью предлагаемого численного метода, осуществлялась путём сравнения результатов интегрирования уравнений движения соответствующих объектов по замкнутому маршруту — от начального момента времени t_0 до заданного момента времени $t_{зад}$ в прямом направлении течения времени и далее с возвратом к моменту t_0 при обратном направлении течения времени — с исходными начальными условиями. Подобное сравнение результатов интегрирования с НУ позволяло оценить ошибку интегрирования на интервале времени $2(t_{зад} - t_0)$ в виде рассогласова-

ния $\Delta\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$, где Δx , Δy , Δz — получившиеся отличия в координатах между результатами прогноза и НУ. Эта схема оценки точности применяемого метода использовалась как в случае интегрирования уравнений движения каждого из рассматриваемых АСЗ, так и в случае интегрирования совместной системы уравнений движения АСЗ и совершающего вблизи него полёт КА.

В табл. 7 приведены ошибки прогнозирования движения астероидов, сближающихся с Землёй, *Апофиса* и *2001 JV1*, для разных интервалов времени, которые охватывали периоды разной динамики движения этих тел.

Таблица 7. Ошибки прогнозирования движения АСЗ предлагаемым численным методом на разных интервалах времени

1 КО	Начало движения	Интервал прогнозирования (вперёд + назад), годы	Ошибка прогноза $\Delta\rho$ (км)
Апофис	24.09.2008	50 (25+25)	1,1
		100 (50 + 50)	6,8
		200 (100 + 100)	17,3
Апофис	01.04.2032	50 (25+25)	0,00009
		100 (50 + 50)	0,00055
		200 (100 + 100)	0,00143
2001 JV1	04.11.2013	50 (25+25)	0,00006
		100 (50 + 50)	0,00011
		200 (100 + 100)	0,00074

В качестве начальных моментов движения астероида *Апофис* были выбраны две даты: 24.09.2008 г. и 01.04.2032 г.

Для астероида *2001 JV1* было рассмотрено только одно начало движения — от даты 04.11.2013 г.

Для каждых начальных условий осуществлялось прогнозирование движения АСЗ на интервалы времени:

- 1) 50 лет (25 лет вперед + 25 лет назад),
- 2) 100 лет (50 лет вперед + 50 лет назад),
- 3) 200 лет (100 лет вперед + 100 лет назад).

Для первых начальных условий астероида *Апофис* все интервалы прогнозирования охватывали тесное сближение этого АСЗ с Землёй в апреле 2029 г., когда орбита этого астероида в результате совершенного гравитационного манёвра будет сильно изменена. Ошибка прогнозирования движения *Апофиса* в этом случае зависела от длительности интервала прогнозирования — по мере возрастания длины этого интервала ошибка прогнозирования увеличивалась. Вместе с тем, как можно заключить из табл. 7, даже в случае такой сложной динамики полёта данного астероида ошибка прогнозирования его движения на момент сближения с Землёй в 2036 г. не будет превосходить единиц километров.

В случае же более спокойной динамики полёта астероида *Апофис* на интервале времени после 2032 г. точность прогнозирования его движения с помощью предлагаемого метода оказалась очень высокой: даже на интервале прогнозирования в 200 лет ошибка прогнозирования оценивалась величиной $\Delta\rho < 2$ м.

Очень высокая точность прогнозирования движения может быть отмечена и для астероида *2001 JV1*. В этом случае ошибка прогнозирования движения астероида на 200 лет вообще была меньше 1 м.

Можно обосновано предположить, что на относительно непродолжительных интервалах прогнозирования точность расчёта параметров движения АСЗ с помощью данного метода численного интегрирования будет оставаться достаточно высокой даже в случаях присутствия в пределах интервала прогнозирования участков полёта АСЗ с тесными их сближениями с Землёй или другими большими планетами (при условии, что исходные параметры орбит этих АСЗ будут определены достаточно точно).

В плане оценки возможности корректного описания относительного движения КА в рамках описанного выше подхода, предусматривающего осуществление совместного интегрирования уравнений движения АСЗ и КА в гелиоцентрической ИСК, был выполнен следующий эксперимент. От начальных условий движения астероида *2001 JV1* и КА, изначально находящегося на орбите искусственного спутника этого астероида, заданных в гелиоцентрической ИСК и представленных в табл. 6, осуществлялось совместное интегрирование уравнений движения обоих этих КО и далее, согласно формулам (8), определялись вектора состояния КА $\{\bar{\rho}_{КА}(t), \bar{v}_{КА}(t)\}$ относительно астероида на интервале времени 1 год с шагом выдачей результатов 1 ч. Результаты расчётов показали, что в течение всего времени на данном интервале прогнозирования КА оставался на околоастероидной орбите, близкой к исходной. Для КА с начальной околокруговой орбитой вокруг астероида, имевшей высоту 500 м, расстояния от астероида до КА в пределах данного интервала времени периодически изменялись в пределах от 500 до 496 м. Результаты прогнозирования вектора положения КА относительно астероида для исходной орбиты с высотой 500 м в плоскости этой орбиты представлены на рис. 7. Здесь показанная синим цветом линия представляет собой множество выходных точек расчёта вектора положения КА $\bar{\rho}_{КА}(t)$ в координатной плоскости xu астероидоцентрической ИСК.

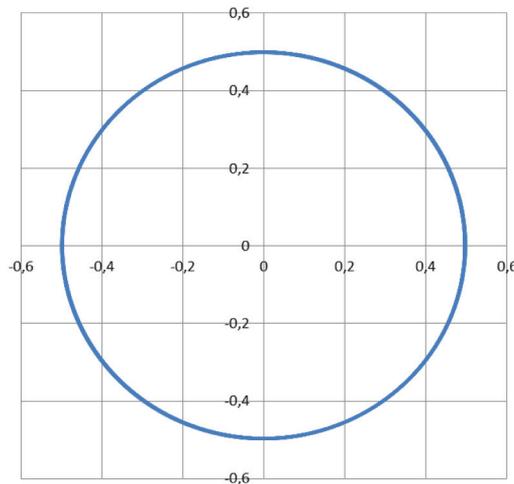


Рис. 7. Результат прогноза относительного движения КА по круговой орбите высотой 500 м в астероидоцентрической системе координат (по вертикальной оси координат y (км), по оси абсцисс — координата x (км))

ЗАКЛЮЧЕНИЕ

Изложены основные положения оригинального высокоэффективного универсального численного метода высокого порядка для интегрирования уравнений движения небесных тел, разработанного в Центре управления полётами и успешно применяемого здесь при решении широкого спектра баллистико-навигационных и эфемеридных задач для различных космических объектов искусственного и естественного происхождения, совершающих полёты в ближнем и дальнем космосе.

Исследованы и оценены возможности применения данного метода для решения прецизионных задач динамики полёта астероидов, сближающихся с Землёй, и направляемых к ним КА на примерах двух АСЗ — *Анофис* и *2001 JV1*, имеющих существенно разные орбиты, а также — гипотетического КА, совершающего полёт вокруг АСЗ *2001 JV1*.

Приведены результаты исследования динамики движения по орбите и сближения с Землёй на интервалах времени 100 лет АСЗ *Анофис* и *2001 JV1*, а также динамики относительного движения КА, находящегося на орбите искусственного спутника астероида *2001 JV1* на интервале 1 год.

Выполненная оценка точности прогнозирования движения АСЗ *Анофис* и *2001 JV1* на разных длительных интервалах времени, вплоть до 200 лет, свидетельствует о высоких точностных характеристиках предлагаемого метода и возможности решения сложных задач динамики полёта небесных тел, требующих высокой точности расчётов, с помощью данного метода.

ЛИТЕРАТУРА

- [Колюка, 2018] *Колюка Ю. Ф.* Высокоэффективный метод численного интегрирования уравнений движения в задачах небесной механики и космической баллистики // *Космонавтика и ракетостроение*. 2018. Вып. 1(100). С. 100–117.
- [Колюка, Афанасьева, 2015] *Колюка Ю. Ф., Афанасьева Т. И.* Проектирование баллистической схемы полёта КА к потенциально опасному астероиду 2001 JV1 с целью изменения его орбиты взрывом // *Космическая баллистика — от истоков к будущему: сб. тр. научного семинара, посвященного 100-летию со дня рождения П. Е. Эльясберга*. М.: ИКИ РАН, 2015. С. 172–193.
- [Brozović et al., 2018] *Brozović M., Benner L. A. M., McMichael J. G., Giorgini J. D., Pravec P., Scheirich P., Magri C., Busch M. W., Jao J. S., Lee C. G., Snedeker L. G., Silva M. A., Slade M. A., Semenov B., Nolan M. C., Taylor P. A., Howell E. S., Lawrence K. J.* Goldstone and Arecibo radar observations of (99942) Apophis in 2012–2013 // *Icarus*. 2018. V. 300. P. 115–128.

APPLICATION OF THE HIGH-EFFECTIVE METHOD FOR NUMERICAL INTEGRATION OF THE DIFFERENTIAL EQUATION OF MOTION TO SOLVE THE PROBLEMS OF FLIGHT DYNAMICS OF ASTEROIDS ENCOUNTER THE EARTH AND THE SPACE CRAFTS DESIGNED TO RENDEZVOUS WITH THESE ASTEROIDS

Y. Kolyuka, T. Afanasieva

Central Research Institute of Machine Building (TsNIIImash)

The questions of flight dynamics of asteroids, that approach close the Earth, and flying near these asteroids spacecrafts are considered. The basic principles of the original high-efficient (in sense of accuracy and operating speed) numerical method for integration of differential equations of motion of various space objects, which can be recommended for solving the complicated asteroid flight dynamics problems is given. The results of application of this method for prediction of the approaches of asteroids Apophis and 2001 JV1 to the Earth within the long-term intervals of time are given. The precision of long-term (up to 200 years) motion prediction of these asteroids by means of the given method is estimated.

Keywords: asteroids approaching the Earth, the minimum orbit intersection distance, flight dynamics, method for numerical integration, prediction the Earth encounters, motion prediction accuracy

Kolyuka Yury — head of laboratory, PhD, senior scientist, yfk@mcc.rsa.ru
Afanasieva Tatiana — leader scientist, PhD, yfk@mcc.rsa.ru

ПРИМЕНЕНИЕ КИНЕТИЧЕСКОГО УДАРА ДЛЯ ВОЗДЕЙСТВИЯ НА МАЛЫЕ ОПАСНЫЕ АСТЕРОИДЫ ПРИ ИХ ПОЗДНЕМ ОБНАРУЖЕНИИ

В. С. Сазонов

Центральный научно-исследовательский институт машиностроения
(ЦНИИМаш), Королёв, МО, Россия

Рассмотрены вопросы отклонения от Земли опасных небесных тел с помощью высокоскоростного удара массивным телом, доставляемым ракетой, начальная масса которой является параметром задачи. Определены условия существования оптимальной стратегии ближнего перехвата, обеспечивающей максимальное отклонение астероида малого размера без нарушения его целостности.

Ключевые слова: опасные небесные тела (ОНТ), ближний перехват, высокоскоростной удар, обнаружение, фрагментация, отклонение

ВВЕДЕНИЕ

Наибольшую опасность с точки зрения частоты падения на Землю представляют собой малые опасные небесные тела размером от нескольких десятков до сотен метров. От них защититься труднее всего, что связано с трудностями заблаговременного обнаружения и отклонения их на достаточно большое расстояние при малом запасе времени, соответствующем ближнему перехвату. К этому классу относятся, в частности, Тунгусский и Челябинский метеориты, продемонстрировавшие не так давно реальность астероидной опасности. Воздействие ядерного взрыва на малые астероиды приводит к нежелательной фрагментации [Сазонов, 2015] и ухудшению экологической обстановки, поэтому представляет интерес исследование возможности применения менее энергоёмкого способа — высокоскоростного удара искусственным телом. Это экологически чистый способ. Ниже предложена модель воздействия, использующая известную аналогию физических явлений, происходящих при высокоскоростном ударе и контактном ядерном взрыве (например, [Станюкович, 1960; Kreyenhausen, Schuster, 1977]). При этом учитываются обширные экспериментальные данные, относящиеся к ядерному взрыву [Адушкин, Спивак, 2007]. Далее решается задача о максимальном отклонении астероида при ближнем его перехвате с учётом найденных условий фрагментации, служащих ограничениями.

Сазонов Валентин Сергеевич — ведущий научный сотрудник, кандидат технических наук, старший научный сотрудник, valsazonov@mail.ru

МОДЕЛЬ ВОЗДЕЙСТВИЯ

В результате высокоскоростного удара из образующегося кратера выбрасывается испарившееся, расплавленное и раздробленное вещество, так что создаётся возмущающий импульс, изменяющий первоначальную траекторию астероида. Наименьшая (критическая) относительная скорость ударника v_k , при которой наблюдается взрывная аналогия, равна

$$v_k = \sqrt{2\varepsilon}, \quad (1)$$

где ε — энергия испарения единицы массы. При скоростях $v \geq v_k$ происходит испарение, плавление, дробление и образование трещин. Для астероида, сложенного, например, из скальной породы, $\varepsilon = 2 \cdot 10^7$ Дж/кг и $v_k \approx 6,3$ км/с [Станюкович, 1960]. В космических экспериментах легко достигается и более высокая скорость столкновения (например, в эксперименте Deep Impact $v \sim 12$ км/с). В таких условиях чисто механическая компонента импульса, пропорциональная относительной скорости ударника, намного меньше взрывной компоненты, пропорциональной квадрату скорости, и ей можно пренебречь. Главным отличием высокоскоростного удара от ядерного взрыва с энергетической точки зрения является гораздо меньший масштаб явления, вследствие чего реализуются меньшие температуры образовавшейся плазмы (отсутствует тепловая волна) и меньшие потери энергии на излучение. В результате, значения коэффициентов передачи энергии астероиду χ и энергии, пошедшей на выброс ξ , при высокоскоростном ударе гораздо больше, чем при ядерном взрыве. В случае удара эти значения соответственно равны $\chi_y = 0,6$ и $\xi_y = 0,5$ [Kreyenhagen, Schuster, 1977], а в случае контактного ядерного взрыва $\chi_b = 0,056$ и $\xi_b = 0,04$ [Адушкин, Спивак, 2007]. Отметим, что высокоскоростной удар обеспечивает больший вклад энергии, чем обычный химический взрыв.

При оценке импульса, передаваемого астероиду, ограничимся вкладом наиболее высокоскоростных малоразмерных частиц, выбрасываемых на ранней стадии из полусферической котловой области, формируемой под действием сильной ударной волны.

Вклад низкоскоростных крупных фрагментов, образующихся на поздней стадии формирования кратера, отличающегося от полусферы, далее не учитывается вследствие его малости (энергия выброшенного вещества, через которую выражается импульс, пропорциональна квадрату скорости и массе). Для оценок соответствующей выброшенной массы M_k возьмём за основу экспериментальные и расчётные данные по подземным ядерным взрывам [Адушкин, Спивак, 2007; Брод, 1975], которые можно пересчитать на случаи контактного взрыва и рассматриваемого высокоскоростного удара. На рис. 1 (см. с. 261) для контактного взрыва построены зависимости радиусов разрушения R от величины заряда E , используемые далее в качестве опорных (вследствие отмеченной аналогии).

Линии на рисунке соответствуют следующим областям: 1 — тепловая волна; 2, 3 — испарение [Адушкин, Спивак, 2007; Сазонов, 2013]); 4 — плавление; 5 — котловая область; 6 — смятие; 7 — дробление; 8 — зона новых трещин; 9 — зона подновлённых трещин (блоковой трещиноватости); 10 — зона неупругих деформаций (единичные новые трещины); 11 — зона локальных неупругих деформаций. Примечательно, что радиусы испарения, посчитанные по инженерной методике [Сазонов, 2013] при плотности астероида $\rho = 3400$ кг/м³

и указанном значении ϵ , совпадают с работой [Адушкин, Спивак, 2007] (линии 2 и 3 на рис. 1). Эти данные можно использовать и в задаче о высокоскоростном ударе с той оговоркой, что энергии удара существенно ниже, чем при ядерном взрыве, и, следовательно, тепловая волна (нижняя линия 1 на рис. 1) отсутствует.

Из соображений размерности можно записать уравнение баланса энергии для зоны испарения радиусом $R_{\text{и}}$

$$\chi E = A_{\text{и}} \rho \epsilon R_{\text{и}}^3, \quad (1)$$

где E — энергия взрыва или удара; ρ — плотность вещества астероида; $A_{\text{и}}$ — подлежащий определению безразмерный коэффициент. Для подземного взрыва с плотным окружением имеем $\chi = \chi_{\text{в}} = 1$, и в случае гранита плотности 2670 кг/м^3 при энергии взрыва 1 кт радиус испарения равен 1,83 м [Бруд, 1975], так что из уравнения (1) получаем $A_{\text{и}} \approx 12,8$. Можно считать, что при воздействии контактного взрыва и высокоскоростного удара на астероид соблюдается геометрическое подобие между радиусами однотипных разрушений. Например, отношение радиуса котловой области $R_{\text{к}}$, из которой выбрасывается вещество массой $M_{\text{к}}$, к радиусу $R_{\text{и}}$ в обоих случаях равно $k_{\text{к}} = R_{\text{к}}/R_{\text{и}} \approx 4,8$ [Адушкин, Спивак, 2007]. Выразив отсюда $R_{\text{и}}$ и подставив в равенство (1), получим соотношение для определения радиуса котловой области $R_{\text{к}}$ при ударе

$$\chi_{\text{у}} E_{\text{у}} = A_{\text{и}} \rho \epsilon \left(\frac{R_{\text{к}}}{k_{\text{к}}} \right)^3, \quad (2)$$

где кинетическая энергия удара $E_{\text{у}}$ зависит от массы ударника $M_{\text{у}}$ и относительной скорости v удара:

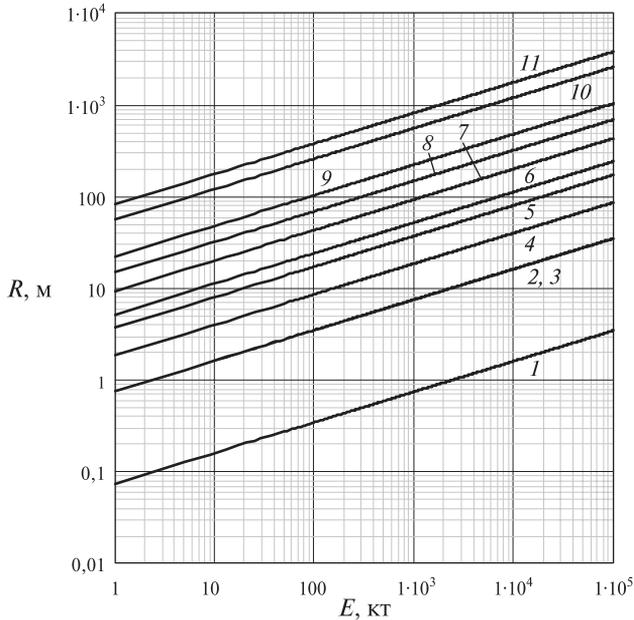


Рис. 1 Зависимости радиусов разрушения астероида от мощности контактного взрыва

$$E_y = \frac{M_y v^2}{2}. \quad (3)$$

Выброшенная из полусферической котловой области масса равна

$$M_k = \frac{2}{3} \pi \rho R_k^3. \quad (4)$$

Среднюю скорость выбросов u определим из уравнения энергии

$$\frac{M_k u^2}{2} = \xi_y E_y. \quad (5)$$

Если M_a — масса астероида, то пренебрегая потерями массы M_k (т.е. считая, что $R_k \ll d$, где d — диаметр астероида), получим из выражений (2)–(5) приращение скорости астероида при ударе:

$$\Delta V = \frac{M_k u}{M_a} = b_y \frac{M_y v^2}{M_a}, \quad \left(b_y = \sqrt{\frac{\pi \chi_y \xi_y k_k^3}{3 A_{и} \varepsilon}} \right). \quad (6)$$

ОПТИМАЛЬНЫЙ ПЕРЕХВАТ АСТЕРОИДА

Будем пренебрегать притяжением Земли и временем активного полёта ракеты, доставляющей к астероиду массивное тело, считая, что перехват происходит на встречном курсе с Земли или с околоземной орбиты [Solem, 1993]. Скорости ракеты относительно Земли и относительно астероида в момент удара, соответственно, равны

$$v_r = g I_y \ln \frac{m_0}{M_y}, \quad v = v_r + v_a, \quad (7)$$

где g — гравитационная постоянная Земли; I_y — удельный импульс ракеты (для двигателя на химическом топливе принимается $I_y = 500$ с); m_0 — начальная масса ракеты, несущей тело ударника массой M_y ; v_a — скорость сближения астероида с Землёй. Если импульс перпендикулярен скорости астероида, то отклонение L с учётом равенств (6) и (7) равно

$$L = t_a \Delta V \left(\frac{v_r}{v_a + v_r} \right) = \frac{b_y c t_a}{M_a} M_y \left(v_a + c \cdot \ln \frac{m_0}{M_y} \right) \ln \frac{m_0}{M_y}, \quad (8)$$

где $c = g I_y$; $t_a = (r_a - v_a t_{п}) / v_a$ — время полёта астероида с момента запуска ракеты, спустя время подготовки $t_{п}$ (далее принимается $t_{п} = 0$); r_a — начальное удаление астероида (предполагается, что оно много больше радиуса Земли). Из уравнения (8) видно, что при некотором $M_y = M_y^*$ отклонение максимально ($L = L^*$), что и определяет оптимальную стратегию. На рис. 2 (см. с. 263) приведены отклонения астероидов L диаметром $d = 200$ м и $d = 100$ м в зависимости от массы ударника M_y при начальном удалении $r_a = 0,3$ а.е. и начальной массы ракеты-носителя $m_0 = 1$ т, 10, 100 и 1000 т.

Отношение m_0 / M_y^* , соответствующее максимуму отклонения астероида, определяется выражением

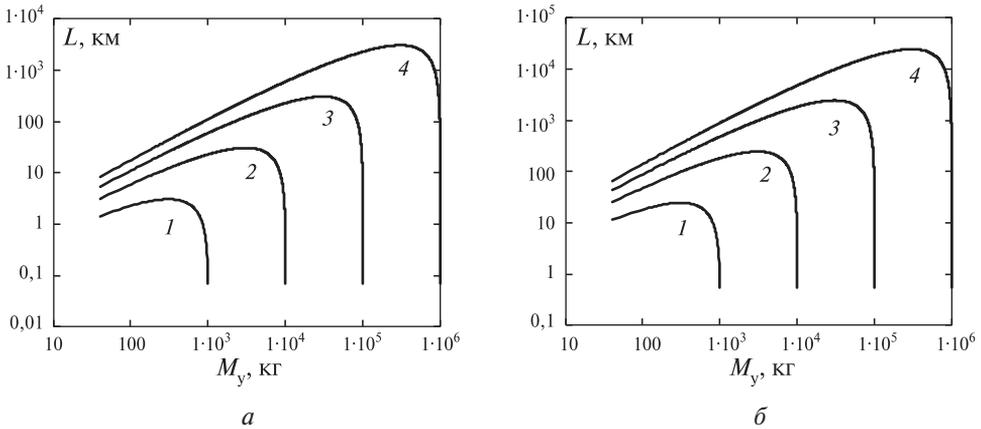


Рис. 2. Отклонения астероида диаметром 200 м (а) и 100 м (б) в зависимости от массы тела-ударника при различной начальной массе ракеты (1 — $m_0 = 1$ т; 2 — $m_0 = 10$ т; 3 — $m_0 = 100$ т; 4 — $m_0 = 1000$ т)

$$q = \ln \frac{m_0}{M_y^*} = 1 - \frac{v_a}{2c} + \sqrt{1 + \left(\frac{v_a}{2c}\right)^2}, \quad (9)$$

причём оптимальная скорость ракеты равна $v_r^* = cq$. Подставив выражение (9) в уравнение (8), получим максимальное отклонение

$$L^* = \frac{b_y t_a M_y^*}{M_a} (v_a + v_r^*) v_r^*, \quad M_y^* = \frac{m_0}{e^q}. \quad (10)$$

УСЛОВИЯ РАЗРУШЕНИЯ АСТЕРОИДА

Запишем условия фрагментации астероида. Радиусы дробления и трещинообразования равны $R_d = k_d R_n$ и $R_T = k_T R_n$, где $k_d \approx 11,1$ и $k_T \approx 31,7$ [Адушкин, Спивак, 2007]. С учётом этого, воспользовавшись формулами (1), (4)–(8) при $L = L^*$, получим относительные величины f_d и f_T , характеризующие степень разрушения при ударе,

$$f_d = \frac{R_d}{d} = \left(\frac{L^* k_d^3}{12 t_a} \sqrt{\frac{3\pi\chi_y}{\xi_y A_n \varepsilon k_k^3}} \cdot \frac{v_a + v_r^*}{v_r^*} \right)^{1/3}, \quad f_T = \frac{R_T}{d} = \left(\frac{L^* k_T^3}{12 t_a} \sqrt{\frac{3\pi\chi_y}{\xi_y A_n \varepsilon k_k^3}} \cdot \frac{v_a + v_r^*}{v_r^*} \right)^{1/3}. \quad (11)$$

Форма этих соотношений та же, что и при действии контактного ядерного взрыва, за исключением значений коэффициентов χ_y и ξ_y , а также оптимальных значений v_r^* и L^* [Сазонов, 2015].

Условия, когда дробление или трещинообразование достигает задней границы астероида, выражаются равенствами $f_d = 1$ и $f_T = 1$. Относительная граница котловой области определяется равенством $f_k = R_k/d = f_d|_{k_d=k_k} = f_T|_{k_T=k_k}$.

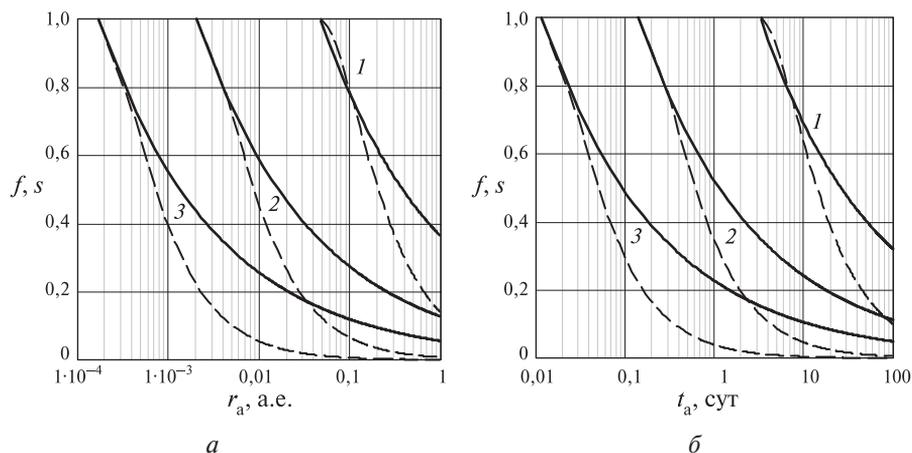


Рис. 3. Зависимости коэффициентов ударного разрушения астероида f (сплошные линии) и s (пунктир) от расстояния до астероида r_a (а) и запаса времени t_a (б): 1 — трещины, 2 — дробление, 3 — котловая область

Охват этой областью всего астероида происходит при $f_k = 1$. О степени разрушения иногда судят по отношению s массы разрушенной области к полной массе тела [Solem, 1993]. На рис. 3 изображены зависимости коэффициентов разрушения по радиусу (f — сплошные линии) и по массе (s — пунктир) от расстояния до астероида r_a (а) и запаса времени t_a (б) при $\rho = 3400 \text{ кг/м}^3$, $v_a = 25 \text{ км/с}$ и $L^* = 1000 \text{ км}$ (отклонение в океан). Кривые 1, 2 и 3 соответствуют областям трещинообразования, дробления и котловой области. Заметим, что они справедливы для тел произвольного диаметра. Приведённые графики показывают, какая часть тела (по радиусу и по массе) в результате удара будет охвачена трещинами (1), будет раздроблена (2) и превращена в выбрасываемый котловой газ (3), если запуск ракеты с телом-ударником осуществить в момент обнаружения t_a при дальности r_a .

Например, астероид будет полностью охвачен трещинами, если $t_a \leq 3,2 \text{ сут}$ и $r_a \leq 0,05 \text{ а.е.}$ При этом будет раздроблена примерно десятая часть массы астероида (точнее, $s_d = 0,127$). Наоборот, при больших значениях t_a и r_a тело не разрушится.

ОЦЕНКИ МАКСИМАЛЬНОГО ОТКЛОНЕНИЯ АСТЕРОИДА

На рис. 4 (см. с. 265) показаны зависимости времени, необходимого для отклонения астероида в океан (на расстояние $L = 1000 \text{ км}$) и в случае пролёта мимо Земли (на расстояние $L = 7000 \text{ км}$), от диаметра при различных массах оптимального перехватчика, а также аналогичные зависимости времени упреждения воздействия. Верхние, средние и нижние горизонтальные пунктирные линии соответствуют полному растрескиванию астероида, его дроблению и переводу в газ в котловой области. Им отвечают моменты 3,2 сут, 0,14 сут и 0,012 сут — при отклонении на 1000 км; и 22,7 сут, 1 сут и 0,08 сут — при отклонении на 7000 км.

Основной вывод, согласно результатам, представленным на рис. 4, заключается в следующем. Чем крупнее астероид, тем раньше (на большем удалении) его нужно обнаружить, чтобы отклонить на требуемое расстояние, причём применение тела-ударника меньшего веса позволяет отклонить без разрушения более мелкие объекты. Так, наименьший диаметр астероида, который можно отклонить без трещинообразования в малонаселённый пункт, например, в океан (на 1000 км) с помощью запуска лёгкой ракеты массой 1 т, равен ~15 м (см. рис. 4а); масса ударника при этом составляет ~310⁻³ кг. Примером такого опасного тела служит Челябинский метеорит. Заблаговременно обнаружить его, а потом отклонить, из-за малых размеров очень сложно. Запуск ракеты должен быть осуществлён не позже, чем за 3,2 сут до столкновения с Землёй, когда расстояние до астероида составляет ~0,05 а.е. Если за критерий разрушения взять дробление, то размер отклоняемого на данное расстояние тела, а также время заблаговременного воздействия и расстояние до объекта, будут ещё меньше — соответственно, 11,5 м, 0,14 сут и 2·10⁻³ а.е. Но масса необходимого ударника при этом возрастёт в 10 раз по сравнению со случаем, когда за критерий разрушения взято трещинообразование, т.е. она равна 3100 кг. Такой массы достаточно, чтобы отклонить без трещинообразования в океан тело диаметром 32,7 м.

Для отклонения на 7000 км того же астероида диаметром $d \sim 15$ м требуется в семь раз большее время, чем в предыдущем случае, а именно $t_a \approx 22,7$ сут (см. рис. 4б). При этом расстояние до астероида в момент запуска ракеты равно ~0,35 а.е.

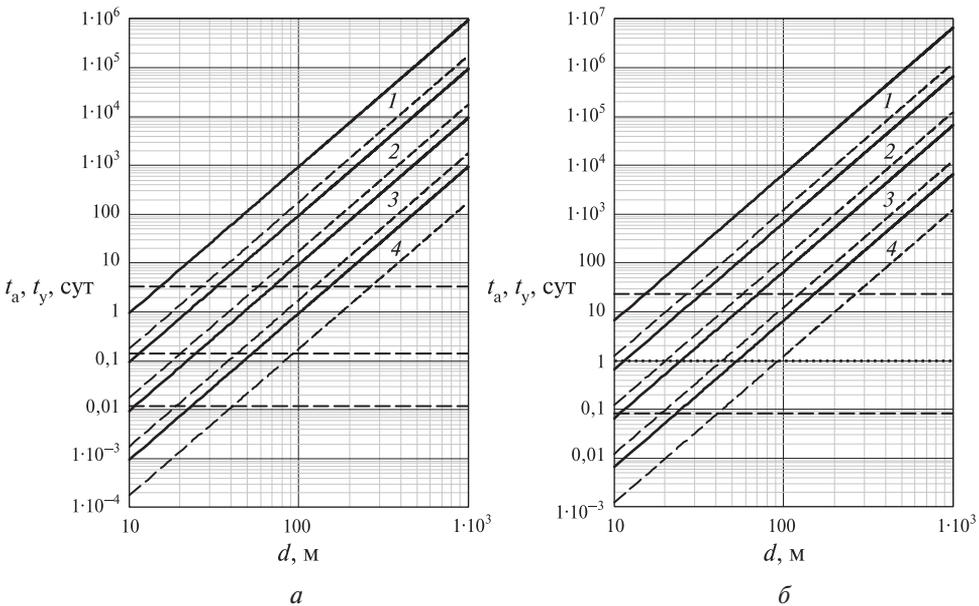


Рис. 4. Зависимости запаса времени t_a (сплошные линии) и времени упреждения t_y (наклонные пунктирные линии), требующихся для отклонения на 1000 км (а) и на 7000 км (б) с помощью удара, от диаметра астероида d при различных массах оптимально сконструированного перехватчика ($v_a = 25$ км/с; 1 — $m_0 = 1$ т, 2 — $m_0 = 10$ т, 3 — $m_0 = 100$ т, 4 — $m_0 = 1$ кг)

ЗАКЛЮЧЕНИЕ

Представленная методика позволяет выбрать массы ракет и тела-ударника, обеспечивающих максимальное отклонение астероида с помощью ударно-кинетического способа противодействия в случае ближнего перехвата, в зависимости от ресурса времени и дальности обнаружения при ограничениях на фрагментацию. Для защиты от тел малого размера ($15 < d < 67$ м) при небольшом запасе времени (примерно $3 < t_a < 5$ сут) предлагается отклонять их без фрагментации, например, в океан — на 1000 км. Масса ударника при этом составляет ~310 кг. Для отклонения на 7000 км (пролёт астероида мимо Земли) с помощью того же ударяющего тела требуется время не менее 23 сут. Применение ударно-кинетического метода для отклонения астероидов диаметром $d > 100$ м при запасе времени $t_a < 5$ сут не является рациональным, так как такой способ требует больших начальных масс ракет-носителей, при которых вероятность разрушения астероида быстро возрастает.

ЛИТЕРАТУРА

- [Адушкин, Спивак, 2007] *Адушкин В. В., Спивак А. А.* Подземные взрывы / Институт динамики геосфер РАН. М.: Наука, 2007. 578 с.
- [Бруд, 1975] *Бруд Г. Л.* Расчёт взрывов на ЭВМ. Подземные взрывы / пер. с англ. М.: Мир, 1975. 163 с.
- [Сазонов, 2013] *Сазонов В. С.* Взрыв как средство отклонения от Земли опасных небесных тел при их раннем и позднем обнаружении // Экологич. вестн. научных центров ЧЭС. 2013. № 4. Т. 3. С. 118–122.
- [Сазонов, 2015] *Сазонов В. С.* Оптимальная стратегия ближнего перехвата опасных для Земли небесных тел в случае применения ядерных средств противодействия // Околосная астрономия 2015: тр. международной конф. 31 августа — 5 сентября 2015, Терскол. 2015. С. 222–227.
- [Станюкович, 1960] *Станюкович К. П.* Элементы теории удара с большими космическими скоростями // Искусственные спутники Земли. 1960. Вып. 4. С. 86–117.
- [Kreyenhagen, Schuster, 1977] *Kreyenhagen K. N., Schuster S. H.* Review and comparison of hypervelocity impact and explosion cratering calculations // Impact and Explosion Cratering / ed. Roddy D. J., Pepin R. O., Merrill R. B. N. Y.: Pergamon Press, 1977. P. 983–1002.
- [Solem, 1993] *Solem J. C.* Interception of Comets and Asteroids on Collision Course with Earth // J. Spacecraft and Rockets. 1993. V. 30. No. 2. P. 222–228.

USE OF KINETIC IMPACT EFFECT ON SMALL DANGEROUS ASTEROIDS AT THEIR LATE DETECTION

V. S. Sazonov

Central Research Institute for Machine building (TSNIIMASH)

The questions of the deviation from the Earth of dangerous celestial bodies by means of high-velocity impact are considered. The conditions of optimum near interception strategy providing the maximum deviation of the small asteroid are defined in case of restrictions on the fragmentation.

Keywords: dangerous celestial bodies (CDB), near interception, high-velocity impact, detection, fragmentation, deviation

Sazonov Valentin Sergeievich — leading scientific employee, PhD, senior researcher, valsa-zonov@mail.ru

АСТЕРОИДЫ КАК РЕСУРС НА БУДУЩЕЕ

М. Д. Сизова, Е. Ефремова

Институт астрономии Российской академии наук (ИНАСАН)

В статье рассмотрено распределение по химическому составу астероидов, у которых, согласно сайту asterank.com, определён таксономический класс, проанализированы астероиды в Главном поясе астероидов и в околоземном пространстве. Осуществлён обзор космических миссий к астероидам, как уже состоявшихся, так и планируемых. Проведена оценка возможностей перелёта к выбранному астероиду, наиболее обоснованному с коммерческой точки зрения.

Ключевые слова: околоземные астероиды, астероиды, добыча полезных ископаемых, таксономические классы, состав астероидов

ВВЕДЕНИЕ

Современные работы по изучению популяции астероидов, сближающихся с Землёй, ведутся не только в направлении астероидно-кометной опасности, но и в целях освоения малых тел как источника ценных природных ресурсов. Прогнозирование запасов и истощения полезных ископаемых на Земле затруднено ввиду большого количества влияющих факторов, таких как рост популяции, открытие новых месторождений и методов обнаружения месторождений, рост динамики потребления вследствие технологических инноваций и прочее. Но, несмотря на возможное истощение тех или иных ресурсов на Земле, в ближайшем будущем добыча полезных ископаемых на астероидах необходимо рассматривать в большей степени как источник сырья именно для освоения космоса, нежели для удовлетворения потребностей на нашей планете. Например, углеродистые астероиды и кометы богаты запасами льда, который можно использовать как источник водорода или питьевой воды для космических миссий (колоний). Например, один астероид диаметром 200 м может стоить миллионы долларов, а общее отношение запасов элементов платиновой группы на астероидах к земным составляет 105:1 (*Маров М. Я.* Астероиды как объект исследований и источник ресурсов: доклад // Расширенный научный семинар «Управление движением естественных небесных тел». 21–23 марта 2018, Таруса), что позволяет говорить о потенциале коммерческого освоения малых тел Солнечной системы. В недалёком будущем использование астероидов в виде полезных ископаемых станет важным шагом в эволюции человечества и позволит распространиться человечеству в пределах Солнечной системы.

АСТЕРОИДЫ, СБЛИЖАЮЩИЕСЯ С ЗЕМЛЁЙ

Очевидно, что для потенциальной разработки полезных ископаемых необходимо рассматривать технически достижимые для нас астероиды, коими являются

Сизова Мария Дмитриевна — младший научный сотрудник, аспирант, sizova@inasan.ru
Ефремова Екатерина Викторовна — младший научный сотрудник, аспирант, efremova@inasan.ru

так называемые околоземные астероиды (астероиды, сближающиеся с Землёй) с перигелием орбиты менее 1,3 а. е. Общее количество открытых околоземных астероидов диаметром от 30 м и крупнее оценивается более чем 800 тыс. объектов. При изучении околоземных астероидов учёные сталкиваются с определёнными сложностями — малые тела размером 50 м трудно обнаруживать на больших расстояниях, дистанционными методами можно приблизительно оценить состав астероида. Но для осуществления миссии необходимо точно знать параметры движения астероида и его физико-химические характеристики, так как без этих данных невозможно ни спланировать посадку на объект, ни оценить коммерческую выгоду.

Для исследования орбитальных и физико-химических характеристик астероида применяют дистанционные (оптические и спектральные) методы и методы исследования с помощью космических аппаратов. Среди наземных методов наблюдения наиболее распространёнными являются астрометрические и спектро-фотометрические наблюдения, которые позволяют уточнить орбиту небесного тела и исследовать отражательные характеристики поверхности астероида — получить спектральные кривые. При этом полученный спектр астероида сравнивается со спектрами изученных метеоритов, и на этом основании делаются выводы о его химическом составе. В некоторых случаях применяют метод радиолокационного зондирования, с помощью которого можно исследовать поступательно-вращательное движение астероидов, уточнять форму и параметры вращения и физические и минералогические характеристики поверхностного слоя.

Геохимический анализ метеоритов, найденных на Земле, показывает соответствие их состава со спектрами наблюдаемых астероидов. С целью классификации малых тел по геохимическим признакам были введены системы, наиболее широко используемые среди которых — системы Толена и SMASS [Tholen, 1989; Bus, Binzel, 2002]. Общей чертой обеих систем является то, что они определяют три широкие категории, включающие большинство известных объектов. Так, астероиды, богатые углеродом, с заметными признаками содержания воды и некоторых органических молекул составляют около 75 % популяции Главного пояса и обозначаются как С-класс. Этот класс астероидов наиболее многочислен в его внешней части [Gradie et al., 1989]. Вторыми по распространённости являются астероиды, состоящие из каменных образований и занимающие 17 % популяции Главного пояса. Этот класс назван S-классом и преобладает в его внутренней части на расстоянии 2,2 а. е., в средней части они встречаются в основном на удалении до 3 а. е., дальше этой границы они встречаются редко. Почти все остальные астероиды Главного пояса состоят из металлических сплавов и названы М-класс. При этом статистика метеоритов среди падений на Землю показывает иную картину — около 90 % составляют каменные тела [Шустов, 2018].

По данным сайта asterank.com, созданного компанией Planetary Resources как база наиболее коммерчески привлекательных астероидов, были проанализированы таксономические классы 932 тел Главного пояса (рис. 1, см. с. 270). Углеродистые астероиды, имеющие малую плотность материала, наиболее многочисленны. Также в выборке присутствуют астероиды, в составе которых есть силикаты, богатые углеродом или органическими веществами, возможно, со следами водяного льда (D-класс). Металлосодержащие астероиды (X-класс) наиболее ценны с точки зрения коммерческого освоения и потому присутствуют в выборке в большом количестве.

Распределение исследованных астероидов Главного пояса по таксономическим классам, 932 астероида

ПРОЦЕНТЫ

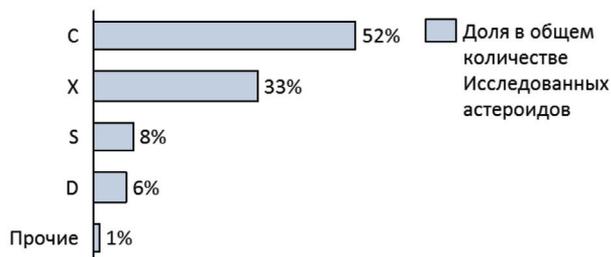


Рис. 1. Распределение исследованных астероидов Главного пояса по таксономическим классам

В околоземном космическом пространстве намного меньше изученных объектов — приблизительно 3000 тел, а таксономический класс определён лишь у 1082 астероидов (рис. 2). Распределение по составу (таксономическим классам) отличается от картины, которую мы наблюдаем в Главном поясе: в околоземном пространстве преобладают каменные астероиды (S-класс) — 41 % от изученной популяции. Вторыми по распространённости являются астероиды, предположительно состоящие из оливина и пироксена с примесью металлов (Q-класс).

Распределение исследованных околоземных астероидов по таксономическим классам, 1082 астероида

ПРОЦЕНТЫ



Рис. 2. Распределение исследованных околоземных астероидов по таксономическим классам

Распределение исследованных околоземных астероидов по диаметрам и периодам, 1082 астероида

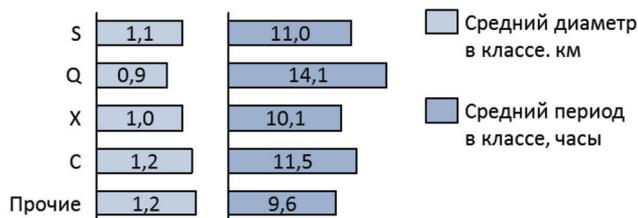


Рис. 3. Распределение исследованных околоземных астероидов по диаметрам и периодам

Эти объекты чрезвычайно редки в главном поясе, и есть причины предполагать, что это углеродистые астероиды, разорванные приливными силами планет, мимо которых они уже совершали движение [Binzel et al., 2010]. Среди исследованных околоземных астероидов металлосодержащими являются 11 % популяции (X-класс). Они представляют наибольший коммерческий интерес и содержат железо, никель, кобальт, алюминий, редкоземельные металлы и элементы платиновой группы. Углеродистые астероиды, в отличие от Главного пояса, в околоземном пространстве менее распространены, но содержат воду, азот, водород и аммиак, что делает их потенциальным источником топлива и питьевой воды для космических миссий. Остальные таксономические классы астероидов в сумме дают 9 % популяции (в среднем 1 % на класс).

При этом размеры и периоды вращения рассмотренных околоземных астероидов говорят о потенциальной технической сложности их освоения. Эти параметры определены у 504 объектов: размеры тел варьируются от 10 м до 4 км, а периоды из вращения — от нескольких минут до 15 дней (рис. 3, см. 270).

ОБЗОР МИССИЙ, ОСУЩЕСТВЛЯЕМЫХ К АСТЕРОИДАМ

Для более точного понимания химического состава астероидов необходимо исследовать вещество астероида. Такие исследования можно проводить как бесконтактными методами с помощью космического зонда, находящегося рядом с астероидом, так и контактными методами с посадкой аппарата на поверхность космического тела. Такие исследования позволяют изучить внешнее и внутреннее строение астероида, плотность, массу, гравитационное поле, а также свойства вещества астероида и реголита. [Фобос Грунт, 2011]. На данный момент было осуществлено пять полётов к астероидам и один — к комете. Целью этих миссий было никак не коммерческое освоение, а изучение астероидов, как строительного материала, оставшегося после образования Солнечной системы.

Первым искусственным спутником астероида и аппаратом, совершившим мягкую посадку на астероид, стал космический аппарат NEAR Shoemaker (NASA) массой 805 кг. Аппарат проработал около года на орбите астероида 433 Эрос, собрав большое количество информации: были уточнены параметры орбиты и движения астероида, создана трёхмерная модель астероида и определён химический состав — на поверхности были обнаружены магний, алюминий, кремний, кальций и железо. После посадки аппарат проработал на поверхности астероида около двух недель; связь была потеряна в декабре 2002 г., и миссия считалась завершённой.

Японский аппарат массой 510 кг Hayabusa в сентябре 2005 г. приблизился к астероиду (25143) Итокава. Аппарат совершил три посадки на астероид, при этом в связи со сбоями оборудования, образцы реголита удалось взять лишь один раз. Проанализировав доставленные на Землю образцы, учёные установили, что они состояли из оливина (минерал, химическая формула которого $(\text{Mg}^{2+}, \text{Fe}^{2+})_2\text{SiO}_4$) с весьма незначительным содержанием железа.

Используя опыт предыдущей миссии, в декабре 2014 г. был запущен КА Hayabusa 2 весом около 600 кг к астероиду (162173) Рюгу углеродистого типа; выход на орбиту объекта состоялся в июле 2018 г., возвращение на Землю с образцами грунта планируется на 2020 г.

В сентябре 2007 г. NASA запустили автоматическую межпланетную станцию Dawn весом около 1200 кг к астероиду (4) Веста и карликовой планете (1 Ceres) Церера. Оба тела — крупнейшие объекты главного пояса, открытые в начале XIX в. В 2012 г. были собраны и проанализированы данные по составу поверхности Весты, изменению её температуры и внутреннему составу (путём отслеживания изменения орбиты аппарата), на их основании составлены подробные карты астероида. В 2015 г. аппарат был захвачен гравитационным полем карликовой планеты и приступил к исследованию Цереры, которое продолжается до сих пор. Обратное возвращение станции не планировалось.

Миссия OSIRIS-REX к углеродному астероиду (101955) Бенну, диаметр которого более 500 м, началась в сентябре 2016 г. Целью межпланетной станции массой 2110 кг (с топливом) является доставка на Землю образцов грунта; аппарат достигнет своей цели в 2020 г.

В марте 2004 г. Европейское Космическое Агентство совместно с NASA запустило миссию Rosetta к комете Чурюмова – Герасименко. В июле 2014 г. аппарат начал передавать данные о состоянии кометы, и проводил исследования до сентября 2016 г., после чего был направлен на столкновение с объектом, и столкнулся с ним на скорости 3 км/ч, завершив миссию.

В октябре 2020 г. ЕКА также планирует запуск мини-спутников в рамках миссии Asteroid Impact Mission (AIM) для изучения астероидов; также NASA планировалась миссия Comet Hopper, но запуск, запланированный на 2016 г., был отменён.

ОЦЕНКА ВОЗМОЖНОСТИ ПЕРЕЛЁТА К АСТЕРОИДУ

Среди наиболее выгодных для коммерческого освоения астероидов с оценённой сайтом asterank.com прибылью в миллиарды долларов были рассмотрены шесть астероидов (рис. 4). Ближайшие возможные даты перелёта, количество дней полёта до астероида и средняя скорость, необходимая космическому аппарату, чтобы покинуть парковочную орбиту Земли, были взяты с сайта sneos.jpl.nasa.gov.

Из приведённых данных видно, что миссия посещения астероида займёт минимум год с учётом времени пребывания на небесном теле и обратного пути на Землю. Также для реального осуществления миссии необходимо, чтобы средняя скорость ухода с парковочной орбиты была не более 5 км/с.

Name	Est. Profit (\$)	Earth Departure dV (km/s)	Outbound Flight Time (d)	Earth departure date
1989 ML	4 380 000 000,00 ₺	3,319	190	2023
Ryugu	30 070 000 000,00 ₺	3,637	100	2020
Bennu	185 000 000,00 ₺	3,928	340	2024
Anteros	1 250 000 000 000,00 ₺	3,987	350	2038
Nereus	1 390 000 000,00 ₺	4,625	150	2022
2001 CC21	29 770 000 000,00 ₺	5,295	340	2022

Рис. 4. Коммерчески выгодные астероиды по версии сайта asterank.com

В рамках проведённой работы были рассмотрены различные варианты перелёта космического аппарата к перспективному в ресурсном плане астероиду 163243 (2002 FB3) [Harris, Drube, 2014] с Земли. Астероид 163243 (2002 FB3) спектрального класса S/Q представляет собой астероид, сближающийся с орбитой Земли, и принадлежит к группе Аполлонов. Диаметр астероида составляет 1,682 км, наклонение орбиты $20,3^\circ$, эксцентриситет 0,6. Скорость движения астероида на момент сближения с орбитой Земли составляет 24 км/с.

Для оценки реализуемости миссии к астероиду был проведён анализ прямого перелёта от Земли к астероиду. Прямой перелёт за счёт топлива КА невозможен вследствие слишком большого (более 10 км/с) необходимого импульса. По примерам предыдущих осуществлённых миссий можно сказать, что оптимальным перелётом с точки зрения наименьшего значения необходимого характеристического импульса dV является перелёт с использованием гравитационного манёвра вокруг Земли. Однако для воплощения коммерческих идей необходимо также учитывать время, необходимое для осуществления миссии. Таким образом, принимая во внимание длительность гравитационного манёвра приблизительно в один год, можно сделать вывод, что астероид 163243 (2002 FB3) и подобные ему астероиды, сближающиеся с орбитой Земли на высоких скоростях, не подходят.

Возможными кандидатами для осуществления коммерческих миссий по добыче ископаемых могут стать астероиды, сближающиеся непосредственно с Землёй и движущиеся на малых скоростях в плоскостях, близких к плоскости орбиты Земли. Для примера был рассмотрен подобный околоземный астероид 2015 DP155 из группы Аполлонов. Скорость астероида в момент сближения с Землёй составляет 5 км/с, наклонение орбиты $5,38^\circ$, эксцентриситет орбиты 0,6. Астероид сближается с Землёй на расстояние 0,02 а.е. В этом случае возможна схема гомановского перелёта между орбитами, лежащими в одной плоскости. Предполагается, что космический аппарат достигает астероид-цели и остаётся на его поверхности, поэтому изменять наклонение плоскости орбиты КА не нужно.

Таким образом, можно сделать вывод, что для реализации коммерческих миссий по добыче полезных ископаемых с околоземных астероидов в качестве астероида-цели стоит выбирать околоземные астероиды, движущиеся на малых скоростях и сближающиеся с Землёй. Эти факторы делают астероиды удобными для достижения как с точки зрения экономии топлива, так и для сокращения общего времени миссии.

ЛИТЕРАТУРА

- [Фобос-Грунт, 2011] Фобос-Грунт: Проект космической экспедиции. Кн. 2. М.: ИКИ РАН, 2011. 261 с.
- [Шустов, 2018] *Шустов Б. М.* Метеориты как свидетели и продукты эволюции Солнечной системы, угроза и ресурс на будущее // Физика Космоса: тр. 47-й Международной студенческой научной конф. Екатеринбург, 29 янв. – 2 февр. 2018. Екатеринбург : Изд-во Урал. ун-та, 2018. С. 135–151.
- [Binzel et al., 2010] *Binzel R. P., Morbidelli A., Merouane S.* Earth encounters as the origin of fresh surfaces on near-Earth asteroids // Nature. 2010. V. 463. P. 331–334.
- [Bus, Binzel, 2002] *Bus S. J., Binzel R. P.* Phase II of the Small Main-belt Asteroid Spectroscopic Survey — A feature-based taxonomy // Icarus. 2002. V. 158. Iss. 1. P. 146–177.

[Gradie et al., 1989] *Gradie J., Chapman C., Tedesco E.* Distribution of taxonomic classes and the compositional structure of the asteroid belt. Asteroid taxonomic classifications. // Asteroids II. Tucson: University of Arizona Press, 1989. P. 316–335.

[Harris, Drube, 2014] Harris A. W., Drube L. How to Find Metal-rich Asteroids // The Astrophysical Journal Letters, Volume 785, Issue 1, article id. L4, 5 pp. (2014).

[Tholen, 1989] *Tholen D. J.* Asteroid taxonomic classifications // Asteroids II. Tucson: University of Arizona Press, 1989. P. 1139–1150.

ASTEROIDS AS A FUTURE RESOURCE

M. D. Sizova, E. Efremova

Institute of Astronomy of Russian Academy of Russian (INASAN)

This article considers the distribution of asteroids by its chemical composition, for which, according to the website asterank.com, has been defined a taxonomic class, analyzed asteroids in the Main belt and in the near-Earth space. The review of space missions to asteroids, both implemented and planned, is carried out. An estimation flight possibilities to the selected asteroid, the most valid by a commercial point of view, was made.

Keywords: near Earth objects, NEO, asteroids, mining, taxonomy, asteroids composition

Sizova Maria Dmitrievna — junior researcher, PhD student, sizova@inasan.ru

Efremova Ekaterina Viktorovna — junior researcher, PhD student, efremova@inasan.ru

055(02)2

ИКИ РАН
117997, Москва, Профсоюзная, 84/32

Подписано к печати 20.11.2018

Заказ 4176

Формат 70×100/16

Тираж 50

22,18 усл.-печ. л.

