

Вычислительный центр имени А.А. Дородницына РАН

Фонд глобальных проблем выживания человечества
имени Н.Н. Моисеева

**Никита
Николаевич
МОИСЕЕВ**

**Избранные
труды**

**Гидродинамика
и механика**

**Оптимизация,
исследование операций
и теория управления**

Москва
«Тайдекс Ко»
2003

ББК 22.1
М74

Ответственный редактор академик РАН А.А. Петров

Издательство благодарит Н.К. Бурову, А.А. Муромского
и Т.В. Сидорову за помощь в подготовке к печати трудов Н.Н. Моисеева

Издание осуществлено при финансовой поддержке
Российского фонда фундаментальных исследований
по проекту № 03-01-14068

Моисеев Н.Н.

М74 Избранные труды в 2-х томах. Т. 1. Гидродинамика и
механика. Оптимизация, исследование операций и теория
управления. — М.: Тайдекс Ко, 2003. — 376 с.
ISBN 5-94702-016-5

Никита Николаевич Моисеев (1917–2000) — выдающийся математик и глубокий мыслитель. Его работы получили высокое научное и общественное признание. В книге собраны важнейшие из трудов Н.Н. Моисеева по гидродинамике и механике, а также исследованию операций, оптимизации и теории управления.

ББК 22.1

Воронина А.В., 2003
Фонд им. Н.Н. Моисеева, 2003
Предисловие: А.А. Петров, 2003
Составление: Ю.Г. Евтушенко,
Ю.Н. Павловский, В.П. Пархоменко,
А.А. Петров, А.М. Тарко,
А.М. Тер-Крикоров, Ю.А. Флеров, 2003
Оформление: Тайдекс Ко, 2003

ISBN 5-94702-016-5

Содержание

| | |
|---|------------|
| Предисловие | 5 |
| Гидродинамика и механика | 17 |
| Движение твердого тела, имеющего полость, частично заполненную идеальной капельной жидкостью | 18 |
| О колебаниях тяжелой идеальной и несжимаемой жидкости в сосуде | 24 |
| Задача о движении твердого тела, содержащего жидкие массы, имеющие свободную поверхность | 28 |
| К теории нелинейных колебаний ограниченного объема жидкости | 76 |
| К теории колебаний упругих тел, имеющих жидкие полости | 91 |
| О неединственности возможных форм установившихся течений тяжелой жидкости при числах Фруда, близких к единице | 115 |
| Теорема существования и неединственности вихревых волн периодического типа | 124 |
| К теории волн в завихренной жидкости | 130 |
| О некоторых особенностях течения жидкости, подверженной действию сил поверхностного натяжения | 146 |
| Асимптотика быстрых вращений | 159 |
| Оптимизация, исследование операций и теория управления | 181 |
| Методы динамического программирования в теории оптимальных управлений. I. Системы, допускающие использование шкалы управлений | 182 |
| Методы динамического программирования в теории оптимальных управлений. II. Общий случай аддитивных функционалов | 196 |

| | |
|--|-----|
| Информационная теория иерархических систем | 214 |
| Иерархические структуры и теория игр. | 267 |
| Имитационные модели | 289 |
| К теории оптимального управления на бесконечном интервале времени | 311 |
| Оптимизация и управление (эволюция идей и перспективы) | 324 |
| Кибернетическое описание эколого-экономических систем | 346 |

Предисловие

XX век был веком канунов и предзнаменований.

В начале века классическая физика была накануне релятивистской и квантовой революции. Концепция макроскопического детерминизма уступила место концепции микроскопической неопределенности, которая качественно изменила рационалистическое мировоззрение. Идею предопределенности гармонии и симметрии в природе и духе сменила идея исходного хаоса, из которого нелинейные взаимодействия выстраивают квазипорядок.

В середине века человечество было на пороге признания того, что овладело силами природы на собственную погибель. Идею неисчерпаемости ресурсов Земли сменила идея ответственности прогресса за судьбу человечества. В конце века настал канун осознания, что технологические нововведения могут погубить социальную природу людей. Способность мыслить и сопереживать выделила людей из живой природы. Информационные технологии позволяют подменить мышление потреблением информации, предложенной по самым высоким стандартам массового производства. Потребителями информации можно манипулировать точно так же, как потребителями товаров, оставляя им ощущение полной свободы выбора и лишая их чувства социальной ответственности.

Две мировые войны были предзнаменованием того, как идеи, овладевшие массами, запускают механизмы взаимного уничтожения в глобальных масштабах. В конце XX века экономический разрыв между странами мира перерос в технологический разрыв, который достиг масштаба цивилизационного разлома. Разлом чреват напряжениями глобального масштаба, которые разрешаются мировыми конфликтами.

Наука активно влияла на перемены XX века, а перемены отразились на отношении общества к науке. Однако редко подобного рода перемены так ярко отражаются в индивидуальной судьбе

ученого, как отразились они в жизни и научном творчестве Никиты Николаевича Моисеева. Поэтому читателям будет интересно проследить эволюцию научных пристрастий, подходов к задачам, взглядов замечательного российского ученого. Мы надеемся, что предлагаемое собрание избранных трудов Н.Н. Моисеева дает достаточно материала для этого. Отбирая работы, мы старались показать, как из конкретного, содержательного научного опыта рождалось мировоззрение и философские взгляды ученого.

* * *

Никита Николаевич Моисеев родился 23 августа 1917 года в Москве. Его отец, Николай Сергеевич, окончил юридический факультет Московского университета и был оставлен при университете «для подготовки к профессорскому званию». Однако в 1918 г. он был уволен из университета, и его научная карьера не состоялась, во многом по вине тогдашнего наркома просвещения А.В. Луначарского. Мать, Елена Николаевна, была приемной дочерью Николая Карловича фон Мекк, сына Надежды Филаретовны фон Мекк, известной по той роли, которую она сыграла в жизни нашего великого композитора П.И. Чайковского. Мать Никиты Николаевича рано умерла, и большую часть жизни он прожил с мачехой, Маргаритой Васильевной, о которой заботился до последних дней ее жизни.

Детство Никита Николаевич провел на Сходне, под Москвой, в доме деда Сергея Васильевича Моисеева, который был инженером-путейцем, дослужился до генеральского чина, и при советской власти занимал высокий пост в наркомате железных дорог. Это была жизнь русской интеллигентной семьи, в которой интересовались многим, в том числе политикой, литературой, говорили о судьбах России. Не принимая большевизм, они надеялись на возрождение России после нэпа. Суждения деда и отца, по признанию Никиты Николаевича, во многом определили его взгляды на всю жизнь. «Дед и отец — они были искренними русскими патриотами в самом цивилизованном понимании этого слова», — вспоминал он.

Однако при вступлении в самостоятельную жизнь социальное происхождение сделало Никиту Николаевича изгоем. Его не приняли в Московский университет, хотя вступительные экзамены были сданы успешно. Только вмешательство И.М. Гельфанда и счастливый случай помогли ему стать студентом МГУ.

На старших курсах Н.Н. Моисеев специализировался по теории функций действительного переменного, но не успел окончить университет, как началась Великая Отечественная война. Он был направлен учиться в Военно-воздушную академию им. Н.Е. Жуковского, окончил ее в мае 1942 г. и в звании лейтенанта уехал на Волховский фронт старшим техником по вооружению самолетов. Закончил войну боевым офицером, награжденным орденами и медалями.

Научную деятельность Никита Николаевич начал в авиационном полку. Случайно ему попала немецкая книга по внешней баллистике ракетных снарядов. Прочитав ее, Моисеев существенно упростил изложенный в книге метод, написал 10-страничный трактат и отправил профессору Ю.А. Победоносцеву, у которого учился в академии Жуковского в 1942 г. Ответом был приказ, которым капитан Моисеев откомандировывался на кафедру реактивного вооружения самолетов академии Жуковского.

Кафедру возглавлял профессор Дмитрий Александрович Вентцель. Его Никита Николаевич считал одним из своих учителей: «Благодаря общению с Дмитрием Александровичем я понемногу начал понимать прелесть прикладной науки и задач, возникающих в инженерной практике, которые требуют и остроумия и изобретательности не меньше, чем любые высокие материи. И постепенно осознал, что наука едина, если она действительно НАУКА. Нет наук первого и второго сорта. Они делятся по совсем другим принципам: есть настоящая, глубокая наука и есть спекуляции на науке. Другого разделения с тех пор я не признаю».

Работая на кафедре, Н.Н. Моисеев предложил простой оригинальный метод расчета рассеивания неуправляемых ракетных снарядов, и эти результаты стали основой его кандидатской диссертации по техническим наукам. За два-три года Н.Н. Моисеев стал известным специалистом в области динамики управляемых ракет и в 1949 г. ушел из армии. Он был принят на работу в ведущий ракетный НИИ и на кафедру МВТУ им. Н.Э. Баумана, выступал с докладами, публиковался, и уже писал книгу, которая должна была стать основой докторской диссертации. Но в 1949 г. его мачеху М.В. Моисееву, учительницу младших классов школы поселка Сходня, арестовали по обвинению «в участии в подготовке вооруженного восстания» и осудили на 10 лет лагерей. Н.Н. Моисеев был лишен допуска к секретным материалам и уволен с работы.

Результаты этого периода его научной деятельности содержались в закрытых отчетах и не были опубликованы.

В 1950 г. Н.Н. Моисеев переезжает в Ростов-на-Дону и, можно сказать, заново начинает научную жизнь в Ростовском университете доцентом на кафедре теоретической механики. Он начинает заниматься гидродинамикой и получает результаты, которые стали содержанием его докторской диссертации по физико-математическим наукам. В 1955 г. Никита Николаевич успешно защитил диссертацию на Ученом совете Математического института им. В.А. Стеклова АН СССР.

В том же 1955 г. Н.Н. Моисеев получает предложение академика М.А. Лаврентьева занять должность профессора кафедры теории взрыва Московского физико-технического института и возвращается в Москву. С 1956 г. в течение тридцати лет он был связан с Московским физико-техническим институтом и Вычислительным центром АН СССР, куда в том же году пришел старшим научным сотрудником в лабораторию А.А. Абрамова. В МФТИ Н.Н. Моисеев был одним из инициаторов создания в 1969 г. факультета управления и прикладной математики и стал его первым деканом. А в ВЦ АН СССР существовало более пятнадцати отделов, созданных по инициативе и при участии Н.Н. Моисеева, когда в 1987 г. Никита Николаевич вышел в отставку с поста заместителя директора ВЦ АН СССР по науке.

* * *

В научном наследии Н.Н. Моисеева легко проследить сочетание высокой математической культуры и острой прикладной направленности. Уже в первых работах, которые принесли ему научный авторитет и известность, отразились главные особенности научного творчества Н.Н. Моисеева: полный математический анализ задачи и интерес к интерпретации математических фактов, способность увидеть те физические и технические следствия, которые эти факты могли иметь. Вокруг Никиты Николаевича собралась большая группа способных молодых исследователей — студентов и аспирантов МФТИ и МГУ им. М.В. Ломоносова. У него было замечательное чутье на новые актуальные задачи. Он видел связь их с теми проблемами, которые были жизненно важны для страны. Особенно ярко это проявилось тогда, когда он искал общие и четкие постановки задач, от решения которых зависит научно-технический и социальный

прогресс нашей страны или судьба планеты в целом. Он разрабатывал методологию междисциплинарных исследований как подход к решению этих задач, искал в методологии место прикладной математики и компьютерных технологий.

Результаты деятельности Н.Н. Моисеева получили высокое научное и общественное признание. В 1966 г. он был избран членом-корреспондентом АН СССР, а в 1984 г. — действительным членом (академиком) АН СССР. Н.Н. Моисеев был награжден многими орденами и медалями, ему была присуждена Государственная премия СССР и премия Совета Министров СССР, был избран членом иностранных академий и научных сообществ, удостоен международных премий.

В научной жизни Н.Н. Моисеева можно выделить несколько периодов, которые определялись его научными интересами. В соответствии с этим предлагаемые читателю «Избранные труды» Н.Н. Моисеева разделены на четыре части.

В 1950-х и начале 1960-х годов исследования Н.Н. Моисеева были посвящены задачам гидродинамики и механики. Работая на кафедре университета в Ростове, он подошел к задачам о колебаниях жидкости в ограниченной полости академически, изучая со студентами классический труд Н.Е. Жуковского о движении твердого тела, полость которого заполнена жидкостью. В это же время Н.Н. Моисеев как математик исследовал спектр собственных колебаний тяжелой идеальной жидкости в неподвижном сосуде. Естественным образом он пришел к обобщению задачи Н.Е. Жуковского — к постановке задачи о движении твердого тела, полость которого не полностью заполнена тяжелой идеальной жидкостью. В линейной постановке система уравнений движения тела с жидкостью была выписана в общей операторной форме. Н.Н. Моисеев выяснил, что линейный оператор задачи расщепляется на бесконечномерный положительно определенный оператор и конечномерный оператор, который не обладает определенными свойствами. Далее он действовал как прикладник, прошедший школу Д.А. Вентцеля. Он усмотрел, что тело с колеблющейся жидкостью можно описать как механическую систему с бесконечным числом степеней свободы. Задача гидродинамики была отделена от задачи динамики эквивалентного тела, с которым связана бесконечная система математических маятников. Динамические характеристики жидкости с «замороженной» свободной поверхностью вычислялись по методу

Н.Е. Жуковского, а колебания свободной поверхности вычислялись через решение задачи о собственных колебаниях жидкости в неподвижной полости. Н.Н. Моисеев доказал полноту главных колебаний системы и обосновал вариационный метод для их вычислений. Это было исчерпывающим решением задачи. Вариационный метод давал возможность эффективно использовать ЭВМ для вычислений динамических характеристик механической системы, эквивалентной телу, содержащему жидкость в полостях сложной формы. В дальнейшем Н.Н. Моисеев распространил эти результаты на задачу о движении упругого тела с полостью, содержащей колеблющуюся идеальную жидкость.

Из теории, построенной Н.Н. Моисеевым, следовало, что для устойчивости равновесия твердого тела, в полости которого может колебаться жидкость, необходима и достаточна устойчивость эквивалентного твердого тела. Это был классический результат, обобщающий известную классическую теорему Н.Е. Жуковского. Из теории следовали и важные практические результаты. На активном участке ракета с жидким топливом может быть описана как твердое (или, точнее, упругое) тело, в полостях которого колеблется жидкость. Уравнения движения такой системы необходимо знать, чтобы стабилизировать ее движение. Так, начав с академической постановки задачи и выполнив полное ее математическое исследование, Н.Н. Моисеев получил прикладные результаты, которые были актуальны для бурно развивавшейся в то время ракетной техники. Методы расчета динамических характеристик тела с жидкостью, которые он предложил, были специально ориентированы на применение ЭВМ в высшей степени актуальной тогда области приложений.

Это — характерная особенность всех работ Никиты Николаевича, которые он выполнил вместе со своими учениками в Вычислительном центре АН СССР. Внимательно следя за развитием вычислительной техники, он не уставал повторять: «Мы должны быть готовы использовать все возможности ЭВМ нового поколения для решения задач народнохозяйственного и государственного значения».

Итогом исследований движения тел с жидкостью стала монография «Динамика тела с полостями, содержащими жидкость», написанная вместе с В.В. Румянцевым и опубликованная в 1965 г.

Задачу о колебаниях тяжелой идеальной жидкости в сосуде Н.Н. Моисеев рассмотрел и в нелинейной постановке, применив

подход, предложенный А. Пуанкаре. Он искал приближенное решение задачи в виде асимптотического ряда по малому параметру (амплитуде волны) и показал, что спектр собственных колебаний является не дискретным, а кусочно-непрерывным, а амплитуда колебаний может быть любой из круга сходимости ряда. Рассмотрев задачу о вынужденных колебаниях жидкости в периодическом поле массовых сил, построил вблизи резонанса и вдали от резонанса приближенное асимптотическое решение, вырождающееся в тривиальное решение.

Основные результаты Н.Н. Моисеева в теории нелинейных волн относятся к ее классическим разделам. При общих предположениях им доказано существование волн на поверхности завихренной жидкости. Он развил метод М.А. Лаврентьева конформных отображений узких полос и применил его для асимптотического анализа целого ряда задач о нелинейных волнах на поверхности жидкости.

Исследования задач о нелинейных колебаниях жидкости в начале 1960-х годов подвели Никиту Николаевича к общей проблеме нелинейных колебаний механических систем. Им были разработаны асимптотические методы, основанные на процедурах усреднения и разделения движений. Асимптотическая обработка задач о движении космических аппаратов дала возможность построить экономные численные методы интегрирования уравнений движения.

Исследования нелинейных колебаний завершились в 1969 г. монографией «Асимптотические методы нелинейной механики».

Начало следующего этапа научной жизни Никиты Николаевича отмечено резкой сменой научных интересов. В начале 1960-х годов он еще продолжал заниматься гидродинамикой и механикой, когда всерьез заинтересовался теорией оптимального управления механическим движением, а затем общей теорией управления сложными системами. В то время на первый план вышли задачи не анализа, а синтеза систем. Первыми из такого рода содержательных задач были задачи оптимального управления движением космических и летательных аппаратов. Н.Н. Моисеев сразу понял актуальность этих задач и оценил перспективы развития этого научного направления. Сначала вместе со своими учениками он осваивал классический принцип максимума Понтрягина и искал эффективные численные методы решения краевых задач, к которым сводился расчет оптимальных траекторий. Затем возникла задача об оптимальной

траектории космического аппарата в облет радиационных поясов Земли. Формулировка задачи содержала ограничения на фазовые переменные аппарата, поэтому к ней нельзя было применить классический принцип максимума Понтрягина, и Никита Николаевич предложил совершенно оригинальный подход к решению таких задач, который он назвал методом перебора в пространстве состояний. По существу он открыл новое направление в теории разностных методов. Метод, предложенный Н.Н. Моисеевым, оказался универсальным для широкого класса задач механики и дал начало новому направлению исследований.

Никита Николаевич Моисеев сделал очень много для развития методов оптимизации и теории управления в нашей стране. С 1965 г. он начал проводить Всесоюзные летние школы по методам оптимизации и теории управления. Они собирались раз в два года на Украине, в Азербайджане, Молдавии, Белоруссии, Эстонии, на Волге, на Урале, в Сибири, и так продолжалось до 1989 г. Школы способствовали росту научной молодежи практически во всех регионах Советского Союза, очень многие участники школ Моисеева выросли в ведущих специалистов, известных во всем мире. С тех пор во всех республиках бывшего Советского Союза сохраняется неформальное научное и духовное единство, которое называется Школой Моисеева.

Собственные исследования задач оптимизации Никита Николаевич завершил монографией «Численные методы в теории оптимальных систем», опубликованной в 1971 г.

Обстановка начала 1960-х годов будоражила умы. Наконец получили общественное признание идеи и методы оптимального планирования народного хозяйства, еще до войны предложенные академиком Л.В. Канторовичем. Отдавая им должное, в то же время Никита Николаевич, пожалуй, первым преодолел иллюзии, связанные с оптимизационным подходом к управлению народным хозяйством. Дело в том, что оптимизационные модели ограничивались только процедурами планирования хозяйства, в них не учитывались реальные механизмы регулирования производства и обращения. Каждый из субъектов экономики имеет собственные интересы и даже в условиях самой жесткой централизации каждый обладает ресурсом (например, дополнительной информацией), чтобы действовать в соответствии со своими интересами. Никита Николаевич очень быстро это почувствовал. В 1967 г. он докладывал о постановке задачи планирования, основанной на

программно-целевом методе. Предлагал, определяя цели развития и способы их достижения, учитывать ограничения, возникающие из-за разнообразия интересов людей, и только после того как определены способы действий, оптимизировать распределение ограниченных ресурсов между работами, из которых состоят способы достижения целей, используя для этого ресурсы ЭВМ.

Так Н.Н. Моисеевым впервые была выдвинута идея сочетания неформальных процедур и математических методов в проблеме управления сложными организационными, технико-экономическими и социально-экономическими системами. Потом она перерастет в общую идею синтеза гуманитарного и естественнонаучного знания. А в начале 1970-х он предложил теорию иерархических систем и метод имитационного моделирования в качестве инструментов изучения сложных систем, в которых люди одновременно и субъекты и объекты управления.

Рассматривая множественность интересов субъектов в условиях организационной и социальной иерархии и неизбежную неполную информированность субъектов, Н.Н. Моисеев построил информационную теорию иерархических систем. Она основана на модели, которая описывает нетождественность целей системы и отдельных ее элементов и принятие решений каждым из них в условиях неполной информированности. Используя теоретико-игровые конструкции, Н.Н. Моисеев предложил подход к численному анализу и синтезу сложных иерархических систем управления.

Метод имитационного моделирования опирался на новые возможности, которые предоставляли ЭВМ третьего поколения. Имитационное моделирование сочетало в себе математическое описание тех процессов в сложной системе, которые поддаются формализации, и неформальные процедуры имитации операторами тех процедур принятия решений, которые формализации не поддаются. Инструментом имитационного моделирования была диалоговая человеко-машинная система с программным обеспечением проведения имитационных экспериментов и обработки их результатов. Метод имитационного моделирования усовершенствовал технологию проектирования сложных систем, заменяя натурные эксперименты с системой имитационными экспериментами с ее полной моделью на ЭВМ.

Исследования проблем управления сложными системами завершились монографией «Математические задачи системного анализа», опубликованной в 1981 г.

Н.Н. Моисеев активно пропагандировал свой подход к проблеме управления народным хозяйством, настойчиво, хотя и безуспешно, пытался внедрить свои идеи в сознание высшего государственного руководства. «...Все было тщетно, — вспоминал Никита Николаевич, — по-иному и быть не могло... Наша государственная, политическая и экономическая система была уникальным созданием Природы. Именно Природы! Ее никто не создавал по какому-то задуманному плану. Она возникла в результате внутренних причин развития организации, тех изначальных стимулов, которые в нее были заложены еще в процессе революции». Однако прежде чем Никита Николаевич пришел к целостной концепции развития сложных систем, прошло несколько лет интенсивной работы над глобальной моделью биосферы. Это был последний период научной жизни Никиты Николаевича в Вычислительном центре АН СССР.

О кибернетическом описании эколого-экономических систем Н.Н. Моисеев писал еще в 1977 г., применяя свою информационную теорию иерархических систем. Проблемами математического моделирования процессов в биосфере он начал интересоваться в конце 1970-х и обсуждал их с Н.В. Тимофеевым-Ресовским. В процессе работы над моделью Н.Н. Моисеев воспринял учение В.И. Вернадского о ноосфере и развил концепцию коэволюции человека и природы. К тому времени уже были широко известны работы Дж. Форрестера и Д. Медоуза по глобальной динамике. Отправной точкой исследований Никиты Николаевича была критика этих работ. Он разработал программу, целью которой было создание системы моделей, описывающих взаимодействия процессов в биосфере и человеческой активности. На первом этапе была построена модель взаимодействия процессов в атмосфере и в океане, влияющих на погоду и климат. Никита Николаевич работал над этой моделью вместе со своим учеником В.В. Александровым, который проводил численное исследование модели и вычислительные эксперименты с ней.

Когда американский астроном К. Саган опубликовал сценарий возможных последствий ядерной войны между США и СССР, Н.Н. Моисеев уже имел модель, на которой можно было проверить сценарий Сагана и дать количественные оценки состояния атмосферы и океана. Расчеты, проведенные Александровым, обнаружили феномен «ядерной зимы». Эта работа принесла Никите Николаевичу мировую известность, но намеченная программа осталась не выполненной: трагически пропал В.В. Александров, началась перестройка,

и финансирование научных исследований сокращалось. Позже Никита Николаевич написал: «Само по себе исследование феномена «ядерной зимы» было более чем второстепенным событием в той большой работе, которую я задумал... Но именно история «ядерной зимы», которая сначала меня особенно не интересовала, получила широкую известность и сделала большую рекламу всему направлению, которое я начал развивать в Вычислительном центре Академии наук СССР. В то же время научные результаты, которые мне представлялись наиболее интересными, так же как и общее понимание смысла проблемы «человек — биосфера» или особенностей самоорганизации материального мира, остались просто незамеченными, а вероятнее всего, и непонятыми».

Результаты исследований глобальных экологических проблем были изложены в монографии «Человек и биосфера», написанной в 1985 г. вместе с В.В. Александровым и А.М. Тарко.

В 1987 г. Никита Николаевич Моисеев ушел в отставку с поста заместителя директора по науке Вычислительного центра АН СССР. Последний период его жизни был связан с экологической организацией Российский Зеленый Крест, Международным эколого-политологическим независимым университетом, Российским комитетом UNEP, президентом которых он был до конца своей жизни. Последние годы его жизни были наполнены интенсивной интеллектуальной и духовной работой, осмысливанием философских проблем естествознания, проблем синтеза естественных и гуманитарных наук, страстной публицистикой, темой которой были судьба России и ответственность интеллигенции.

Исследования проблем глобальной экологии, стабильности биосферы и антропогенных воздействий на процессы в ней привели Н.Н. Моисеева к критическому осмыслению представлений о значении природы в развитии общества и о роли человечества в процессах планетарного масштаба. Развивая идеи А. Пуанкаре, Н. Бора, В.И. Вернадского, он создал оригинальную концепцию, названную универсальным рационализмом. В рамках этой концепции Моисеев пришел к новому пониманию принципов коэволюции человека и природы и понял необходимость новых нравственных императивов как условий сохранения человека как биологического вида на планете.

Свою концепцию Никита Николаевич строил как математик и как философ. Он построил и исследовал математическую модель регуляторных свойств биоты и в то же время настойчиво указывал на

необходимость синтетического подхода естественных и гуманитарных наук к проблемам природы и человечества. В концепции универсального рационализма он предложил схему мирового процесса самоорганизации, которая сочетала в себе подходы естествознания и философии. Концепция покоилась на эмпирических обобщениях и переварила в себе фундаментальные достижения науки XX века. Основные гипотезы о связанности всех элементов Вселенной, о множественности представлений сложной системы, о случайности и неопределенности, которые свойственны явлениям мира, об отборе и т. д. являются философскими осмыслениями фундаментальных положений современных наук о неживой и живой природе. Основываясь на исходных гипотезах, Н.Н. Моисеев представил общую картину коэволюции природы и общества. Применяя общие положения универсального эволюционизма к анализу современного антропогенеза и цивилизационных разломов, он не только предупреждал о катастрофах, грозящих техногенной цивилизации, но выявлял их эволюционные и экологические основания, намечал пути общественного развития, на которых их можно было бы избежать.

* * *

Никита Николаевич Моисеев был патриотом России в самом истинном, высоком смысле. С болью переживал он смутное время, наставшее в России на рубеже 1990-х годов. Но не только. Он активно пытался противостоять пошлости недоучек, объявивших себя элитой нашего общества. Это нашло отражение в его публицистике, многочисленных публичных выступлениях. Он писал и говорил о единстве естественнонаучного и гуманитарного знания, о формировании мировоззрения, о традициях российской науки и роли Учителя, о русской интеллигенции и о месте России в мире.

Никита Николаевич Моисеев прожил долгую, яркую и нелегкую жизнь. В конце ее сверкнула еще одна грань его таланта: в 1994 г. он написал книгу «Как далеко до завтрашнего дня... Свободные размышления. 1917-1993». Эта книга о его судьбе и его работе, которые так тесно переплелись с судьбой его страны.

Он был романтиком и оптимистом, несмотря ни на что верил, что слово, обращенное к людям, найдет отклик и даст плоды.

А.А. Петров,
академик РАН

Гидродинамика и механика

1952–1962

Движение твердого тела, имеющего полость, частично заполненную идеальной каплевой жидкостью

В своем известном сочинении [1] Н.Е. Жуковский рассмотрел задачу о движении твердого тела, имеющего полости, полностью заполненные каплевой жидкостью. Возможно обобщить некоторые из его результатов на тот случай, когда жидкость имеет свободную поверхность.

1. Для простоты полагаем, что твердое тело (сосуд) имеет одну односвязную полость произвольной формы. Движения системы сосуд — жидкость около положения равновесия считаем малыми. Предполагаем, что вектор напряженности массовых сил постоянен по направлению, а его величина является функцией только времени. Тогда свободная поверхность жидкости в положении равновесия представляет собой плоскость. Если движение жидкости в начальный момент потенциальное, то оно им будет и в любой момент времени. Внешние силы, действующие на сосуд, считаем такими, что волны, возникающие в жидкости, обладают малыми амплитудами.

Свяжем с твердым телом подвижную систему координат так, что невозмущенная поверхность жидкости совпадает с плоскостью O_1 . Далее предполагаем, что из множества функций, гармонических внутри полости и удовлетворяющих условию

$$\frac{\partial^2 \psi}{\partial n^2} = 0$$

на поверхности той части полости, которая в положении равновесия занята жидкостью, можно выбрать полное семейство функций ψ_i , ортогональных на плоской фигуре, являющейся сечением полости невозмущенной поверхностью жидкости.

Введем следующие обозначения: M — масса системы; \mathbf{v}_0 — скорость точки O , связанной с сосудом и совпадающей с центром

Для опубликования в Докладах Академии наук СССР представлено академиком М.В. Келдышем.

тяжести системы, в положении равновесия; \mathbf{v} — скорость движения жидкости относительно подвижной системы координат, движущейся поступательно вместе с точкой O ; \mathbf{Q} — главный вектор внешних сил; V_0 — часть полости, занятая жидкостью в положении равновесия; $F_0(x, y, z) = 1$ — эллипсоид инерции сосуда в точке O ; $F_1(x, y, z) = 1$ — эллипсоид инерции эквивалентного (в смысле Жуковского) твердого тела для сосуда, имеющего форму объема V ,

$$F_1 = \sum_{i,j=1}^3 \dots_{i,j} \dots_{i,j},$$

где $\omega_1, \omega_2, \omega_3$ — проекции мгновенной угловой скорости на подвижные оси x, y, z , имеющие начало в точке O ;

$$uv = (u \cdot v)d;$$

$\omega_1, \omega_2, \omega_3$ — гармонические функции, удовлетворяющие граничным условиям: на границах жидкости в положении равновесия (границах объема V)

$$\frac{1}{n} \frac{\partial u}{\partial z}, \frac{2}{n} \frac{\partial z}{\partial x}, \frac{3}{n} \frac{\partial x}{\partial y};$$

$\cos \alpha, \cos \beta, \cos \gamma$ — направляющие косинусы внутренней нормали \mathbf{n}° ; \mathbf{L} — главный момент внешних сил; $\mathbf{g}(t)$ — вектор напряженности массовых сил; g_x, g_y, g_z — его проекции на оси O_1 и O_2 ; F_1, F_2 — инерциальные силы, равные по величине и обратные по знаку проекциям соответствующих ускорений подвижного начала отсчета O на оси O_1 и O_2 ; \mathbf{r} — радиус-вектор относительно подвижного центра O .

В принятых предположениях справедлива следующая теорема:

Теорема. Движение системы сосуд — жидкость описывается бесконечной системой линейных дифференциальных уравнений:

1) уравнение количества движения:

$$M \frac{d\mathbf{v}_0}{dt} = \sum_{i=1}^3 \frac{df_i}{dt} \mathbf{e}_i + \mathbf{Q}; \tag{1}$$

2) уравнение моментов:

$$\frac{d\mathbf{G}}{dt} = \sum_{i=1}^3 \frac{df_i}{dt} (\mathbf{r}_i \times \mathbf{e}_i) + \mathbf{L}, \tag{2}$$

где

$$\mathbf{G} = \frac{1}{2}\Phi(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{2}F_0(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{2}F_1(\alpha_1, \alpha_2, \alpha_3);$$

3) уравнения для коэффициентов разложения в ряд Фурье потенциала скоростей:

$$\frac{d}{dt} \frac{1}{g(t)} \frac{df_i}{dt} = \sum_{k=1}^3 a_{ik} f_k - \frac{d(F \cdot g)}{dt} \frac{b_i (F \cdot g) c_i}{g(t)}$$

$$\frac{d}{dt} \frac{d_{i_1} \quad d_{i_2} \quad d_{i_3}}{g(t)} = e_{i_1} \quad e_{i_2} \quad e_{i_3}, \quad (3)$$

где числа $a_{ik}, b_i, c_i, d_{i_k}, e_{i_k}$ зависят только от формы полости и выбора функций f_i .

Если вектор $\mathbf{g}(t)$ не зависит от времени, то система (1), (2) (3) — система уравнений с постоянными коэффициентами.

Уравнение (2) может быть приведено к виду

$$\frac{d\mathbf{G}}{dy} = L, \quad (4)$$

где

$$\mathbf{G} = \frac{1}{2}\Phi(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{2}\tilde{\Phi}(\alpha_1, \alpha_2, \alpha_3),$$

$$\tilde{\Phi} = 2 \sum_{i=1}^3 \sum_{j=1}^3 f_i \alpha_i \alpha_j.$$

Уравнение (4) будем называть обобщенным уравнением Жуковского.

2. В качестве примера рассмотрим плоские колебания открытого сосуда прямоугольной формы под действием упругой силы — $Mk^2 X$, где X — рабочее поджатие пружины.

В этом случае система уравнений имеет вид:

$$X_{n=0}^{2n-1} f_{2n-1} k^2 X = 0, \quad (5)$$

$$f_{2n-1}^2 = \frac{2}{2n-1} f_{2n-1} X b_{2n-1} = 0,$$

где

$$b_{2n-1} = 2 \frac{(-1)^n \operatorname{th} \frac{2n-1}{2l} h}{(2n+1)^2 h \left(1 + \frac{m_1}{m_2}\right)},$$

$$\frac{2}{2n-1} = \frac{2n-1}{2l} g \operatorname{th} \frac{2n-1}{2l} h, \quad (6)$$

$$b_{2n-1} = (-1)^n \frac{8l}{(2n-1)^2},$$

h — глубина жидкости, $2l$ — ширина сосуда, m_1 — масса сосуда, m_2 — масса жидкости.

Уравнение частот имеет вид:

$$k^2 = 4 \frac{b_{2n-1}}{2} \frac{2n-1}{2n-1}. \quad (7)$$

Распределение корней легко исследовать графически (рис. 1).

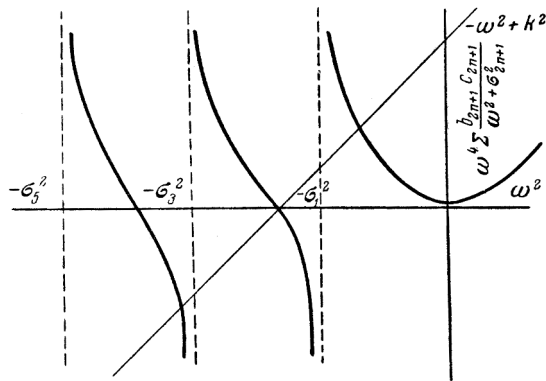


Рис. 1

Ряд b_{2n-1}^{2n-1} сходится, и если его сумма меньше или равна единице, то все корни уравнения чисто мнимые. Это, однако, всегда имеет место. В самом деле, оценим величину

$$D = \frac{b_{2n-1}^{2n-1}}{1 - \frac{2}{2n-1} / 2^2}.$$

Так как

$$D = \frac{\text{th} \frac{2n-1}{2l} h}{h(2n+1)^2} = \frac{16}{3^3},$$

то

$$D = \frac{8}{2} = \frac{1}{(2n+1)^2} < 1.$$

Детальному исследованию этого частного примера посвящена статья Л.Н. Сретенского [2], в которой он, подсчитывая давление на стенки движущегося сосуда, приходит к аналогичному уравнению частот. Однако далее он делает вывод о возможности развивающихся колебаний. Последнее, как нами показано, неверно.

Ряд $\frac{b_{2n-1}^{2n-1}}{2^{\frac{2n-1}{2n-1}}}$, как это видно из (6), для любого $2^{\frac{2n-1}{2n-1}}$ сходится как $\frac{1}{n^4}$. Быстрая сходимость этого ряда обуславливает эффективность метода при решении прикладных задач.

Для определения собственных частот может оказаться полезной следующая теорема:

Теорема. Сумма $\sum_n \left| \frac{2}{2n-1} - \frac{2}{2n-1} \right|$ стремится к нулю с возрастанием числа n как величина $\frac{1}{n^2}$.

Через $2n-1$ мы обозначили корень уравнения частот, удовлетворяющий неравенству $\frac{2}{2n-3} < \frac{2}{2n-1} < \frac{2}{2n-1}$.

Каждой из собственных частот ω_i может быть поставлена в соответствие приведенная масса сосуда равенством

$$M_i = \frac{K}{|i|^2}.$$

В общем случае M_i есть функция приведенной возвращающей силы K .

Литература

1. Жуковский Н.Е. О движении твердого тела, имеющего полости, наполненные однородной капельной жидкостью. Т. 3, с. 21. — М., 1936.

2. Сретенский Л.Н. Колебания жидкости в подвижном сосуде./Изв. АН СССР, ОТН, № 10 (1951), 1483–1494.

1952 г.

О колебаниях тяжелой идеальной и несжимаемой жидкости в сосуде

1. Как известно, потенциал скоростей малых движений жидкости в сосуде x, y, z, t является функцией, гармонической внутри области, ограниченной стенками сосуда и плоскостью $z = 0$ свободной поверхности жидкости в положении равновесия, и удовлетворяющей граничным условиям:

на стенках:

$$\frac{\partial \phi}{\partial n} = 0; \tag{1}$$

при $z = 0$:

$$\frac{\partial \phi}{\partial z} = \frac{1}{g} \frac{\partial^2 \phi}{\partial t^2}, \tag{2}$$

где g — ускорение силы тяжести.

Свободная поверхность определяется в этом случае условием:

$$x, y, z = \frac{1}{g} \frac{\partial \phi}{\partial t}(x, y, 0, t). \tag{3}$$

2. Пусть $\phi_n(x, y, z)$ — система функций, гармонических внутри, удовлетворяющих условию (1) и образующих полную ортонормированную систему функций на плоской фигуре S , являющейся сечением полости сосуда плоскостью $z = 0$.

Тогда потенциал можно искать в виде ряда:

$$\phi = \sum_i f_i(t) \phi_n. \tag{4}$$

Обозначая через a_{nm} коэффициенты Фурье в разложении при $z = 0$ функции ϕ_n / z :

Для опубликования в Докладах Академии наук СССР представлено академиком М.В. Келдышем.

$$\frac{\partial}{\partial z} a_{nm}(x, y, 0) \quad (5)$$

и используя условия (2), получим для определения функций $f_n(t)$ следующую бесконечную систему линейных уравнений:

$$f_n(t) - g_{nm} f_m(t) = 0, \quad n, m = 1, 2, \dots \quad (6)$$

3. Пусть теперь функции $q_n(x, y)$ образуют некоторую полную ортонормированную на S систему. Положим

$$(x, y, z, t) = q_n(t) q_n(x, y). \quad (7)$$

Тогда потенциал можно выразить формулой:

$$(x, y, z, t) = \oint H(x, y, z, \xi, \eta) \frac{(\xi, \eta, t)}{n} d\xi d\eta, \quad (8)$$

где H — функция Грина задачи Неймана, а интеграл распространен на всю поверхность, ограничивающую объем V .

В силу (1) и (7) выражение (8) принимает вид:

$$(x, y, z, t) = \sum_n q_n(t) q_n(x, y),$$

где

$$q_n(x, y, z, 0) = \int_S H(x, y, z, \xi, \eta, 0) q_n(\xi, \eta) d\xi d\eta.$$

Обозначая через b_{nm} коэффициенты Фурье в разложении

$$q_n(x, y, 0) = \sum_m b_{nm} q_m(x, y), \quad (9)$$

получим из условия (3) следующую бесконечную систему уравнений для определения функций $q_n(t)$:

$$\sum_m b_{nm} q_m - g q_n = 0, \quad n, m = 1, 2, \dots \quad (10)$$

4. Собственными частотами главных колебаний будут корни характеристических уравнений (6) или (10), а для определения главных колебаний в общем случае нам надо разрешить одну из этих

систем. Возникает вопрос: каким должен быть базис $\{ \varphi_n \}$ или $\{ \psi_n \}$, чтобы система (6) или (10) распадалась на отдельные уравнения? Как нетрудно видеть, таким базисом является нормированная система собственных функций интегрального уравнения:

$$\int_S \varphi_n(x, y) H(x, y, 0, \dots, 0) \varphi_n(x, y) dx dy = b_{nn}, \quad (11)$$

так как в этом случае $b_{nm} = \delta_{nm} \frac{1}{b_{nn}}$, где δ_{nm} — символ Кронекера, и система (10) приводится к виду

$$g_n q_n - g_n q_n = 0. \quad (12)$$

Аналогично показывается, что $a_{nm} = \delta_{nm} \frac{1}{a_{nn}}$ т. е. функции $f_n(t)$ удовлетворяют той же системе дифференциальных уравнений (12).

5. Таким образом, мы приходим к следующей теореме:

Теорема. Квадрат собственных частот свободных колебаний жидкости в сосуде равен произведению

$$g_n \varphi_n,$$

где φ_n — собственные числа интегрального уравнения (11), а форма свободной поверхности, соответствующая главным колебаниям, определяется формулой

$$\varphi_n = A_n \sin \sqrt{g_n} t \varphi_n(x, y),$$

где A_n и φ_n — произвольные постоянные, а φ_n — собственные функции того же интегрального уравнения (11).

6. Вычислим энергию системы в переменных $q_n(t)$.

Кинетическая энергия T :

$$T = \frac{\rho}{2} \int_V \dot{q}_n^2 dx dy dz = \frac{\rho}{2} \int_V \dot{q}_n^2 dx dy dz;$$

потенциальная энергия U :

$$U = \frac{g}{2} \int_S dx dy \frac{g}{2} q_n^2.$$

Составляя функцию Лагранжа $L = T - U$, мы убеждаемся что уравнения (12) суть уравнения Лагранжа второго рода малых движений жидкого континуума, если в качестве обобщенных координат возьмем счетную систему функций $q_n(t)$.

1952 г.

Задача о движении твердого тела, содержащего жидкие массы, имеющие свободную поверхность

Н.Е. Жуковским была рассмотрена в самой общей постановке задача о движении твердого тела, полости которого целиком наполнены несжимаемой жидкостью [1]. Данная работа посвящена анализу случая движения твердого тела при частичном заполнении полости. Более общая постановка задачи заставляет сузить класс рассматриваемых движений; мы ограничимся исследованием только малых движений около положения равновесия, т. е. внешние силы, действующие на сосуд, будем считать такими, что волны, возникающие в жидкости, обладают малыми амплитудами.

Рассматриваемая задача принципиально отличается от задачи Жуковского. Так, например, в случае, изученном Н.Е. Жуковским, можно ввести понятие эквивалентного твердого тела, т. е. такого твердого тела, движение которого при данных внешних силах эквивалентно движению тела с жидкостью. В нашем случае такого твердого тела не существует. Система сосуд — жидкость при наличии свободной поверхности будет системой с бесконечным числом степеней свободы, а это приводит к качественному отличию движения сосуда. Так, например, колебание физического маятника всегда устойчиво, как только его центр тяжести лежит ниже точки подвеса. В изучаемом случае последнее уже может быть не справедливо. Можно указать еще ряд простых примеров, иллюстрирующих это качественное различие.

Мы сохраняем основные предположения, принятые Н.Е. Жуковским: жидкость считаем идеальной и несжимаемой, движение жидкости в покоящемся сосуде — потенциальным. Таковым его можно считать и в последующие моменты времени относительно любой поступательно движущейся системы координат, если только массовые силы, действующие на жидкость, имеют силовую функцию. Последнее мы будем всегда предполагать. Более того, будем считать вектор напряженности массовых сил постоянным по направлению, а

его величину — функцией только времени. Таким образом, в качестве массовых сил рассматривается либо сила тяжести, либо сила инерции поступательного (параллельного некоторой прямой) движения.

Следовательно, в положении равновесия свободная поверхность представляет собой плоскость, перпендикулярную единичному вектору напряженности поля массовых сил g^0 .

В первых параграфах работы исследуется общий случай движения жидкости в подвижном и неподвижном сосудах; затем выясняется вид уравнений моментов и количества движения для случая малых движений сосуда около положения равновесия и выводится система дифференциальных уравнений движения системы сосуда с жидкостью. В последних параграфах рассматриваются несколько специальных случаев, имеющих определенный интерес для приложений. Конкретизация задач позволяет со значительной полнотой исследовать движение.

§ 1. Малые колебания тяжелой жидкости в сосуде*

1. Как известно, потенциал скоростей малых движений жидкости в сосуде $(x, y, z; t)$ является функцией, гармонической внутри области, ограниченной стенками сосуда и плоскостью $z = 0$ свободной поверхности в положении равновесия и удовлетворяющей следующим граничным условиям:

$$\text{на стенках сосуда } \frac{\partial \phi}{\partial n} = 0, \quad (1.1)$$

$$\text{на поверхности } z = 0 \quad \frac{\partial \phi}{\partial z} = \frac{1}{g} \frac{\partial^2 \phi}{\partial t^2}, \quad (1.2)$$

где g — ускорение силы тяжести.

Свободная поверхность определяется в этом случае условием

$$\phi(t; x, y) = \frac{1}{g} \frac{\partial^2 \phi}{\partial t^2} \Big|_{z=0}. \quad (1.3)$$

2. Пусть $\phi_n(x, y, z)$ — система функций, гармонических внутри удовлетворяющих условию (1.1) и образующих полную ортонормированную систему функций на плоской фигуре S , являющейся сечением полости сосуда плоскостью $z = 0$.

Потенциал можно искать в виде ряда

* Основные результаты этого параграфа изложены в нашей заметке [3].

$$(x, y, z; t) = \sum_n f_n(t) \varphi_n(x, y, z).$$

Обозначая через a_{nm} коэффициенты Фурье в разложении при $z = 0$ функции $\frac{\varphi_n}{z}$:

$$\frac{\varphi_n}{z} = \sum_m a_{nm} \varphi_m \quad (1.4)$$

и используя условие (1.2), получим для определения функций f_n следующую бесконечную систему линейных дифференциальных уравнений второго порядка:

$$f_n - g \sum_m a_{nm} f_m = 0, \quad n = 1, 2, \dots \quad (1.5)$$

В общем случае система (1.5) представляет собой систему совместных уравнений. В том частном случае, когда функции $\varphi_n(x, y, z)$ могут быть представлены в виде

$$\varphi_n(x, y, z) = \varphi_n(x, y) \varphi_n(z),$$

система (1.5) распадается на отдельные уравнения, так как в этом случае

$$a_{nm} = \delta_{nm} \frac{d \varphi_n}{dz} \Big|_{z=0},$$

где δ_{nm} — символ Кронекера, и мы сразу получаем уравнения главных колебаний

$$f_n - g \frac{d \varphi_n}{dz} \Big|_{z=0} f_n = 0, \quad n = 1, 2, \dots$$

3. Пусть функции $\varphi_n(x, y)$ образуют некоторую полную ортонормированную на S систему. Положим

$$(t; x, y) = \sum_n q_n(t) \varphi_n(x, y). \quad (1.6)$$

Тогда потенциал можно выразить следующей формулой:

$$(x, y, z; t) = H(x, y, z; \varphi_n, \varphi_n) \frac{(\varphi_n, \varphi_n)}{n} d, \quad (1.7)$$

где H — функция Грина задачи Неймана и интеграл распространен по поверхности S , ограничивающей объем V .

В силу соотношений (1.1) и (1.6), выражение (1.7) принимает вид

$$u(x, y, z; t) = \sum_n q_n(t) \bar{u}_n(x, y, z), \quad (1.8)$$

где

$$\bar{u}_n(x, y, z) = \int_S H(x, y, z; \xi, \eta, 0) \bar{u}_n(\xi, \eta) d\xi d\eta. \quad (1.8)$$

Обозначая через b_{nm} коэффициенты Фурье в разложении

$$\bar{u}_n(x, y, 0) = \sum_m b_{nm} u_m(x, y),$$

получим из условия (1.3) следующую бесконечную систему уравнений для определения функций $q_n(t)$:

$$\sum_m b_{nm} q_m - g q_n = 0. \quad (1.9)$$

4. Для определения главных колебаний и собственных частот нам надо разрешить одну из систем (1.5) или (1.9), которые в общем случае являются бесконечными системами совместных линейных уравнений. Если это возможно, то собственными частотами будут корни характеристических уравнений (1.5) или (1.9).

Возникает вопрос: каким должен быть базис $\bar{u}_n(x, y, z)$ или $u_n(x, y)$, чтобы система (1.5) или (1.9) в общем случае распалась на отдельные уравнения? Как нетрудно установить, таким базисом является нормированная система собственных функций интегрального уравнения

$$u_n(x, y) = \int_S H(x, y, 0; \xi, \eta, 0) u_n(\xi, \eta) d\xi d\eta. \quad (1.10)$$

В самом деле, в этом случае

$$b_{nm} = \delta_{nm} \frac{1}{g_n}, \quad (1.11)$$

где δ_{nm} — символ Кронекера, и система (1.9) приводится к виду

$$q_n - g_n q_n = 0. \quad (1.12)$$

Далее, положим

$$\eta_n(x, y) = \eta_n(x, y, 0), \quad q_n = f_n.$$

Тогда в линейном приближении тождество

$$\frac{d}{dt} \eta_n = -\omega_n \eta_n$$

примет вид

$$\eta_n(t) = \eta_n(x, y, 0) e^{-i\omega_n t},$$

откуда находим, что

$$a_{nm} = \delta_{nm} \eta_n,$$

т. е. функции $\eta_n(t)$ удовлетворяют той же системе дифференциальных уравнений (1.12).

5. Таким образом, мы приходим к следующей теореме:

Теорема. Квадрат собственной частоты свободных колебаний, жидкости в сосуде ω_n^2 равен произведению

$$\omega_n^2 = g \lambda_n,$$

где λ_n — собственные числа интегрального уравнения (1.10), а форма свободной поверхности, соответствующая главным колебаниям, определяется формулой

$$\eta_n(x, y; t) = A_n \sin \sqrt{g \lambda_n} t \lambda_n(x, y),$$

где A_n и λ_n — произвольные постоянные и λ_n — собственные функции того же интегрального уравнения.

6. Вычислим энергию жидкости в переменных $q_n(t)$. Кинетическая энергия T выражается формулой

$$T = \frac{\rho}{2} \int \dot{q}_n^2 dV.$$

Применяя теорему Грина и равенство (1.6), приведем эту формулу к виду

$$T = \frac{g}{2} \sum_n q_n^2. \quad (1.13)$$

Делая замену (1.6) в выражении потенциальной энергии U :

$$U = \frac{g}{2} \int_S dx dy,$$

получим:

$$U = \frac{g}{2} \sum_n q_n^2. \quad (1.14)$$

Составляя функцию Лагранжа $L = T - U$, мы убеждаемся, что уравнения (1.12) суть уравнения Лагранжа второго рода для малых движений жидкости, если в качестве обобщенных координат взять счетную систему функций $q_n(t)$.

Примечание. Так как q_n — коэффициенты разложения свободной поверхности $(x, y; t)$ в ряд Фурье, то для любого момента времени ряд (1.14) сходится. Имея в виду, что

$$q_n(t) = A_n \sin \sqrt{g} t, \quad n = 1, 2, \dots,$$

убеждаемся также в сходимости ряда (1.13).

7. Результаты данного параграфа легко распространить на тот случай, когда сосуд движется поступательно с некоторым ускорением параллельно линии действия силы тяжести (или в случае отсутствия притяжения — параллельно некоторой прямой).

Система уравнений (1.12) будет при этом иметь вид

$$q_n'' + g q_n = g(t) q_n \quad (n = 1, 2, \dots), \quad (1.12)$$

где $g(t)$ следует считать заданной функцией времени. Так, например, если мы будем изучать колебания тяжелой жидкости в сосуде, стоящем на платформе, вибрирующей по синусоидальному закону, то система уравнений (1.12) примет вид

$$q_n = g_0 a \cos(t) q_n = 0, \quad (1.12)$$

т. е. мы увидим, что каждая из переменных q_n удовлетворяет уравнению Матье. При определенных значениях a и $g(t)$ стоячие волны будут иметь увеличивающуюся со временем амплитуду.

Однако развивающиеся колебания в жидкости могут возникнуть и при монотонном изменении ускорения. Так, например, можно показать, что если $g(t) < 0$, где $g(t)$ — некоторая отрицательная постоянная, то в сосуде колебания будут развивающимися.

В самом деле, уравнение

$$q_n = g(t) q_n = 0$$

при помощи замены независимого переменного

$$\tau = \int_0^t g(t) dt$$

может быть приведено к виду

$$\frac{d^2 q_n}{d\tau^2} + \frac{1}{2} \frac{d}{d\tau} (\ln g) \frac{dq_n}{d\tau} + q_n = 0.$$

Составляя выражение энергии, отнесенной к координате q_n :

$$2V_n = \frac{dq_n}{d\tau} + q_n^2,$$

убеждаемся, что если $\frac{dg}{d\tau} < 0$, то движение устойчиво, так как $V_n < 0$.

В противном случае энергия V_n , отнесенная к каждой из координат, будет увеличиваться.

Характер изменения амплитуды стоячих волн в ряде случаев можно оценить количественно, пользуясь известными формулами асимптотических представлений:

$$q_n(t) = \frac{A_n}{\sqrt[4]{g(t)}} \cos \int_0^t \sqrt{g(t)} dt + \dots$$

где A_n и α_n — произвольные постоянные, так как во многих практически важных задачах произведение $\alpha_n g$ велико даже для малых номеров n .

§ 2. Движение жидкости в движущемся сосуде

1. Свяжем с твердым телом подвижную систему координат так, чтобы невозмущенная поверхность жидкости совпала с плоскостью O_1 . Обозначим через O точку, совпадающую в положении равновесия с центром тяжести системы сосуд — жидкость и неподвижную относительно сосуда. Свяжем с точкой O поступательно движущуюся систему координат x, y, z , такую, что в положении равновесия орты обеих систем коллинеарны. Будем предполагать, что оси подвижной системы координат x, y, z , все время параллельны осям некоторой «неподвижной» системы отсчета X, Y, Z .

Если переносное движение сосуда — поступательное, то движение жидкости относительно сосуда будет потенциальным, так как силы инерции, определяющие относительное движение жидкости, будут обладать силовой функцией

$$U(x, y, z) = Xx + Yy + Zz,$$

где X, Y, Z — координаты начала подвижной системы отсчета.

Если переносное движение будет вращательным, то относительное движение будет вихревым.

2. Рассмотрим сначала движение жидкости относительно сосуда, который движется поступательно. Это задача эквивалентна задаче о колебаниях жидкости в неподвижном сосуде под действием сил, имеющих потенциал

$$V = F_1 x + F_2 y + g(t) z.$$

Пользуясь интегралом Коши, можно представить уравнение свободной поверхности в виде

$$\left(x, y, z; t \right) = \frac{1}{g} \left[F_1 x + F_2 y + g(t) z \right]_0. \quad (2.1)$$

Обозначая через $\varphi_n(x, y, z)$ систему собственных функций интегрального уравнения (1.10), а через α_n — его собственные числа, положим:

$$\begin{aligned}
 (\dots; t) &= \sum_n q_n(t) \dots_n(\dots), \\
 (\dots; 0; t) &= \sum_n q_n(t) \frac{\dots_n(\dots)}{\dots_n}.
 \end{aligned}
 \tag{2.2}$$

Разлагая функции \dots и \dots в ряды Фурье на S по собственным функциям $\{\dots_n\}$:

$$\begin{aligned}
 \dots &= \sum_n \dots_n \dots_n, \\
 \dots &= \sum_n \dots_n \dots_n.
 \end{aligned}$$

и сравнивая коэффициенты при одинаковых функциях в уравнении (2.1), получим следующую бесконечную систему уравнений для определения функций q_n :

$$q_n \dots_n g_n q_n = F_{\dots_n} \dots_n - F_{\dots_n} \dots_n.
 \tag{2.3}$$

Из уравнений (2.3) видно, что резонансными являются частоты лишь тех индексов $n = \frac{2}{n} g_n$, для которых один из коэффициентов \dots_n или \dots_n отличен от нуля. Так, например, если сосуд имеет форму параллелепипеда, то

$$\dots_{2n} = \dots_{2n} = 0,^*$$

т. е. резонансными являются только частоты нечетных индексов, а амплитуды стоячих волн четных индексов не могут быть изменены движением сосуда в горизонтальной плоскости.

3. В случае произвольно движущегося сосуда относительное движение жидкости будет вихревым. Поэтому будем рассматривать движение жидкости относительно системы координат x, y, z . Следуя Н.Е. Жуковскому, выразим потенциал этого движения в координатах

$$(\dots, \dots; t).$$

* См. § 7.

Эта функция на стенках полости должна удовлетворять граничному условию

$$\frac{\partial \psi}{\partial n} = v_n,$$

где v_n — нормальная составляющая скорости соответствующей точки поверхности полости относительно системы координат x, y, z . В силу предположения о малости колебаний, скорость частиц жидкости при $\omega \rightarrow 0$ будет равна сумме

$$v_n = \frac{d\psi}{dt}.$$

Поэтому потенциал скоростей можно представить в виде

$$\psi(x, y, z; t) = H(x, y, z, t) + v_n(x, y, z) d + \dots \quad (2.4)$$

Первое слагаемое — не что иное как потенциал Н.Е. Жуковского [1]. Так как скорость v_n является линейной функцией проекций угловой скорости сосуда ω , то первое слагаемое может быть представлено в виде

$$\psi(x, y, z; t) = \omega(t) \psi_1(x, y, z) + \omega(t) \psi_2(x, y, z) + \omega(t) \psi_3(x, y, z).$$

Функции ψ_1, ψ_2, ψ_3 мы будем называть *функциями Жуковского*. Н.Е. Жуковский показал [1], что они являются гармоническими внутри V и удовлетворяют на S следующим граничным условиям:

$$\begin{aligned} \frac{\partial \psi_1}{\partial n} &= \cos \alpha_0 \psi_1(x_0, y_0, z_0) - \cos \beta_0 \psi_2(x_0, y_0, z_0), \\ \frac{\partial \psi_2}{\partial n} &= \cos \beta_0 \psi_1(x_0, y_0, z_0) - \cos \alpha_0 \psi_2(x_0, y_0, z_0), \\ \frac{\partial \psi_3}{\partial n} &= \cos \alpha_0 \psi_1(x_0, y_0, z_0) - \cos \beta_0 \psi_2(x_0, y_0, z_0), \end{aligned} \quad (2.5)$$

где $\cos \alpha_0, \cos \beta_0$ — направляющие косинусы внутренней нормали, а x_0, y_0, z_0 — координаты точки O .

Итак, потенциал скоростей движения жидкости относительно системы отсчета x, y, z в случае произвольного движения сосуда может быть представлен выражением

$$q_n - n,$$

где $-n(, ,)$ определяются по формулам (1.8).

Для определения функций $q_n(t)$ мы воспользуемся уравнением свободной поверхности, которое в этом случае будет иметь вид

$$(, ; t) \frac{1}{g} - t \Big|_0 [F g] [F g], \tag{2.6}$$

где g и g — проекции ускорения силы тяжести на оси O_1 и O_1 . Определим составляющие векторов \bar{F} и \bar{g} . Положение единичных векторов \bar{x}^0, \bar{y}^0 и \bar{z}^0 относительно неподвижной системы координат определим параметрами θ_1, θ_2 и θ_3 (см. фиг. 1). Они аналогичны углам скольжения, атаки и крена, вводимым обычно в теории самолета. Пренебрегая малыми второго порядка, получим:

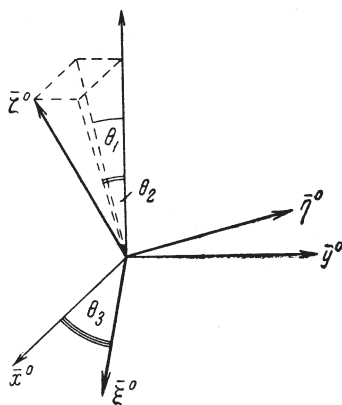
$$\begin{aligned} \bar{x}^0 &= \bar{x}^0 + \theta_2 \bar{z}^0 + \theta_3 \bar{y}^0, \\ \bar{y}^0 &= \bar{y}^0 + \theta_1 \bar{z}^0 + \theta_3 \bar{x}^0, \\ \bar{z}^0 &= \bar{z}^0 + \theta_2 \bar{x}^0 + \theta_1 \bar{y}^0, \\ \bar{x}^0 &= \bar{x}^0 + \theta_1 \bar{y}^0 + \theta_2 \bar{z}^0, \end{aligned} \tag{2.7}$$

и, следовательно, с той же точностью

$$\begin{aligned} X &= \bar{X}^0 + X^0, \\ Y &= \bar{Y}^0 + Y^0, \end{aligned}$$

т. е. $F = X, F = Y$, и аналогично

$$g = g_2, g = g_1.$$



Фиг. 1

Разлагая в ряды Фурье функции Жуковского:

$$\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, 0 \right)_{n-1} \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, 0 \right)_n \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, 0 \right)_{n+1} \quad (2.8)$$

из уравнения (2.6) получим следующую бесконечную систему уравнений для определения функций $q_n(t)$:

$$q_n - \frac{2}{n} q_n X_{n-1} Y_{n-1} - \frac{1}{n} g_{n-1} - \frac{2}{n} g_n = 0, \quad n = 1, 2, \dots \quad (2.9)$$

Каждое из уравнений этой системы может быть проинтегрировано независимо от других.

§ 3. Основные уравнения динамики системы сосуд — жидкость

В этом параграфе мы выясним тот вид, который принимают основные теоремы динамики в применении к данной системе в проекциях на выбранные нами оси.

1. *Количество движения сосуда с жидкостью и уравнение количества движения.* Как мы видели, движение твердого тела определяет движение заключенной в нем жидкости. В свою очередь, жидкость оказывает давление на стенки полости, которое зависит от скоростей частиц жидкости и от формы свободной поверхности. В общем случае главный вектор сил давления отличен от нуля, поэтому движение жидкости влияет на движение твердого тела. Так как свободная поверхность жидкости определяется счетным множеством коэффициентов Фурье ее разложения, то система сосуд — жидкость, если жидкость лишь частично его заполняет, представляет собой систему со счетным числом степеней свободы. Движение твердого тела, содержащего жидкие массы, мы будем представлять как сумму движений, поступательного и вращательного вокруг некоторой точки, неподвижной относительно сосуда. Такая точка зрения, как мы это видели, полностью применима и к исследованию жидкости в сосуде.

Введем следующие обозначения:

$$\begin{aligned} m_1 & \text{ — масса сосуда,} \\ m_2 & \text{ — масса жидкости,} \\ M = m_1 + m_2 & \text{ — масса системы,} \\ \bar{v}_0 & \text{ — абсолютная скорость точки } O, \\ \bar{v} & \text{ — скорость частиц относительно системы } x, y, z. \end{aligned}$$

Тогда количество движения системы может быть представлено в следующем виде:

$$\bar{K} = M \bar{v}_0 + \sum_i m_i \bar{v}_i \sim d, \quad (3.1)$$

где сумма распространена на все частицы сосуда. Так как

$$\sum_i m_i \bar{v}_i = m_1 \bar{v}_1, \quad \text{где } \bar{v}_1 \text{ — относительная скорость центра тяжести}$$

сосуда, то (3.1) можно переписать в виде

$$\bar{K} = M \bar{v}_0 + m_1 \bar{v}_1 \sim d.$$

Слагаемое

$$\sim d$$

представляет собой количество движения, которое имела бы жидкость, если бы свободную поверхность мы заменили твердой крышкой; поэтому

$$\sim d = m_2 \bar{v}_2,$$

где \bar{v}_2 — относительная скорость центра тяжести отвердевшей жидкости в объеме . В силу несжимаемости жидкости имеем:

$$m_1 \bar{v}_1 = m_2 \bar{v}_2 = 0.$$

Поэтому получим окончательно:

$$\bar{K} = M \bar{v}_0 \sim d. \quad (3.2)$$

Таким образом, количество движения системы сосуд — жидкость равно количеству движения, которое имел бы сосуд, если бы жидкость в нем отвердела, плюс количество движения, происходящего вследствие волн. Последнее слагаемое равно произведению массы жидкости на скорость движения ее центра тяжести относительно точки O .

Таким образом, уравнение количества движения мы можем представить в виде

$$M \frac{d\bar{v}_0}{dt} = \frac{d}{dt} \bar{Q}, \quad (3.3)$$

где \bar{Q} — главный вектор внешних сил.

Обозначая через X, Y, Z координаты точки O в инерциальной системе отсчета и замечая, что

$$\begin{aligned} \frac{d}{dt} \sum_n q_n \frac{d}{dt} q_n, \\ \frac{d}{dt} q_n \end{aligned} \quad (3.3)$$

$$\frac{d}{dt} \frac{d}{dt} \text{ — бесконечно малая второго порядка,}$$

перепишем окончательно уравнение (3.3) в проекциях на неподвижные оси:

$$\begin{aligned} MX &= \sum_n q_n Q_x, \\ MY &= \sum_n q_n Q_y, \\ MZ &= Q_z, \end{aligned} \quad (3.4)$$

где величины $\sum_n q_n$, $\sum_n q_n$ не зависят от движения сосуда и определяются только геометрией той части сосуда, которая занята жидкостью.

Уравнения (3.4) представляют собой выражение теоремы количества движения сосуда с жидкостью в выбранных переменных.

2. *Кинетический момент.* Кинетический момент мы можем представить в виде

$$\bar{G} = \bar{G}_1 + (\bar{r} \times \bar{\omega}) d = \bar{r} \times \sum_n q_n \bar{\omega}_n d,$$

где \bar{r} — радиус-вектор жидкой частицы относительно точки O и \bar{G}_1 — кинетический момент сосуда. Преобразуем второе слагаемое; для этого введем вектор \bar{a} , имеющий компоненты

$$a = 0, \quad a = (\quad 0), \quad a = 0.$$

Тогда проекцию этого слагаемого на ось O_1 мы можем представить в виде

$$G_2 = (\sim \bar{a}) d.$$

Так как $\operatorname{div} \bar{a} = 0$, то, применяя формулу Грина, получим:

$$G_2 = \sim a_n d,$$

где — поверхность, ограничивающая объем . Но

$$a_n = (\quad 0) (\quad 0) \frac{1}{n},$$

где , , — направляющие косинусы внутренней нормали. Применяя еще раз формулу Грина, получим:

$$G_2 = \sim (\quad) d.$$

Вводя обозначения Жуковского:

$$uv = (u \ v) d$$

и полагая

$$\sim = \dots,$$

получим окончательно:

$$G_2 = \dots] \dots] \dots]$$

и совершенно аналогично

$$G_2 = \dots] \dots] \dots],$$

$$G_2 = \dots] \dots] \dots].$$

Следовательно,

$$\bar{G}_2 = \frac{1}{2} f_2(\dots),$$

где

$$f_2(\dots) = \dots$$

Пусть эллипсоид инерции сосуда задан уравнением

$$f_1(x, y, z) = 1.$$

Положим

$$\bar{G}_0 = \bar{G}_1 = \bar{G}_2 = \frac{1}{2} (\dots),$$

где

$$f_1 = f_2.$$

Таким образом, эллипсоид

$$(x, y, z) = 1 \tag{3.5}$$

является эллипсоидом инерции эквивалентного твердого тела (в смысле Жуковского), т. е. твердого тела, эквивалентного сосуду с жидкостью в том случае, если мы заменим свободную поверхность твердой горизонтальной стенкой.

Преобразуем теперь третье слагаемое в выражении кинетического момента:

$$\bar{G}_3 = \sum_n q_n (\bar{r} \sim_n) d.$$

Повторяя приведенные выше рассуждения, мы получим:

$$G_3 = q_n [\dots] ; G_3 = q_n [\dots] ; G_3 = q_n [\dots] ;$$

другими словами,

$$\bar{G}_3 = \frac{1}{2} f_3(\dots),$$

где

$$f_3 = 2 q_n [\Gamma_n] [\Gamma_n] [\Gamma_n],$$

и, следовательно, общее выражение кинетического момента системы сосуд — жидкость может быть записано в виде

$$\bar{G} = \frac{1}{2} \Phi(\dots) + \frac{1}{2} f_3(\dots). \quad (3.6)$$

Функция Φ — однородная положительно определенная квадратичная форма, а функция f_3 — однородная линейная форма от проекций угловой скорости.

Таким образом, кинетический момент системы сосуд — жидкость равняется сумме кинетического момента абсолютно твердого тела, эквивалентного системе при замене свободной поверхности горизонтальной стенкой, и кинетического момента волнового движения жидкости в покоящемся сосуде.

Эллипсоид (3.5) может быть приведен к главным осям. Теперь его уравнение будет иметь вид

$$Ax^2 + By^2 + Cz^2 = 1, \quad (3.7)$$

где A, B, C — моменты инерции эквивалентного твердого тела. Как показал Н.Е. Жуковский, они меньше моментов инерции соответствующих жидких масс.

3. *Уравнение моментов.* Так как благодаря волнам на свободной поверхности центр тяжести системы перемещается относительно сосуда, то уравнение моментов относительно центра O должно быть представлено в виде

$$\bar{G} = \bar{r}_O M \frac{dv_O}{dt} + \bar{L}, \quad (3.8)$$

где \bar{G} — момент количества движения тела с жидкостью в поступательно движущейся системе координат x, y, z относительно центра O , \bar{v}_O — скорость точки O , \bar{r}_O — радиус-вектор центра тяжести,

\bar{L} — главный момент внешних сил относительно того же центра.
Произведение

$$\bar{r}_O M \frac{dv_O}{dt}$$

задает момент сил инерции относительно центра O .

Так как мы изучаем движение только с малыми ускорениями и малыми волнами, то второе слагаемое в левой части уравнения (3.8) будет малой второго порядка. Исходя из этих соображений мы можем пренебречь моментом сил инерции и представить уравнение моментов в виде

$$\frac{d\bar{G}}{dt} = \bar{L}. \quad (3.8)$$

Если в качестве подвижной системы координат мы возьмем главные оси инерции эквивалентного твердого тела, т. е. главные оси эллипсоида (3.7), и напишем уравнение (3.8) в проекциях на эти оси, то получим уравнения, аналогичные уравнениям, выведенным Н.Е. Жуковским.

Пусть матрица $\|i_j\|$ определяет положение новых осей $\bar{p}^0, \bar{q}^0, \bar{r}^0$ относительно старых p, q, r . Тогда

$$f_3 = 2 \left[q_n \begin{bmatrix} - & p \end{bmatrix} p + \begin{bmatrix} - & q \end{bmatrix} q + \begin{bmatrix} - & r \end{bmatrix} r \right] \\ 2 \left[\begin{bmatrix} - & p \end{bmatrix} p + \begin{bmatrix} - & q \end{bmatrix} q + \begin{bmatrix} - & r \end{bmatrix} r \right],$$

где $\begin{matrix} p & 11 & 21 & 31, & q & 12 & 22 & 32, \\ r & 13 & 23 & 33, \end{matrix}$ p, q, r — проекции угловой скорости на оси эллипсоида (3.7). Следовательно, выражение (3.6) примет вид

$$\bar{G} = \frac{1}{2} \Phi(p, q, r) \left(\begin{bmatrix} - & p \end{bmatrix} d \bar{p}^0 + \begin{bmatrix} - & q \end{bmatrix} d \bar{q}^0 + \begin{bmatrix} - & r \end{bmatrix} d \bar{r}^0 \right).$$

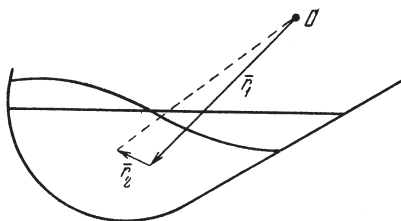
Так как $\begin{bmatrix} - & p \end{bmatrix} d = \sum_s \frac{d}{dt} d$ и т. д., то, принимая во внимание обозначения (2.8), получим:

$$\bar{G} = \frac{1}{2} \Phi(p, q, r) = q_n (p_n \bar{p}^0 + q_n \bar{q}^0 + r_n \bar{r}^0), \quad (3.9)$$

и, следовательно, скалярные уравнения моментов будут иметь вид

$$\begin{aligned} Ap &= \sum_n p_n q_n L_p, \\ Bq &= \sum_n q_n q_n L_q, \\ Cr &= \sum_n r_n q_n L_r, \end{aligned} \quad (3.10)$$

где A, B, C — моменты инерции эквивалентного твердого тела. Величины p_n, \dots определяются только геометрией полости. В правой части этих уравнений стоят компоненты главного момента внешних сил. Этот момент мы будем рассматривать как состоящий из двух слагаемых. Первое слагаемое L_1 будет зависеть от обобщенных координат и скоростей сосуда и времени. Это слагаемое может быть не консервативным.



Фиг. 2

Второе слагаемое L_2 — это момент массовых сил, действующих на жидкость, и, следовательно, зависит от обобщенных координат q_n . Это слагаемое всегда будет консервативным. Выясним его структуру. Момент силы тяжести относительно точки O будет равен

$$(\bar{r}_1 - \bar{r}_2) m_2 \bar{g},$$

где \bar{r}_1 — радиус-вектор центра тяжести жидкости при отсутствии волн, \bar{r}_2 — смещение центра тяжести вследствие волн.

Первое слагаемое будет зависеть только от положения сосуда и, следовательно, войдет в L_1 , вторым слагаемым и будет L_2 . Пусть в положении равновесия центр тяжести жидкости лежит на оси O_1 ; тогда вектор \bar{r}_2 имеет проекции

$$r = \frac{d}{m_2}, \quad r = \frac{d}{m_2}, \quad r = \frac{d}{m_2} \quad y, \quad (3.11)$$

где y — аппликата центра тяжести жидкости в положении равновесия. Интегрирование в этих выражениях распространено на весь объем, занимаемый жидкостью.

Очевидно, что

$$r = \frac{d}{m_2} \frac{d}{m_2} \frac{d}{m_2} q_n \dots$$

Аналогично

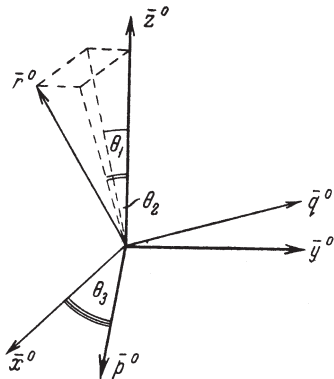
$$r = \frac{q_n}{m_2}$$

Нетрудно видеть, что величина r будет величиной второго порядка малости. Поэтому

$$\bar{L} = g \frac{q_n}{m_2} \dots \bar{Z}^0.$$

Прежде чем вычислить проекции вектора \bar{L}_2 на направления $\bar{p}^0, \bar{q}^0, \bar{r}^0$, выведем несколько вспомогательных соотношений.

Пусть углы θ_1, θ_2 и θ_3 определяют положение осей p, q, r относительно системы отсчета x, y, z и в положении равновесия обе системы совпадают. Тогда с принятой точностью получим формулы перехода:



$$\begin{aligned} \bar{p}^0 &= \bar{x}^0 \cos \theta_1 + \bar{z}^0 \sin \theta_1, \\ \bar{q}^0 &= \bar{y}^0 \cos \theta_2 + \bar{z}^0 \sin \theta_2, \\ \bar{r}^0 &= \bar{x}^0 \sin \theta_3 + \bar{y}^0 \sin \theta_3 + \bar{z}^0 \cos \theta_3. \end{aligned} \quad (3.12)$$

В свою очередь, положение системы x, y, z относительно неподвижной системы будет задано матрицей направляющих косинусов $\| \cos_{ij} \|$. Вычислим проекции единичного вектора \bar{Z}^0 на оси p, q, r . Так как

$$\bar{Z}^0 = \cos_{31} \bar{x}^0 + \cos_{32} \bar{y}^0 + \cos_{33} \bar{z}^0, \quad (3.12)$$

то, пользуясь формулами (3.12), получим:

$$\begin{aligned} \bar{Z}^0 & \left(\begin{matrix} 31 & 32 & 33 \\ 2 & 3 & 2 \end{matrix} \right) \bar{p}^0 \left(\begin{matrix} 31 & 3 & 32 \\ 33 & 1 & 1 \end{matrix} \right) \bar{q}^0 \\ & \left(\begin{matrix} 31 & 2 & 32 \\ 2 & 1 & 33 \end{matrix} \right) \bar{r}^0 \end{aligned} \quad (3.12)$$

и, следовательно, отбрасывая малые второго порядка,

$$\begin{aligned} L_{2p} & \quad g \quad q_n \left(\begin{matrix} n a_1 & n a_2 \end{matrix} \right), \\ L_{2q} & \quad g \quad q_n \left(\begin{matrix} n b_1 & n b_2 \end{matrix} \right), \\ L_{2r} & \quad g \quad q_n \left(\begin{matrix} n c_1 & n c_2 \end{matrix} \right), \end{aligned} \quad (3.13)$$

где

$$\begin{aligned} a_1 & \quad 33 \quad 12 \quad 13 \quad 32, \\ a_2 & \quad 22 \quad 33 \quad 23 \quad 32, \\ b_1 & \quad 13 \quad 31 \quad 11 \quad 33, \\ b_2 & \quad 23 \quad 31 \quad 21 \quad 33, \\ c_1 & \quad 11 \quad 32 \quad 12 \quad 31, \\ c_2 & \quad 21 \quad 31 \quad 22 \quad 32. \end{aligned} \quad (3.14)$$

Таким образом, уравнения моментов мы можем представить в виде

$$\begin{aligned} A_1 & \quad q_n \quad p_n \quad g \quad q_n \left(\begin{matrix} n a_1 & n a_2 \end{matrix} \right) L_{1p}, \\ B_2 & \quad q_n \quad q_n \quad g \quad q_n \left(\begin{matrix} n b_1 & n b_2 \end{matrix} \right) L_{1q}, \\ C_3 & \quad q_n \quad r_n \quad g \quad q_n \left(\begin{matrix} n c_1 & n c_2 \end{matrix} \right) L_{1r}. \end{aligned}$$

Правые части этих уравнений зависят только от обобщенных координат сосуда $1, 2, 3$, их производных и времени, и не зависят от переменных q_n .

Полученные уравнения являются обобщением соответствующих уравнений Н.Е. Жуковского для малых движений сосуда с

жидкостью. Последние получаются из наших уравнений, если положить все $q_n = 0$.

§ 4. Общие уравнения движения сосуда с жидкостью.

Интеграл энергии

1. Во втором параграфе мы вывели уравнения, которым удовлетворяют переменные q_n . Полученная система (2.9) совместно с уравнением моментов и уравнением количества движения образует замкнутую систему уравнений. Нам осталось только несколько изменить обозначения соответственно выбранной в предыдущем параграфе системе отсчета.

Для этого вычислим g и g , пользуясь обозначениями предыдущего параграфа. Так как

$$g = g(\bar{Z}^0), \quad g = g(\bar{Z}^0),$$

то, пользуясь равенствами

$$\bar{p}^0 = \bar{p}^0, \quad \bar{q}^0 = \bar{q}^0, \quad \bar{r}^0 = \bar{r}^0 \text{ и т. д.}$$

и равенством (3.12) и принимая во внимание условие ортогональности ортов \bar{e}_1^0, \bar{e}_2^0 и \bar{e}_3^0 , получим:

$$g = g(a_{11}, b_{12}, c_{13}), \quad g = g(a_{21}, b_{22}, c_{23}).$$

Используя принятые обозначения, мы можем сформулировать полученные результаты в виде следующей теоремы (см. [2]):

Теорема. *Малые движения твердого тела, содержащего жидкие массы, имеющие свободную поверхность, около положения равновесия описываются следующей бесконечной системой линейных дифференциальных уравнений:*

1) Уравнения количества движения:

$$\begin{aligned} MX &= \sum_n q_n Q_x, \\ MY &= \sum_n q_n Q_y, \\ MZ &= Q_z. \end{aligned} \tag{4.1}$$

2) Уравнения моментов:

$$\begin{aligned}
 A \quad & \ddot{q}_n + p_n \dot{q}_n + g_n q_n (n a_1 + n a_2) = L_{1p}, \\
 B \quad & \ddot{q}_n + q_n \dot{q}_n + g_n q_n (n b_1 + n b_2) = L_{1q}, \\
 C \quad & \ddot{q}_n + r_n \dot{q}_n + g_n q_n (n c_1 + n c_2) = L_{1r}.
 \end{aligned} \tag{4.2}$$

3) Уравнения для коэффициентов Фурье q_n разложения свободной поверхности жидкости в ряд по собственным функциям интегрального уравнения (1.10):

$$\begin{aligned}
 q_n + \frac{2}{n} \ddot{q}_n + X_{nn} \dot{Y}_{nn} + 1 p_n + 2 q_n + \\
 + 3 r_n + g_n [n a_1 + n a_2]_1 + [n b_1 + n b_2]_2 + \\
 + [n c_1 + n c_2]_3 = 0, \quad n = 1, 2, \dots,
 \end{aligned} \tag{4.3}$$

где числа $A, B, C, p_n, q_n, a_1, a_2, b_1, b_2, c_1, c_2$ зависят только от геометрии сосуда и плотности жидкости.

Таким образом, содержащиеся в твердом теле жидкие массы не только приводят к изменению моментов инерции по сравнению с соответствующими моментами инерции твердого тела, но и изменяют характер системы дифференциальных уравнений. Последнее принципиально отличает нашу задачу от задачи, рассмотренной Жуковским, и для случая малых колебаний полученные результаты являются прямым обобщением результатов Жуковского. Для того чтобы убедиться в этом, достаточно положить $q_n = 0$.

Можно показать, что полученная система уравнений (4.1), (4.2) и (4.3) является системой уравнений Лагранжа, если в качестве обобщенных координат принять параметры \dot{p}, q_n и координаты точки O , т. е. X, Y, Z .

2. Подсчитаем потенциальную энергию жидкости. Потенциальная энергия системы складывается из потенциальной энергии внешних сил, зависящих от обобщенных координат сосуда P_1 , и потенциальной энергии жидкости P_2 , зависящей от обобщенных координат q_n . Нас будет интересовать только второе слагаемое.

Пользуясь обозначениями предыдущего параграфа, можно его представить в виде

$$P_2 = (h_1 - h_{от}) m_2 g, \quad (4.4)$$

где $h_{от}$ — высота поднятия центра массы жидкости в ее движении относительно системы O, X, Y ,

$$h_1 = \int_0^t (\bar{v}_2 - \bar{v}_1) \bar{Z}^0 dt,$$

$\bar{\omega}$ — угловая скорость вращения сосуда.

После очевидных преобразований выражение (4.4) примет вид

$$P_2 = g \left[\bar{p}^0 + \bar{q}^0 + \bar{r}^0 \right] \frac{1}{2} \bar{Z}^0 \frac{g}{2} \bar{q}_n^2.$$

Пользуясь формулами (3.12), (3.12'), получим окончательно:

$$P_2 = g \left[\bar{p}_n^0 + \bar{q}_n^0 + \bar{r}_n^0 \right] \frac{1}{2} \bar{Z}_n^0 \frac{g}{2} \bar{q}_n^2. \quad (4.5)$$

3. Подсчитаем кинетическую энергию сосуда с жидкостью. Она равна

$$T = \frac{1}{2} \int (\bar{v}_0 - \bar{v})^2 dV, \quad (4.6)$$

где \bar{v}_0 — скорость точки O и \bar{v} — относительная скорость. Интеграл в выражении (4.6) распространен на весь объем, занимаемый сосудом с жидкостью. Энергию T можно представить в следующем виде:

$$T = M \frac{\bar{v}_0^2}{2} + m_1 \bar{v}_1^2 + \dots - d$$

$$\frac{1}{2} \bar{v}_0^2 + \dots - d + \frac{1}{2} \bar{v}_1^2 + \dots - d + T_1,$$

где \bar{v}_1 — относительная скорость центра массы сосуда и T_1 — кинетическая энергия относительного движения сосуда.

В силу несжимаемости жидкости, очевидно, что

$$m_1 \bar{v}_1 \sim d = 0.$$

Согласно соотношениям (3.3), отбрасывая малые второго порядка, получим:

$$\bar{v}_0 \sim d \left(X_n q_n + Y_n q_n \right).$$

Сумма $T_0 = \frac{1}{2} \sim^2 d T_1$ представляет собой кинетическую энергию эквивалентного твердого тела в его относительном движении:

$$T_0 = \frac{1}{2} (Ap^2 + Bq^2 + Cr^2).$$

Далее найдем:

$$\sim -d \left(p_n q_n + q_n q_n + r_n q_n \right).$$

Наконец, предпоследнее слагаемое, являющееся энергией волнового движения жидкости в покоящемся сосуде, выражается формулой

$$\frac{1}{2} \sim^2 d \left(\frac{q_n^2}{n} \right).$$

Таким образом, кинетическая энергия системы сосуд — жидкость имеет вид

$$T = \frac{M}{2} (X^2 + Y^2 + Z^2) + X_n q_n + Y_n q_n + \frac{1}{2} (Ap^2 + Bq^2 + Cr^2) + q_n p_n + q_n q_n + q_n r_n + \frac{1}{2} \frac{q_n^2}{n}. \quad (4.7)$$

Пользуясь выражениями энергии (4.6) и (4.7), мы непосредственно убеждаемся в том, что уравнения (4.1)—(4.4) суть уравнения Лагранжа в обобщенных координатах X, Y, Z, q_1, q_2, \dots

4. Таким образом, если внешние силы консервативны и их потенциальную энергию обозначим через Π_1 , то система уравнений (4.1)—(4.4) допускает первый интеграл — интеграл энергии:

$$H = T + \Pi_1 + \Pi_2.$$

§ 5. Малые колебания при потенциальных силах

1. Будем рассматривать колебания сосуда с жидкостью в предположении, что внешние силы, зависящие от обобщенных координат сосуда, имеют силовую функцию, и примем, что положению равновесия соответствуют нулевые значения переменных X, \dots, q_1, \dots

Тогда, вводя обозначения

$$X = X_1, Y = X_2, Z = X_3, q_1 = X_4, q_2 = X_5, q_3 = X_6,$$

мы можем представить выражение потенциальной энергии внешних сил Π_1 в виде

$$\Pi_1 = \frac{1}{2} \sum_{i,j=1}^6 b_{ij} X_i X_j. \quad (5.1)$$

Следовательно, система уравнений (4.1)—(4.3) примет вид

$$\begin{aligned} a_j X_j - \sum_{i=1}^6 b_{ji} X_i - (q_n - j_n - q_n - j_n) = 0, \quad j = 1, 2, \dots, 6, \\ \sum_{i=1}^6 (X_i - i_n - X_i - i_n) - q_n = g - q_n = 0, \quad n = 1, 2, \dots, \end{aligned} \quad (5.2)$$

где

$$\begin{aligned} 1_n &= p_n, \quad 2_n = q_n, \quad 3_n = r_n, \\ 4_n &= p_n, \quad 5_n = q_n, \quad 6_n = r_n, \\ i_n &= 0, \quad i = 1, 2, 3, \\ 4_n &= g(n a_1 - n a_2), \quad 5_n = g(n b_1 - n b_2), \\ 6_n &= g(n c_1 - n c_2). \end{aligned}$$

Аналогично будет упрощено выражение потенциальной энергии:

$$U = U_2 + U_1 - \frac{1}{2} \sum_{m=1}^6 Y_m^2 - \sum_{n=1}^6 q_n \sum_{m=1}^6 v_{mn} Y_m - \frac{1}{2} \sum_n g_n q_n^2. \quad (5.5)$$

Коэффициенты v_{mn} и g_n выражаются линейно через коэффициенты Y_{m1} и Y_{m2} . Таким образом, уравнения (5.2) примут вид

$$Y_m - \sum_{n=1}^6 Y_{mn} - (q_n v_{mn} - q_n) = 0, \quad (5.6)$$

$$\sum_{m=1}^6 (Y_{m1} v_{mn} - Y_{m2} v_{mn}) - q_n - g_n = 0, \quad n = 1, 2, \dots, 6.$$

Ограничиваясь конечным числом n членов разложения в ряд Фурье, мы получим обычную консервативную систему. Ее уравнение частот можно представить в виде

$$\left[\begin{array}{c} \overbrace{\quad\quad\quad}^{6 \text{ столбцов}} \quad \overbrace{\quad\quad\quad}^{n \text{ столбцов}} \\ \begin{array}{cccccccc} 1 & 0 & \dots & \frac{v_{11}}{2} & \frac{v_{12}}{2} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{v_{11}}{2} & \frac{v_{21}}{2} & \dots & 1 & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{v_{12}}{2} & \frac{v_{22}}{2} & \dots & 0 & 1 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{array} \\ \dots \end{array} \right] = 0.$$

Вычисляя корни этого уравнения мы можем привести систему уравнений к нормальному виду:

$$Z_k - \omega_k^2 Z_k = 0 \quad (k = 1, 2, \dots, n + 6), \quad (5.7)$$

где ω_k — главные частоты колебаний системы *сосуд — жидкость*.

Для любого сколь угодно большого числа n система (5.6) может быть разрешена и приведена к нормальному виду (5.7). Переходя к пределу при $n \rightarrow \infty$, покажем, что бесконечная система алгебраических уравнений (5.3) остается разрешимой и значения собственных частот ω_i , найденные из укороченной системы уравнений, стремятся к определенным пределам при неограниченном увеличении числа уравнений. Для этого нам достаточно показать, что детерминант

$$D_n(\omega) = \det \|g_{ij}(\omega)\|$$

удовлетворяет условиям теоремы Коха, т. е. произведение диагональных элементов и суммы квадратов всех прочих элементов сходится [6]. Нам достаточно доказать лишь второе из этих утверждений. Но оно очевидно. В самом деле, $D_n(\omega)$ является мероморфной функцией, имеющей полюсы в точках

$$\omega = \pm \sqrt{g_{jj}}, \quad j = 1, 2, \dots, n.$$

Для всех прочих значений ω каждый из рядов

$$\sum_{j=1}^n \frac{g_{ij}^2 + v_{ij}^2}{g_{jj}^2},$$

или

$$\sum_{j=1}^n \frac{g_{ij}^2 + v_{ij}^2}{g_{jj}^2}, \quad i = 1, 2, \dots, n,$$

сходится, в силу того что g_{ij}, v_{ij} — коэффициенты Фурье (т. е. ряды

$$\sum_j \frac{g_{ij}^2}{j^2}, \quad \sum_j \frac{v_{ij}^2}{j^2}$$
 являются сходящимися).

3. Из разрешимости системы (5.6) и возможности привести ее к нормальному виду следует справедливость теоремы Лагранжа, которая в данном случае может быть сформулирована следующим образом:

Теорема. Для устойчивости тривиального решения системы (5.6) необходима и достаточна положительная определенность квадратичной формы (5.5).

В самом деле, возможность приведения к нормальному виду эквивалентна одновременному приведению к диагональному виду квадратичных форм (5.4) и (5.5), при котором квадратичная форма (5.5) принимает вид

$$U = \frac{1}{2} \sum_{n=1}^6 (\lambda_n^2) Z_n^2.$$

Следовательно, если все λ_n — чисто мнимые, т. е. если квадратичная форма (5.5) — положительно определенная, то каждое из нормальных колебаний будет устойчиво. Обратно, если все нормальные колебания устойчивы, то λ_n^2 должно быть отрицательным.

Найдем условия знакоопределенности квадратичной формы (5.5). Для этого введем новые переменные

$$q_n = U_n \sum_{m=1}^6 \frac{v_{mn}}{g} Y_m. \quad (5.8)$$

После очевидных преобразований выражение (5.5) примет вид

$$2U = g \sum_{n=1}^6 U_n^2 \sum_{i,j=1}^6 Y_i Y_j,$$

где

$$ij = \sum_{n=1}^6 \frac{v_{in} v_{jn}}{g}, \quad ii = \sum_{n=1}^6 \frac{v_{in}^2}{g}.$$

Следовательно, имеет место следующее предложение:

Для положительной определенности квадратичной формы (5.5) необходима и достаточна положительная определенность квадратичной формы

$$L = \sum_{n,m=1}^6 Y_n Y_m. \quad (5.9)$$

Таким образом, критерием устойчивости тривиального решения системы (5.5) является выполнение неравенств Сильвестра для квадратичной формы (5.9).

4. Пусть корни уравнения частот ω_j нами определены. Тогда, обозначая через M_{kj} минор k -го элемента первой строки характеристического определителя, вычисленный в предположении $\omega_j = i$, мы получим значения переменных в явном виде:

$$\begin{aligned}
 Y_1 &= M_{11} Z_1, \\
 Y_2 &= M_{21} Z_1, \\
 &\dots \dots \dots \\
 q_n &= M_{n1} Z_1, \\
 &\dots \dots \dots
 \end{aligned}
 \tag{5.10}$$

Найдем форму свободной поверхности:

$$\eta(x, y, t) = \sum_n q_n(t) \cos(k_n x),$$

или, принимая во внимание формулы (5.10),

$$\eta(x, y, t) = \sum_{n,m} M_{n1} Z_m(t) \cos(k_n x).$$

Введем обозначения:

$$M_m(x, y) = \sum_n M_{n1} \cos(k_n x),$$

Получим, далее:

$$\eta(x, y, t) = \sum_n Z_m(t) M_m(x, y),
 \tag{5.11}$$

т. е. свободная поверхность может быть представлена как сумма стоячих волн $M_m(x, y) Z_m(t)$, каждая из которых колеблется с частотой $|\omega_m|$.

§ 6. Случай, когда сосуд имеет одну степень свободы.

Влияние диссипативных сил

1. Этот случай представляет значительный практический интерес, так как к нему сводится ряд конкретных задач: задачи о

колебаниях разного рода маятников с жидкостью, о качке (бортовой или килевой) наливных сосудов и т. д. К тому же этот случай является наиболее простым и поэтому может быть детально изучен. Исследованию поступательных колебаний прямоугольного сосуда с жидкостью была посвящена работа Л.Н. Сретенского [5]. Он вычислял суммарное давление на стенки сосуда и затем составлял уравнение его движения. Пользуясь нашими результатами, мы проведем исследование движения сосудов произвольной формы для произвольных малых колебаний как поступательных, так и вращательных.

2. Предположим сначала, что внешние силы имеют потенциал, и обозначим через Y обобщенную координату сосуда. Тогда система (5.6) примет вид

$$\begin{aligned}
 Y & \quad {}^2 Y \quad (q_n \quad q_n v_n) \quad 0, \\
 {}_n Y \quad v_n Y & \quad - q_n \quad g \quad q_n \quad 0, \quad n = 1, 2,
 \end{aligned}
 \tag{6.1}$$

Соответствующее уравнение частот будет иметь вид

$$0, \tag{6.2}$$

где

$$\begin{vmatrix}
 1 & \frac{{}_1^2 v_1}{2} & \frac{{}_2^2 v_2}{2} & \dots \\
 \frac{{}_1^2 v_1}{g} & 1 & 0 & \dots \\
 \frac{{}_2^2 v_2}{g} & 0 & 1 & \dots \\
 \dots & \dots & \dots & \dots
 \end{vmatrix} = 0.$$

Вычисляя определитель, приведем уравнение (6.2) к виду

$$\Phi_1(s^2) = \frac{n(n^2 - \nu_n^2)^2}{(s^2 - \frac{\nu_n^2}{n})^2} \quad (6.3)$$

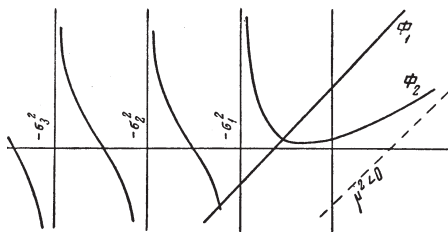
Функция

$$\Phi_1(s^2) = \frac{n(n^2 - \nu_n^2)^2}{(s^2 - \frac{\nu_n^2}{n})^2}$$

является мероморфной функцией s^2 , имеющей простые полюсы в точках

$$s^2 = g_n = \frac{\nu_n^2}{n}.$$

Распределение частот проще всего исследовать графически. Из чертежа (фиг. 4) видно, что корни уравнения (6.2) лежат между полюсами функции $\Phi_2(s^2)$. Однако корни уравнения (6.2) могут лежать и справа от оси ординат, если квадратичная форма (5.9) не является положительно определенной. В данном случае



Фиг. 4

$$L_{11} Y^2,$$

и, следовательно, единственным условием устойчивости будет выполнение следующего неравенства:

$$L_{11} \frac{\nu_n^2}{g} > 0.$$

3. Во многих технических задачах ряд $\frac{n(n^2 - \nu_n^2)^2}{(s^2 - \frac{\nu_n^2}{n})^2}$ оказывается настолько быстро сходящимся, что его вычисление при фиксированном s не представляет затруднений. Это обстоятельство позволяет эффективно применить метод последовательных приближений. Перенумеруем собственные частоты в порядке возрастания их абсолютных величин: ν_1, ν_2, \dots так, что

$$\nu_1 < \nu_2 < \nu_3 < \dots$$

При вычислении ξ_0 в качестве первого приближения имеет смысл брать решение биквадратного уравнения

$$\xi^2 - \xi + \frac{1}{2} \left(\frac{v_1}{\xi} \right)^2 = 0. \quad (6.4)$$

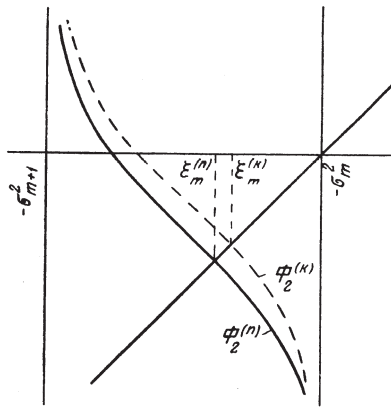
В ряде случаев уже это приближение может дать удовлетворительную точность.

Обозначим через $\xi_m^{(n)}$ корень уравнения

$$\xi^2 - \xi + \frac{1}{2} \left(\frac{v_i}{\xi} \right)^2 = 0, \quad (6.5)$$

удовлетворяющий неравенству

$$\xi_m^{(n)} \geq \xi_m^{(k)}. \quad (6.6)$$



Фиг. 5

В дальнейшем нам будет необходима следующая

Теорема. $\xi_m^{(n)}$ — монотонная функция n и стремится к определенному конечному пределу при неограниченном возрастании n .

Нам, очевидно, достаточно доказать только монотонность функции $\xi_m^{(n)}$, так как корни уравнения (6.2) удовлетворяют неравенству (6.6) для любого n .

Фиксируем некоторое значение m . Тогда, очевидно,

$$\Phi_2^{(n)} = \frac{1}{2} \left(\frac{v_i}{\xi_m^{(n)}} \right)^2 < \Phi_2^{(k)} = \frac{1}{2} \left(\frac{v_i}{\xi_m^{(k)}} \right)^2,$$

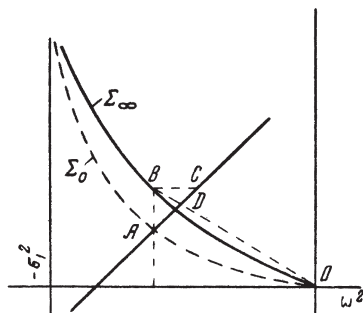
если только $n < k$. Поэтому кривая $\Phi_2^{(n)}(\xi)$ пройдет ниже кривой $\Phi_2^{(k)}(\xi)$ (см. фиг. 5), и, следовательно, точка пересечения с прямой $\Phi_1(\xi)$ сдвигается вправо, т. е. $\xi_m^{(n)} < \xi_m^{(k)}$.

Таким образом, заменяя исследование характеристического уравнения бесконечной системы (6.1) исследованием характеристического уравнения конечной системы и вычисляя корни характеристического уравнения этой системы, мы всегда получаем приближенные значения корней характеристического уравнения (6.2).

Доказанное предложение помогает оценить интервал, в котором содержится искомое значение корня.

Для определения границ изменения корня ω_0^2 можно использовать следующее построение.

Найдем наибольший корень квадратного уравнения (6.4), который равен абсциссе точки пересечения A . Обозначим его через ω_0^{*2} и вычислим $\Phi_2(\omega_0^{*2})$; получим точку B , ордината которой, по доказанному, больше ординаты точки A . Соединим точку B с началом координат. Точку пересечения этого отрезка с прямой $\Phi_1(\omega^2)$ обозначим через D . Как это видно из фиг. 6, искомое значение корня будет лежать между абсциссами точек B и D . Вычисление $\Phi_2(\omega_0^{*2})$ не представляет



Фиг. 6

труда ввиду быстрой сходимости данного ряда. Следует заметить, что значение абсциссы точки D ближе к значению корня, чем значение второго приближения, соответствующего абсциссе точки C .

4. Для вычисления собственных частот больших индексов может иметь значение следующая теорема, характеризующая поведение ω_n^2 при $n \rightarrow \infty$.

Теорема. Величина ω_n^2 может быть представлена выражением $|\omega_n^2| \sim g^{-n} [k_n^2]$, где $k_n \rightarrow 0$ при $n \rightarrow \infty$.

Функция $\Phi_2(\omega^2)$ — мероморфная функция, простые полюсы которой находятся в точках ω_n^2 . Следовательно, в окрестности этих точек ее разложение имеет вид

$$\Phi_2(\omega^2) = \frac{k_1}{\omega^2 - \omega_n^2} + k_0 + k_1(\omega^2 - \omega_n^2) + \dots,$$

где $k_1 = -n \binom{2}{n} v_n^2$.

Обратная функция имеет в этой точке простой нуль; следовательно, ее разложение будет иметь вид

$$\binom{2}{n} \frac{l_1}{\Phi_2(\binom{2}{n})}$$
 члены высшего порядка,

где $l_1 = k_1$. Таким образом,

$$n \frac{l_1}{\Phi_2(\binom{2}{n})}$$
 члены высшего порядка.

Так как

$$\binom{2}{n-1} \Phi_2(\binom{2}{n}) = \binom{2}{n},$$

то с точностью до малых высшего порядка получаем:

$$\frac{n \binom{2}{n} g_n v_n^2}{\binom{2}{n-1} g} = n \frac{n \binom{2}{n} g_n v_n^2}{\binom{2}{n} g}.$$

Полученное неравенство полностью доказывает теорему.

Таким образом, мы нашли характеристику асимптотического поведения собственных частот колебаний сосуда с жидкостью.

5. До сих пор мы изучали колебания сосуда около положения равновесия только под действием консервативных внешних сил. Рассмотрим теперь общий случай, предполагая также наличие внешней силы, пропорциональной первой степени обобщенной скорости сосуда. Жидкость по-прежнему будем считать идеальной. Уравнение частот в этом случае будет иметь вид

$$a = n \frac{n \binom{2}{n} v_n^2}{\binom{2}{n}}, \quad (6.7)$$

где a — некоторое положительное число. Исследуем зависимость знаков действительных частей корней уравнения (6.7) от распределения корней уравнения (6.2). Вместо уравнения (6.7) рассмотрим уравнение

$$f(\binom{2}{n}) = n \frac{n \binom{2}{n} v_n^2}{\binom{2}{n}} = a, \quad (6.8)$$

которое переходит в уравнение (6.7) при $\mu = 1$.

Заметим, что

$$\frac{df}{dz^2} = 0,$$

где z_n — корень уравнения $f(z^2) = 0$, удовлетворяющий неравенству

$$z_{n-1}^2 < z_n^2 < z_{n+1}^2.$$

В самом деле, в интервале (z_{n-1}^2, z_n^2) функция $f(z^2) = \Phi_1 - \Phi_2$ один раз обращается в нуль, т. е. прямая $\Phi_1(z^2)$ один раз пересекает кривую $\Phi_2(z^2)$, которая на этом участке изменяется от $+$ до $-$.

Как нами было показано выше, при $\mu = 0$ корни уравнения (6.8) либо все лежат на мнимой оси, либо, если не выполнено условие устойчивости, некоторые из них находятся в правой полуплоскости. При изменении μ от 0 до 1 корни уравнения (6.8) будут изменять свое положение на плоскости. Пусть сначала μ — некоторое сколько угодно малое положительное число, тогда с точностью до малых первого порядка

$$z^*(\mu) = z^*(0) + \mu z^*,$$

где $z^* = \frac{dz^*}{d\mu}$, z^* — некоторый корень уравнения (6.8) при $\mu = 0$.

Поэтому с той же точностью

$$2 z^* z^* f'(z^{*2}) = a z^*,$$

т. е. для малых значений μ все корни сдвигаются влево. Сделаем теперь в уравнении (6.8) замену переменного:

$$\frac{1}{z},$$

после чего получим:

$$\frac{1}{z^2} \left(1 - 2 z^2 \sum_n \frac{n(n-1) v_n z^2}{(1 - \frac{2}{n} z^2)^2} \right) = a z = 0. \quad (6.9)$$

Каждому корню z^* уравнения (6.8) соответствует корень z^* уравнения (6.9) и обратно. Кроме того, знаки z^* и z^* будут одинаковы, так как

$$z^* = \frac{1}{z^{*2}} z^*,$$

где z^* — чисто мнимая величина.

Исключая из рассмотрения случай безразличного равновесия, т. е. случай, когда

$$\frac{v_n^2}{n g} = 0,$$

заметим, что для конечных значений v_n корни уравнения (6.9) будут лежать в круге конечного радиуса. Поэтому чисто мнимые корни уравнения (6.9) при $v_n = 0$ при изменении v_n сначала сдвигаются влево, при дальнейшем же его изменении могут перейти в правую полуплоскость, только пересекая мнимую ось. Но последнего также не может быть, так как уравнение (6.9) ни при каких отличных от нуля значениях v_n не может иметь чисто мнимых корней. Очевидно, в силу тех же соображений, что положительные корни уравнения (6.9) (если они имеются) при любом конечном v_n остаются в правой полуплоскости. Таким образом, для данного случая нами в полном объеме доказана теорема, аналогичная известной теореме Кельвина:

Теорема. 1. *Неустойчивое при потенциальных силах положение равновесия остается неустойчивым при добавлении диссипативных сил.*

2. *Устойчивое при потенциальных силах положение равновесия становится асимптотически устойчивым при добавлении диссипативных сил с полной диссипацией (так как каждое из главных колебаний будет затухающим).*

Полная диссипация в нашей задаче будет означать, что все коэффициенты разложения Фурье v_n , v_n отличны от нуля, так как в противном случае затухать будут волны только тех индексов n , для которых v_n или v_n отличны от нуля.

В заключение представим себе характер колебания сосуда при наличии диссипативной силы, пропорциональной первой степени обобщенной скорости сосуда.

Пусть в начальный момент сосуд находится в покое, а в жидкости были возбуждены волны. В дальнейшем энергия волн заставит колебаться весь сосуд и поэтому будет рассеиваться внешней диссипативной силой. Так как начальную энергию мы должны считать конечной, то с течением времени сосуд будет стремиться к состоянию покоя. Волны первого типа, чья энергия была передана сосуду, также будут успокаиваться. Волны второго типа, соответствующие тем индексам, для которых v_n и v_n равны нулю (нетрудно показать, что эти волны характеризуются тем, что не изменяют положения центра тяжести сосуда), не зависят от движения сосуда; поэтому с течением времени они не изменятся и сохранят свою интенсивность.

§ 7. Движение твердого тела, имеющего полость в форме прямоугольного параллелепипеда

В качестве примера рассмотрим случай параллелепипедальной полости.

1. *Определение системы собственных функций.* Будем предполагать, что в состоянии покоя плоскость свободной поверхности жидкости перпендикулярна боковой поверхности параллелепипеда. Примем эту плоскость в качестве плоскости O_1 . Начало отсчета O_1 поместим в центр симметрии невозмущенной свободной поверхности. Собственные функции мы можем разыскать методом Фурье. Обозначим через $2l$ и $2d$ длину и ширину полости, через h — ее глубину. Граничные условия на стенках, которым должен удовлетворять потенциал скоростей, в этом случае будут следующими:

$$\begin{aligned}
 \text{при } x = l & \quad \varphi = 0, \\
 \text{при } x = -l & \quad \varphi = 0, \\
 \text{при } y = d & \quad \varphi = 0, \\
 \text{при } y = -d & \quad \varphi = 0, \\
 \text{при } z = h & \quad \varphi = 0.
 \end{aligned} \tag{7.1}$$

Поэтому система функций $\varphi_n(x, y, z) = \varphi_n(x, y) \varphi_n(z)$, нормированная условием

$$\int_{-l}^l \int_{-d}^d \int_0^h \varphi_n^2(x, y, z) dx dy dz = 1,$$

будет иметь вид

$$F_{nm}(x, y, z) = \frac{\sin \frac{n}{2l} x \sin \frac{m}{2d} y \operatorname{ch} \frac{h}{2} \sqrt{\frac{n^2}{l^2} + \frac{m^2}{d^2}}}{\sqrt{ld} \operatorname{ch} \frac{h}{2} \sqrt{\frac{n^2}{l^2} + \frac{m^2}{d^2}}}, \quad (7.2)$$

где \sin соответствует нечетным, а \cos — четным значениям индексов n и m . Следовательно, системой собственных функций будет система функций $\{ F_{nm}(x, y, z) \}$, определяемых равенством

$$F_{nm} = \frac{1}{\sqrt{ld}} \frac{\sin \frac{n}{2l} x \sin \frac{m}{2d} y}{\cos 2l \cos 2d}. \quad (7.3)$$

Собственные числа λ_{nm} могут быть найдены из условия

$$\lambda_{nm} = \frac{d}{d} \left(\frac{\lambda_{nm}}{\lambda_0} \right),$$

т. е.

$$\lambda_{nm} = \frac{1}{2} \sqrt{\frac{n^2}{l^2} + \frac{m^2}{d^2}} \operatorname{th} \frac{h}{2} \sqrt{\frac{n^2}{l^2} + \frac{m^2}{d^2}}. \quad (7.4)$$

2. Перейдем теперь к определению потенциала Жуковского:

~

Функции — гармонические внутри объема и удовлетворяют граничным условиям (2.5).

Определим сначала функцию . Граничные условия, которым она удовлетворяет, имеют вид

при $d = 0$ — (0),

при 0 и $h = 0$.

Следуя Жуковскому, будем искать решение в виде

$$F^{(0)}(x, y, z) = C_0(x_0)(y_0) C,$$

где функция $F^{(\cdot)}$ — гармоническая внутри полости и удовлетворяет на ее стенках граничным условиям

$$\begin{aligned} \text{при } d &= \frac{F^{(\cdot)}}{0}, \\ \text{при } 0 \text{ и } h &= \frac{F^{(\cdot)}}{2(\cdot_0)}. \end{aligned} \tag{7.5}$$

Положим

$$F^{(\cdot)} = \sum_{n=1} C_n^{(\cdot)} \frac{\sin \frac{n}{2d}}{\cos \frac{n}{2d}} \operatorname{sh} \frac{n}{2d} \frac{h}{2} \cdot 2 \cdot_0, \tag{7.6}$$

где \sin по-прежнему соответствует нечетным, а \cos — четным значениям n . Функция, определенная рядом (7.6), является гармонической

и удовлетворяет условию $\frac{F^{(\cdot)}}{0}$ при d .

Выберем коэффициенты $C_n^{(\cdot)}$ так, чтобы ряд (7.6) удовлетворял второму из граничных условий (7.5). Так как в рассматриваемой области имеет место разложение

$$2 \sum_{n=0} a_{2n+1} \sin \frac{2n+1}{2d}, \tag{7.7}$$

где коэффициенты Фурье определяются формулами

$$a_{2n+1} = (-1)^n \frac{16d}{(2n+1)^2 \cdot 2}, \tag{7.7}$$

то, дифференцируя ряд (7.6) и сравнивая коэффициенты в разложениях (7.6) и (7.7), мы приходим к следующим формулам для определения величин $C_n^{(\cdot)}$:

$$\begin{aligned} C_{2n}^{(\cdot)} &= 0, \\ C_{2n+1}^{(\cdot)} &= (-1)^n \frac{32d^2}{(2n+1)^3 \operatorname{ch} \frac{2n+1}{2d} \frac{h}{2}}, \end{aligned} \tag{7.7}$$

и, следовательно, искомая функция имеет вид

$$F^{(\cdot)} = 2 \cdot_0 \sum_{n=0} (-1)^n \left(\frac{h}{2} \right) C_n,$$

где

$$C_{2n-1}^{(0)} \sin \frac{2n-1}{2d} \operatorname{sh} \frac{2n-1}{2d} = \frac{h}{2}.$$

Постоянную C определим из условия несжимаемости:

$$\int_0^d (\sigma, 0) d = 0.$$

Замечая, что $\int_0^d \sigma d = 0$, мы найдем:

$$C = 0.$$

Следовательно, получим окончательно:

$$\sigma = 0, \quad \tau = 0. \quad (7.8)$$

Аналогично определим функцию τ :

$$\tau = 0. \quad (7.8)$$

где

$$C_{2n-1}^{(0)} \sin \frac{2n-1}{2l} \operatorname{sh} \frac{2n-1}{2l} = \frac{h}{2},$$

$$C_{2n-1}^{(0)} = (-1)^n \frac{32l^2}{(2n-1)^3 \operatorname{ch} \frac{2n-1}{2l} \frac{h}{2}}.$$

Точно так же для функции τ легко получим ее выражение:

$$\tau_{n,m} = \frac{(-1)^{n-m} 16}{\frac{2n-1}{2l} \frac{2m-1}{2d}} \frac{l}{d(2n-1)^2} \frac{d}{l(2m-1)^2} \sin \frac{2n-1}{2l} \sin \frac{2m-1}{2d} = 0. \quad (7.10)$$

3. Найдем далее разложения в ряды Фурье функций σ , τ , ϵ и η :

$$a) \quad \rho_{n, 2n-1, 2n-1, 0},$$

$$\text{где } \rho_{n, 2n-1} = \frac{(1)^n 8d^{\frac{3}{2}}}{(2n-1)^2 \cdot 2};$$

$$б) \quad \rho_{n, 2n-1, 0, 2n-1},$$

$$\text{где } \rho_{n, 2n-1} = \frac{(1)^n 8d^{\frac{3}{2}}}{(2n-1)^2 \cdot 2};$$

$$в) \quad (\rho_{n, 0}, 0)_{n, 0, 2n-1, 0, 2n-1}, \quad (7.11)$$

$$\text{где } \rho_{n, 2n-1} = \frac{32d^{\frac{5}{2}}}{2(2n-1)^3} \operatorname{th} \frac{2n-1}{2d} \frac{h}{2} (1)^n \frac{8d^{\frac{3}{2}}}{(2n-1)^2 \cdot 2};$$

$$г) \quad (\rho_{n, 0}, 0)_{n, 0, 2n-1, 2n-1, 0},$$

$$\text{где } \rho_{n, 2n-1} = \frac{32l^{\frac{5}{2}}}{2(2n-1)^3} \operatorname{th} \frac{2n-1}{2l} \frac{h}{2} (1)^n \frac{8l^{\frac{3}{2}}}{(2n-1)^2 \cdot 2};$$

$$д) \quad (\rho_{n, m, 0})_{n, m, 0, 2n-1, 2m-1, 2n-1, 2m-1},$$

$$\text{где } \rho_{n, 2n-1, 2m-1} = \frac{(1)^{n+m} 16}{4 \frac{2n-1}{2l} \frac{2m-1}{2d}} \frac{l^{\frac{3}{2}}}{d(2n-1)^2} \frac{d^{\frac{3}{2}}}{l(2m-1)^2}$$

$$\frac{8}{2} \frac{(1)^n l^{\frac{3}{2}}}{(2n-1)^2} \frac{(1)^m d^{\frac{3}{2}}}{(2m-1)^2}.$$

Примем для простоты, что оси инерции сосуда совпадают с осями системы x, y, z , т. е. что матрица $\| \delta_{ij} \|$ — единичная. Тогда, подставляя полученные разложения в систему (4.3), мы получим

Как показал Н.Е. Жуковский [1], для любых цилиндрических полостей с образующей вдоль оси O_1 имеют место равенства

$$[\quad] [\quad] = 0.$$

Так как в нашем случае в качестве образующей может быть принято любое из ребер параллелепипеда, будем иметь, кроме того:

$$[\quad] = 0.$$

Поэтому A, B, C — моменты инерции эквивалентного твердого тела — могут быть вычислены по формулам

$$A = A_0 [\quad], \quad B = B_0 [\quad], \quad C = C_0 [\quad],$$

где A_0, B_0, C_0 — моменты инерции сосуда.

Система уравнений (7.12)—(7.14) описывает движение сосуда параллелепипедальной формы при заданных \bar{L} и Q как функции координат сосуда и времени.

5. В качестве примера рассмотрим задачу о плоских колебаниях прямоугольного сосуда под воздействием упругой силы (ср. [4] и [5]).

Возвращающую силу пружины положим равной

$$Q_x = Mk^2 X.$$

Уравнения движения в нашем случае примут вид

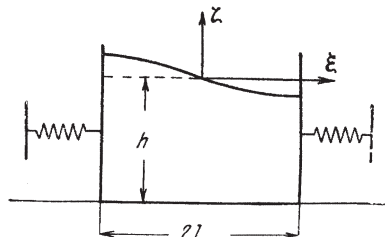
$$MX + \sum_{n=0}^{2n-1} q_{2n-1} k^2 MX = 0, \tag{7.15}$$

$$q_{2n-1} = \frac{2}{2n-1} q_{2n-1} X + \sum_{n=1}^{2n-1} q_{2n-1} = 0,$$

где

$$q_{2n-1} = (1)^n \frac{8l^{\frac{3}{2}}}{(2n-1)^2 \cdot 2},$$

$$q_n = g_n - g \frac{2n-1}{2l} \operatorname{th} \frac{2n-1}{2l} h.$$



Фиг. 7

Составляя уравнение частот, получим:

$$k^2 = \frac{32l^2}{M(2n-1)^3} \operatorname{th} \frac{2n-1}{2l} h \frac{h^4}{n^2}.$$

Так как

$$M = 2lh \frac{m_1}{m_2},$$

то это уравнение окончательно представим в виде

$$k^2 = \frac{16l \operatorname{th} \frac{2n-1}{2l} h}{(2n-1)^3 h \left(\frac{h^2}{n^2} - 1 \right) \frac{m_1}{m_2}}. \quad (7.16)$$

Из предыдущего нам известно, что корни уравнений (7.16) — чисто мнимые. Однако это легко показать и непосредственно. В самом деле, для того чтобы справа от оси $x^2 = 0$ (см. фиг. 4) не было точек пересечения кривых $\Phi_1(x^2)$ и $\Phi_2(x^2)$ достаточно, чтобы для любого $x^2 > 0$ было выполнено неравенство

$$S = \frac{16l \operatorname{th} \frac{2n-1}{2l} h}{(2n-1)^3 h \left(\frac{h^2}{n^2} - 1 \right) \frac{m_1}{m_2}} > 1.$$

Но последнее всегда имеет место. Действительно, так как $\frac{\operatorname{th} x}{x} > 1$, то

$$S = \frac{8}{n^2} \frac{1}{(2n-1)^2} > \frac{8}{n^2} \frac{1}{n^2} \frac{1}{(2n)^2} = \frac{8}{n^2} \frac{1}{6} \frac{1}{4} \frac{1}{6} > 1.$$

6. Пользуясь теоремой, доказанной в предыдущем параграфе, выясним поведение величины $\left| \frac{h^2}{2n-1} - \frac{h^2}{2n+1} \right|$ больших n . По доказанному, $\left| \frac{h^2}{2n-1} - \frac{h^2}{2n+1} \right| \sim \frac{h^2}{2n-1} \frac{h^2}{2n+1}$.

Следовательно,

$$\left| \frac{\omega_0^2}{2n-1} - \frac{\omega_0^2}{2n+1} \right| \sim \frac{(2n-1)^2}{4l^4} - \frac{64l^3}{(2n-1)^4} \operatorname{th} \frac{2n-1}{2l} h;$$

отсюда эта величина стремится к нулю, как $\frac{1}{(2n-1)^2}$, т. е. высшие

собственные частоты колебаний сосуда с жидкостью стремятся к собственным частотам колебаний жидкости в неподвижном сосуде.

7. Значение корней уравнения частот зависит только от трех параметров: характеристики возвращающей силы k^2 , отношения

глубины жидкости к ширине сосуда $\frac{h}{l}$ и отношения массы сосуда и

массы жидкости $\frac{m_1}{m_2}$. Характер этой зависимости показывает

прилагаемый график. Все рас-

четы приведены для $k^2 = 1$.

Каждой собственной частоте ω_{2n-1} можно поставить в соответствие значение некоторой приведенной массы при помощи равенства

$$\tilde{M}_i = \frac{k^2 M}{\left| \frac{\omega_i^2}{i} \right|},$$

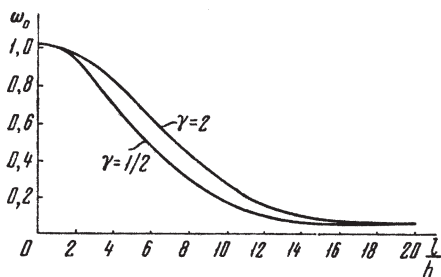
где M — масса системы со-

суд — жидкость. В данном

примере отношение $\frac{\tilde{M}_0}{M}$ меняется от 1 при $\frac{l}{h} = 0$ до $\frac{1}{2}$ при $\frac{l}{h} = \infty$,

что соответствует тому факту,

что с увеличением l — ширины сосуда — влияние волнового движения жидкости на движение сосуда увеличивается (увеличивается доля энергии волн в общей суммарной энергии системы сосуд — жидкость).



Фиг. 8. Зависимость первой собственной частоты ω_0 от $\frac{l}{h}$

Литература

1. Жуковский Н.Е. Движение твердого тела, имеющего полости, наполненные однородной капельной жидкостью. Избранные сочинения. Т. I, с. 31–153. — М.-Л., 1948.

2. Моисеев Н.Н. Движение твердого тела, имеющего полости, частично наполненные идеальной жидкостью./ДАН СССР, т. LXXXV, № 4 (1952), 719–722.

3. Моисеев Н.Н. О колебаниях тяжелой идеальной и несжимаемой жидкости в сосуде./ДАН СССР, т. LXXXV, № 5 (1952), 963–96.

4. Моисеев Н.Н. Задача о малых колебаниях открытого сосуда с жидкостью под действием упругой силы./Укр. матем. журнал, т. IV, № 2 (1952), 168–173.

5. Сретенский Л.Н. Колебания жидкости в подвижном сосуде./Изв. АН СССР, ОТН, № 10 (1951), 1483–1494.

6. Канторович Л.В., Крылов В.И. Приближенные методы высшего анализа. — М.-Л., 1949.

1952 г.

К теории нелинейных колебаний ограниченного объема жидкости

Работа посвящена расчету нелинейных колебаний ограниченного объема жидкости. Однако работа имеет ряд недостатков, вследствие которых автор долгое время считал нецелесообразным ее опубликование.

Так, следует отметить, что результаты получены формальным путем, сходимость процесса не доказана, вычисления, связанные с реализацией процесса, очень трудоемки.

Возможность распространения этих результатов на случай колебаний тела с жидкостью, имеющей свободную границу, была неясна, так как отсутствовала теорема о полноте главных колебаний такого тела.

Так как при нелинейных колебаниях амплитуда очень быстро достигает своего предельного значения и волны разрушаются, то казалось, что практически нужна либо линейная теория, либо теория, учитывающая диссипацию энергии при разрушении волн.

Обоснование процесса не удалось получить и к данному времени. Тем не менее следующие обстоятельства, как нам кажется, оправдывают теперь публикацию теории.

1. Применение быстродействующих машин позволяет сейчас без особого труда проводить требуемые вычисления, тем более что в последние годы разработаны эффективные методы численного решения соответствующих краевых задач.

2. Вопрос о полноте главных колебаний тела с жидкостью вполне выяснен, распространение развиваемой теории на случай колебания тела с жидкостью, имеющей свободную границу, не встречает принципиальных трудностей.

3. Исследование резонансных явлений в жидкости принципиально невозможно провести в рамках линейной теории, а этот вопрос в настоящее время начал интересовать практиков.

§ 1. Свободные колебания жидкости

1°. Задача сводится к определению функции ζ , гармонической в объеме τ (см. фиг. 1, где приведены обозначения), ограниченном неподвижной поверхностью Σ и свободной границей $z = \zeta(x, y, t)$ и удовлетворяющей условиям

$$\frac{\partial \zeta}{\partial n} = 0 \text{ на } \Sigma, \quad (1.1)$$

$$\frac{\partial \zeta}{\partial t} + g \frac{1}{2} (\zeta)^2 = 0 \text{ для } z = \zeta, \quad (1.2)$$

где функция ζ определяется из кинематического соотношения

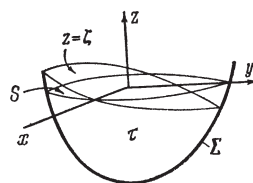
$$\frac{d}{dt} \zeta = \frac{\partial \zeta}{\partial z} \zeta. \quad (1.3)$$

Обозначим через λ_n и $\mu_n(x, y)$ собственные числа и функции интегрального уравнения

$$\mu_n(x, y) = \int_{\Sigma} H(x, y, 0; x', y', 0) \mu_n(x', y') dS, \quad (1.4)$$

где H — функция Грина задачи Неймана для объема τ , ограниченного поверхностью Σ и плоскостью $z = 0$.

Имеет место следующий результат [4]. Пусть ϕ_n^* и ψ_n^* — потенциал скоростей и свободная граница n -го свободного колебания жидкости бесконечно малой амплитуды и пусть



$\phi_n^* = \mu_n(x, y) \sin \lambda_n t.$

Тогда

$$\psi_n^* = \mu_n \Phi_n(x, y, z) \cos \lambda_n t. \quad (1.5)$$

Фиг. 1 Здесь

$$\Phi_n(x, y, z) = \int_{\Sigma} H(x, y, z; x', y', 0) \mu_n(x', y') dS,$$

т. е.

$$\Phi_n(x, y, 0) = \frac{\mu_n(x, y)}{\lambda_n}.$$

Частоты собственных колебаний ω_n и числа ν_n связаны формулой

$$\omega_n^2 = g \nu_n. \quad (1.6)$$

2°. Пусть α — некоторый параметр; положим

$$\nu_n = \alpha^n, \quad \omega_n = \alpha^n. \quad (1.7)$$

Введем новое независимое переменное

$$t = \frac{t}{h_n \alpha^n}, \quad (1.8)$$

где h_n — подлежащие определению постоянные.

Подставляя ряды (1.7) в условия (1.2) и (1.3) и переходя к новому переменному (1.8), получим следующую систему уравнений для определения неизвестных функций ν_n и ω_n :

$$\frac{1}{m} g_0 = 0, \quad (1.9)$$

$$\frac{1}{m} g_1 = A_1(\alpha, \alpha) \frac{h_1}{m} g_0,$$

$$\frac{1}{m} g_2 = A_2(\alpha, \alpha, \alpha, \alpha, h_1) \frac{h_2}{m} g_0,$$

$$\frac{1}{m} \frac{1}{z} = 0, \quad (1.10)$$

$$\frac{1}{m} \frac{1}{z} = B_1(\alpha, \alpha, h_1).$$

Функции A_i и B_i легко могут быть вычислены. Символ $()_0$ означает, что значения функции взяты при $z = 0$.

3°. Найдем нулевое приближение. Дифференцируя частным образом по α первое из уравнений системы (1.9) и используя первое из условий (1.10), получим

$$\frac{z_0^2}{2} - \frac{1}{2} g \frac{z_0}{m} = 0. \quad (1.11)$$

Положим

$$f_{n0}(t) = f_n(x, y, z), \quad f_n(x, y, 0) = f_n(x, y).$$

Тогда

$$f_n(x, y, 0) = \int_S H(x, y, 0, x, y, 0) \frac{z_0^n}{z_0} dS,$$

т. е.

$$\frac{z_0^n}{z_0} = \int_S H(x, y, 0, x, y, 0) \frac{z_0^n}{z_0} dS. \quad (1.12)$$

Поэтому для функции f_{n0} получим уравнение

$$f_{n0} - \frac{z_0^n}{2} f_{n0} = 0.$$

В силу произвольности начала отсчета времени, единственное решение этой системы периода 2π имеет вид*:

$$f_{n0} = 0 \text{ при } n \neq m, \quad f_{m0} = c \cos \dots$$

Примем $c = ag/m$; тогда

$$f_{m0} = \frac{g}{m} \cos \dots \quad (1.13)$$

Используя первое из уравнений (1.10) и равенство (1.12), найдем форму волновой поверхности:

$$z_0 = a \cos \dots \quad (1.14)$$

Здесь a — произвольная постоянная.

4°. Найдем первое приближение. Дифференцируя второе из уравнений системы (1.9), заменяя z_0 его значением из второго

* Предполагается, что z_0^2/m не равно целому числу, если $n \neq m$.

уравнения системы (1.10) и принимая во внимание значения φ_0 и $z=0$, получим при $z=0$

$$\frac{1}{2} - \frac{1}{2} g \frac{1}{z} = \frac{h_1}{m} g a \cos m \quad a^2 \sin 2 F_1^{(2)}(x, y).$$

Здесь $F_1^{(2)}(x, y)$ — уже известная функция. Решение будем искать в виде

$$f_{n1} = f_{n1}^{(2)}$$

Если принять

$$F_1^{(2)}(x, y) = b_{1n}^{(2)},$$

то для определения f_{n1} придем к следующей системе:

$$f_{n1} = \frac{2}{m} f_{n1} - a^2 b_{1n}^{(2)} \sin 2 \quad (n \neq m), \tag{1.15}$$

$$f_{m1} = f_{m1} - \frac{h_1}{m} g a \cos m \quad a^2 b_{1m}^{(2)} \sin 2.$$

Для существования периодического решения системы (1.15) необходимо и достаточно, чтобы $h_1 \neq 0$. Амплитуда a может быть любой; не ограничивая общности, примем ее равной единице.

Если определены f_{n1} , то сразу могут быть вычислены функции f_{1n} .

5°. Для определения первой поправки на частоту h_2 надо рассмотреть второе приближение. Повторяя рассуждения, придем к следующему условию для функции f_{2n} при $z=0$:

$$\frac{1}{2} - \frac{1}{2} g \frac{1}{z} = \frac{h_2}{m} g \cos m \quad F_2^{(1)} \cos m \quad F_2^{(3)} \cos 3m.$$

Положим

$$f_{2n} = f_{2n}^{(1)} - F_2^{(1)} \cos m \quad b_{n2}^{(1)} \cos m, \quad F_2^{(3)} \cos 3m \quad b_{n2}^{(3)} \cos 3m.$$

Для функций $f_{n2}(t)$ получим уравнения

$$f_{n2} = \frac{2}{m} f_{n2} b_{n2}^{(1)} \cos \quad b_{n2}^{(3)} \cos 3 \quad (n = m),$$

$$f_{m2} = f_{m2} \frac{h_2}{m} g b_{m2}^{(1)} \cos \quad b_{m2}^{(3)} \cos 3 \quad (n = m).$$

Для существования периодических решений этой системы необходимо и достаточно, чтобы

$$h_2 = \frac{b_{m2}^{(1)} m}{g}. \quad (1.16)$$

Если равенство (1.16) имеет место, то легко находится искомое решение.

6°. Легко провести индукцию и доказать возможность вычисления любого приближения. Условие для функции f_k будет:

при k нечетном

$$\frac{2}{m} f_{nk} = \frac{1}{m} g \frac{h_k}{z} - \frac{h_k}{m} g \cos \quad m \quad F_k^{(s)}(x, y) \sin s \quad (s = 2, 4, 6, \dots),$$

при k четном

$$\frac{2}{m} f_{nk} = \frac{1}{m} g \frac{h_k}{z} - \frac{h_k}{m} g \cos \quad m \quad F_k^{(s)}(x, y) \cos s \quad (s = 3, 5, 7, \dots).$$

Примем

$$f_{nk} = f_{nk} n, \quad F_k^{(s)} = b_{nk}^{(s)} n.$$

Для функций f_{nk} получим следующую систему уравнений:

$$f_{nk} = \frac{2}{m} f_{nk} b_{nk}^{(s)} \frac{\sin s}{\cos s} \quad (n = m),$$

$$f_{mk} = f_{mk} \frac{h_k}{m} g \cos \quad b_{mk}^{(s)} \frac{\sin s}{\cos s} \quad (n = m).$$

Здесь \sin соответствует нечетным, а \cos четным значениям k (при этом индекс s принимает также соответственно четные или нечетные значения).

Отсюда следует, что если k нечетно, то необходимо $h_k = 0$, если k четно, то

$$h_k = \frac{b_{km}^{(1)}}{g}. \quad (1.17)$$

7°. Используя полученное решение, можно вычислить потенциал скоростей и все характеристики течения. В частности, уравнение свободной границы можно представить в виде

$$m(x, y) \sin \frac{m^2 t}{h_2^2} \dots \quad (1.18)$$

Таким образом, параметр m представляет собой амплитуду волны. Можно сделать также ряд общих выводов.

а) Частота является функцией амплитуды:

$$m = \frac{m}{Q(m)}. \quad (1.19)$$

Таким образом, спектр является не дискретным, а кусочно-непрерывным.

б) Периодические колебания возможны с любой амплитудой, лежащей внутри круга сходимости ряда (1.18). В этом проявляется одна из аналогий изучаемых колебаний и колебаний консервативных систем конечного числа степеней свободы.

Для более детального изучения свойств свободных колебаний необходимо конкретизировать форму полости.

В частном случае цилиндрических полостей можно показать, что:

а) нельзя указать момента t , в который свободная поверхность представляет собой плоскость;

б) не существует неподвижных узлов.

Доказательство сходимости предложенного процесса встречает ряд трудностей. В частности, реализация процесса возможна только тогда, когда n/m не есть целое число, если только $n \neq m$.

Как частный случай из этих результатов получается решение задачи Я.И. Секерж-Зеньковича [5], рассмотревшего стоячие волны в безграничной жидкости. Его теория эквивалентна теории колебаний жидкости в сосуде, имеющем форму параллелепипеда.

§ 2. Вынужденные колебания, резонансные явления

1°. Задача о вынужденных колебаниях жидкости под действием поля массовых сил отличается от рассмотренной тем, что условие (1.2) не является однородным:

$$\frac{1}{t} g \frac{1}{2} (\)^2 U(t, x, y) \text{ при } z = 0. \quad (2.1)$$

Примем для простоты, что

$$U = \frac{-\sin p t f(x, y)}{p}. \quad (2.2)$$

Поставим задачу об отыскании периодических решений этой системы периода $2/p$.

2°. При $\epsilon = 0$ задача будет описывать колебание некоторой консервативной системы, поэтому обычная квазилинейная трактовка может оказаться недостаточной. Будем рассматривать нашу систему как систему, близкую к системе Ляпунова, и разыскивать решения, которые при $\epsilon = 0$ переходят в периодические решения задачи (1.1)—(1.3).

Выше было установлено, что формально могут существовать периодические решения этой задачи, период которых зависит от амплитуды. В данном случае период T задан, — он равен периоду внешней силы $2/p$ и, следовательно, амплитуда должна быть определена из соотношения

$$T = \frac{2}{pn}, \quad (2.3)$$

где n — любое целое число.

Пользуясь формулой (1.19), можно уравнение (2.3) представить в виде

$$^2 h_2 \quad ^4 h_4 \quad \dots \quad \frac{m}{pn} p^n. \quad (2.4)$$

Мы видим, что поставленная задача заведомо не имеет единственного решения: могут существовать решения периода $2/p$, которые при $\epsilon = 0$ могут переходить в тривиальные; могут существовать решения, которые при $\epsilon = 0$ переходят в нетривиальные решения задачи, рассмотренной в предыдущем параграфе. В этом параграфе мы остановимся только на построении решений первого типа.

3°. Рассмотрим колебания вдали от резонанса. Положим

$$h_{n1} = \frac{1}{p_n}, \quad h_{n2} = \frac{1}{p_n}. \quad (2.5)$$

Повторя рассуждения предыдущего параграфа, при $z = 0$ приходим к следующим уравнениям относительно f_n и f_{n1} :

$$\frac{1}{t} g_{n1} = \frac{1}{p} \sin pt f(x, y), \quad \frac{2}{t} g_{n2} = A_1(\alpha_1, \beta_1), \dots, \quad (2.6)$$

$$\frac{1}{t} = \frac{1}{z}, \quad \frac{2}{t} = \frac{2}{z} B_1(\alpha_1, \beta_1), \dots \quad (2.7)$$

Дифференцируя первое из уравнений (2.6) и используя первое из уравнений (2.7), получим

$$\frac{2}{t^2} g_{n1} = \frac{1}{y} f(x, y) \cos pt \quad (z = 0).$$

Положим далее

$$f_{n1}(t) = f_n(x, y, z), \quad f(x, y) = c_n f_n(x, y).$$

Функции f_{n1} будут удовлетворять следующей системе уравнений:

$$f_{n1} = \frac{2}{n} f_{n1} = c_n \cos pt.$$

Следовательно,

$$f_{n1} = \frac{c_n}{p_n^2} \cos pt. \quad (2.8)$$

Вычисление последующих приближений не встречает трудностей.

4°. Рассмотрим решение вблизи резонанса. Решение в форме (2.8) теряет смысл, если $p_n \approx m$. Для того чтобы изучить характер колебаний, будем считать «расстройку» $p_n^2 - m^2$ малой:

$$p^2 = m^2 + \epsilon.$$

Тогда, пользуясь формулой (1.6), положим

$$g = \frac{p^2}{m} - \frac{p^2}{m}.$$

Можно показать, что при $\omega = 0$ не существует периодических решений задачи (2.1), имеющих вид:

$$r_1^n e^{an/b}, \quad r_1^n e^{-an/b},$$

если только a и b — целые числа и $b \neq 3a$. Поэтому положим

$$r_1^n e^{i/3 n}, \quad r_1^n e^{-i/3 n}.$$

Системы уравнений (2.6) и (2.7) будут в этом случае иметь вид ($z = 0$):

$$\begin{aligned} \frac{1}{t} \frac{p^2}{m} r_1 &= 0, \quad \frac{2}{t} \frac{p^2}{m} r_2 = A_1, \\ \frac{3}{t} \frac{p^2}{m} r_3 &= A_3, \quad \frac{f(x, y)}{p} \sin pt, \dots, \end{aligned} \quad (2.9)$$

$$\frac{1}{t} \frac{1}{z}, \quad \frac{2}{t} \frac{2}{z} = B_1, \dots \quad (2.10)$$

В этих уравнениях функции A_i и B_i определяются формулами ($z = 0$):

$$A_1 = \frac{1}{t} \frac{1}{z} \frac{1}{2} \left(\frac{1}{z} \right)^2, \quad (2.11)$$

$$A_2 = \frac{1}{t} \frac{2}{z} \frac{2}{z^2} \frac{1}{z} r_2 \frac{1}{z} \left(\frac{1}{z} \frac{1}{z} \right), \dots$$

$$B_1 = \frac{1}{x} \frac{1}{x} \frac{2}{z^2}, \quad (2.12)$$

$$B_2 = \frac{1}{x} \frac{2}{x} \frac{1}{z} \frac{2}{x} \frac{2}{z^2} \frac{2}{z^2} \frac{3}{z^3} \frac{1}{z^2}, \dots$$

Дифференцируя по t первое из уравнений (2.9) и используя первое из уравнений (2.10), получим ($z = 0$)

$$\frac{\partial^2}{\partial t^2} \frac{1}{m} \frac{\partial^2}{\partial z} \frac{1}{z} = 0. \quad (2.13)$$

Положим, как и ранее,

$$f_{n1} = f_n. \quad (2.14)$$

Для функций f_{n1} получим систему уравнений

$$f_{n1} - \frac{\partial^n}{\partial m^n} p^2 f_{n1} = 0, \quad n = 1, 2, \dots \quad (2.15)$$

Единственным периодическим решением системы (2.15), имеющим период $2\pi/p$, будет

$$f_{n1} = 0 \text{ при } n \neq m, \quad f_{m1} = M \sin pt + N \cos pt. \quad (2.16)$$

Здесь M и N — постоянные, подлежащие определению.

Пользуясь формулой (2.14) и первой из формул (2.10), находим

$$f_{n1}(x, y, z) = (M \sin pt + N \cos pt), \quad (2.17)$$

$$f_{m1}(x, y) = \frac{1}{p} (M \cos pt + N \sin pt).$$

Продифференцируем вторую формулу (2.9) частным образом по t :

$$\frac{\partial^2}{\partial t^2} \frac{1}{m} \frac{\partial^2}{\partial z} \frac{1}{z} = \frac{A_1}{t} - \frac{B_1}{t}. \quad (2.18)$$

Здесь A_1 и B_1 вычисляются соответственно по формулам (2.11) и (2.12).

Производя выкладки, найдем, что

$$\frac{A_1}{t} - \frac{B_1}{t} = {}_{11}(M, N)A_{11}(x, y)\sin 2pt - {}_{12}(M, N)A_{12}(x, y)\cos 2pt. \quad (2.19)$$

Здесь ${}_{11}$ и ${}_{12}$ суть квадратичные формы своих переменных.

Для определения потенциала ϕ_2 положим

$$\phi_2 = f_{n2} = f_n, \quad A_{ij} = A_{ij}^{(n)} = f_n.$$

Тогда относительно функций f_{n2} получим систему уравнений

$$f_{n2} = \frac{n}{m} p^2 f_{n2} \quad {}_{11}^{(n)} \sin 2pt \quad {}_{12}^{(n)} \cos 2pt, \quad n = 1, 2, \dots \quad (2.20)$$

Так как по предположению $n/m = k^2$, где k — целое число (если $n = m$), то единственным периодическим решением периода $2/p$ системы (2.19) будет

$$f_{n2} = {}_{11} d_{11}^{(n)} \sin 2pt \quad {}_{12} d_{12}^{(n)} \cos 2pt, \quad n = m,$$

$$f_{m2} = {}_{11} d_{11}^{(m)} \sin 2pt \quad {}_{12} d_{12}^{(m)} \cos 2pt \quad M_1 \sin pt \quad N_1 \cos pt. \quad (2.21)$$

Здесь

$$d_{ik}^{(n)} = \frac{{}_{ik}^{(n)}}{p^2 4 \frac{2}{m}},$$

постоянные M_1 и N_1 подлежат определению.

Повторяя указанный выше порядок вычислений, получим следующее уравнение для определения третьего приближения ($z = 0$):

$$\frac{3}{t^2} \frac{p^2}{m} \frac{3}{z} \sin pt \quad {}_{i=0}^3 a_{i,3} (x, y) M^i N^{3-i}$$

$$\cos pt \quad {}_{i=0}^3 b_{i,3} (x, y) M^i N^{3-i} \quad f(x, y)$$

$$\sin 2pt F_1(x, y, M, N, M_1, N_1) \quad \dots \quad \cos 3pt F_4(x, y, M, N, M_1, N_1).$$

Введем обозначения

$${}_3 f_{k3} \quad {}_k a_{ij} \quad {}_{ij}^{(k)} \quad {}_k b_{ij} \quad {}_{ij}^{(k)} \quad {}_k,$$

$$F_s \quad {}_k \quad {}_{sk} M \quad N \quad M_1, N_1) \quad {}_k(x, y), \quad f(x, y) \quad c_k \quad {}_k.$$

Аналогично системе (2.20) получим систему, n -е уравнение которой будет иметь вид:

$$f_{n3} = \frac{n}{m} p^2 f_{n3} \sin^3 pt \cos^3 pt e_n$$

$$\sin 2pt \cos 2pt \sin 3pt \cos 3pt \quad (2.22)$$

Для того чтобы уравнение (2.22) имело периодическое решение периода $2/p$, необходимо и достаточно, чтобы числа M и N удовлетворяли следующей системе уравнений:

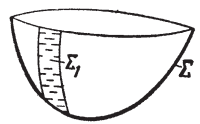
$$\sin^3 pt \cos^3 pt e_n = 0 \quad (2.23)$$

Таким образом, мы пришли к системе двух кубических уравнений, определяющих амплитуду и фазу функции φ . Функция φ будет содержать две постоянные M_1 и N_1 , которые должны быть определены из уравнения 4-го приближения.

§ 3. О задаче Ю.А. Кравченко

В работах [1—3] экспериментально и теоретически изучаются сейшеобразные колебания воды в портовых акваториях, индуцированные волнами, приходящими из открытого моря.

Предполагается, что акватория порта имеет форму цилиндра или параллелепипеда и соединена каналом с открытым морем. Предполагается, что в канале распространяются волны вполне определенных параметров. Требуется определить характер волн, возникающих на поверхности жидкости внутри порта. Эта задача отличается от задачи (1.1)—(1.3) тем, что в ней условие (1.1) будет заменено следующим (фиг. 2):



Фиг. 2

$$\frac{\partial \varphi}{\partial t} = F(P) \cos pt \quad \text{для } P = 1,$$

$$\frac{\partial \varphi}{\partial t} = 0 \quad \text{для } P = 0. \quad (3.1)$$

Эта задача может быть сведена к задаче, изученной в § 2. Для этого достаточно положить $\varphi = \varphi_1 + \varphi_2$ где φ_1 — произвольная гармоническая функция, удовлетворяющая условиям (3.1). Тогда на φ_2 будет удовлетворять условию $\frac{\partial \varphi_2}{\partial t} = 0$.

Условие постоянства давления переписывается в виде

$$\frac{1}{2} g \frac{1}{z} (\frac{1}{z})^2 F(t) \Phi(\sigma, \omega_1, \omega_2), \quad (3.2)$$

где функция Φ выражается через функции ω_1 и ω_2 .

Кинематическое условие примет вид:

$$\frac{d}{dt} \frac{1}{z} = \frac{1}{z} \frac{1}{z} \frac{1}{z}. \quad (3.3)$$

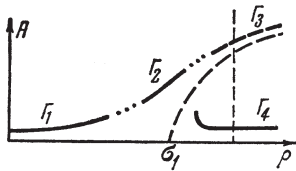
Для построения функции ω_2 и определения резонансного режима может быть применена процедура, изложенная в § 2.

В заключение сделаем несколько замечаний относительно трудностей, которые встречается построение нелинейной теории, и желательных направлений дальнейших исследований.

Автором была проделана серия опытов для проверки теории вынужденных колебаний маятников, обладавших полостью в форме цилиндра и параллелепипеда. Эти опыты количественно подтвердили ту аналогию с нелинейными колебаниями механических систем конечного числа степеней свободы, которая развивается в этой работе.

Однако при этом обнаружилось новые обстоятельства. Пусть, например, частота внешней силы p постепенно увеличивается, приближаясь слева к значению ω_1 (фиг. 3), тогда, если разность $\omega_1 - p$

не очень мала, то зависимость амплитуды A от частоты вынужденных колебаний довольно хорошо следует кривой Γ_1 . Если эта разность мала, то изменение амплитуды следует кривой Γ_2 . Если бы имела место полная аналогия с системой конечного числа степеней свободы, то изменение $A(p)$ должно было бы следовать пунктирной кривой Γ_3 . Далее этот режим был бы неустойчив и система перешла бы в режим, который описывается кривой Γ_4 . Согласно изложенному это было бы уже двухузловое колебание. В опытах этот факт не наблюдался. На режим Γ_3 система никогда не выходила: уже для $p < \omega_1$ и достаточно близких к ω_1 достигалась предельная амплитуда Стокса и волны разрушались (для высокочастотных колебаний это разрушение наступало еще раньше, когда $p < \omega_1$). Это разрушение приводит к необратимой потере энергии. Поэтому для исследования колебаний реальных систем в режимах, близких к резонансу (например, в



Фиг. 3

Фиг. 3

теории флаттера крыла, несущего баки), необходимо создать модель, учитывающую возможность разрушения волн и увеличение энтропии.

Сложность вычислений рассматриваемых задач приводит к необходимости пользоваться приближенными схемами, учитывающими, однако, существенно нелинейную природу явления. Заметим, что во многих случаях требуется учитывать только первые собственные колебания, при которых форма границы обладает малой кривизной, поэтому, по-видимому, целесообразно развить здесь методы, аналогичные вариационным методам Лаврентьева.

В нелинейной постановке эта задача теоретически почти не изучена. Основной вопрос о существовании периодических решений этой задачи остается открытым.

Литература

1. *McNown J.S.* Sur l'entretien des oscillations des eaux portuaires sous l'action de la haute mer. Publication Scientifiques et Techniques Ministere de l'Air. Paris, 1953.

2. *Aptij A.S.* Recher ches thйoriques et experimentales sur les mouvements des liquids pesants aves surface libre. Publication Scientifiques et Techniques Ministere de l'Air. Paris, 1955.

3. *Kravtchenko J. and McNown J.S.* Seiche in rectangular ports. *Quart. Appl. Math.*, 1955, № 1, 19–26.

4. *Мусеев Н.Н.* Задача о движении твердого тела, содержащего жидкие массы, имеющие свободную поверхность./*Мат. сб.*, т. 32, вып. 1, 1953.

5. *Секерж-Зенькович Я.И.* К теории стоячих волн конечной амплитуды на поверхности тяжелой жидкости./*ДАН СССР*, т. 58, вып. 4, 1947.

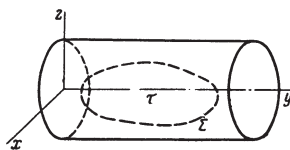
1958 г.

К теории колебаний упругих тел, имеющих жидкие полости

В современной технике встречаются разнообразные задачи, требующие исследования совместных колебаний упругого тела и жидкости. В общей постановке исследование такой задачи достаточно сложно. Здесь излагается приближенная теория, основанная на следующих упрощающих допущениях: а) линейность задачи: все смещения и скорости предполагаются бесконечно малыми, соответственно этому уравнения движения и граничные условия линеаризуются; б) схема балки: реальное упругое тело заменяется балкой с прямолинейной осью жесткости, предполагается справедливость гипотезы плоских сечений; в) жидкость — идеальная и несжимаемая, ее движение — безвихревое; г) массовая сила — сила тяжести; д) внешние силы консервативны. Здесь приводятся: вывод общих уравнений, разрешимость основных задач, анализ спектра, формулировка вариационных принципов и их обоснование.

§ 1. Плоские изгибные колебания балки, полость которой целиком заполнена жидкостью

1. Введем систему координат следующим образом (фиг. 1): ось y направим вдоль оси жесткости, оси x и z расположим в сечении, перпендикулярном оси y , так, чтобы система координат была правой. Длину балки обозначим через l , через $m(y)$ — погонную массу, через $c(y)$ — жесткость на изгиб, (y) — напряжение внешних консервативных сил. Будем рассматривать лишь тот случай, когда колебания происходят в плоскости yz . Через $Z(y, t)$ обозначим прогиб.



Фиг. 1

Пусть V — объем, который занимает жидкость, S — поверхность, его ограничивающая, $\Phi(x, y, z, t)$ — потенциал скоростей абсолютного движения жидкости. Функция Φ — гармоническая в (x, y, z, t) . На Φ она удовлетворяет условию

$$\dot{z} / n = v_n, \quad (1.1)$$

где v_n — проекция скоростей точек на внешнюю нормаль к S . В данном случае

$$v_n = \frac{z}{t} (\mathbf{z}^0 \cdot \mathbf{n}^0) = \frac{Z}{t} (x, y, z).$$

Здесь \mathbf{n}^0 — орт внешней нормали, $\cos(\mathbf{z}^0 \cdot \mathbf{n}^0)$.

Пусть $H(P, Q)$ — функция Грина задачи Неймана для области V . Введем оператор Неймана

$$Nu = H(P, Q) u(Q) ds_Q, \quad N \frac{Z}{t}. \quad (1.2)$$

Составим выражения кинетической и потенциальной энергий балки с жидкостью

$$T = \frac{1}{2} \int_0^l m Z_t^2 dy = \frac{1}{2} \int_0^l (HZ_t)^2 dy, \quad \Pi = \frac{1}{2} \int_0^l c Z_{yy}^2 dy = \frac{1}{2} \int_0^l Z^2 dy, \quad (1.3)$$

где ρ — плотность жидкости. Первое слагаемое в последнем равенстве — это потенциальная энергия упругих сил, второе — внешних сил, которые предполагаются консервативными.

2. Поставим задачу об отыскании свободных колебаний балки. Для этого сначала положим

$$Z(y; t) = \cos t \Phi(y), \quad \sin t \Phi(x, y, z), \quad \Phi = H. \quad (1.4)$$

Имея в виду применить метод Ритца, выпишем уравнение принципа Гамильтона

$$L = \int_0^t (T - \Pi) dt = 0. \quad (1.5)$$

Подставляя в это уравнение выражение функций Z и Φ согласно (1.4), а также $t = 2/\omega$ и отбрасывая несущественный множитель, приведем выражение L к виду

$$L = \int_0^l \left[\frac{1}{2} m \dot{\Phi}^2 dy - \frac{1}{2} (H \Phi)_{yy}^2 dy - \frac{1}{2} c \Phi^2 dy \right] dy. \quad (1.6)$$

Пусть теперь $\{ \varphi_n \}$ — некоторая полная и ортонормированная на $[0, l]$ система функций. Тогда согласно методу Ритца мы должны принять

$$a_n = \int_0^l \varphi_n^2 dy, \quad (1.7)$$

где a_n — числа, подлежащие определению. Подставляя ряд (1.7) в выражение функционала L , получим

$$L = \sum_{n,m} a_n a_m \left(\int_0^l \varphi_n \varphi_m dy \right) + \sum_{n,m} a_n a_m (c_{nm} - \int_0^l \varphi_n \varphi_m dy),$$

где

$$\int_0^l \varphi_n \varphi_m dy = \int_0^l c_{nm} \varphi_n \varphi_m dy, \quad \int_0^l \varphi_n \varphi_m dy = \int_0^l H_n H_m dy, \quad (1.8)$$

Таким образом, уравнение (1.5) будет приведено к системе алгебраических уравнений

$$\frac{L}{a_n} - \sum_m a_m \left(\int_0^l \varphi_n \varphi_m dy \right) - (c_{nm} - \int_0^l \varphi_n \varphi_m dy) = 0, \quad n = 1, 2, \dots \quad (1.9)$$

Характеристическое уравнение этой системы

$$\left| \sum_m a_m \left(\int_0^l \varphi_n \varphi_m dy \right) - (c_{nm} - \int_0^l \varphi_n \varphi_m dy) \right| = 0 \quad (1.10)$$

будет уравнением частот.

Коэффициенты c_{nm} могут быть названы коэффициентами присоединенных масс жидкости относительно системы функций φ_n . Таким образом, инерционные свойства жидкости определяются симметричной матрицей бесконечного порядка. Из того факта, что в общем случае $c_{nm} \neq 0$, если $n \neq m$, следует, что наличие жидкости внутри полости изменяет не только собственные частоты, но и формы главных колебаний.

3. В предыдущем разделе была изложена формальная схема применения метода Ритца и было установлено, что задача гидромеханики может быть решена независимо от решения задачи динамики системы. (Функции H_n зависят только от геометрии полости и выбора функций φ_n , они не зависят от движения балки.)

При реализации изложенной выше схемы вычислений встречаются два вопроса. Во-первых, как разумно определять систему функций u_n , и, во-вторых, как эффективно строить функции H_n .

Что касается второго вопроса, то очень трудно дать какие-либо общие рекомендации. Говоря о первом вопросе, можно во многих случаях рекомендовать выбирать в качестве функций $\{u_n\}$ собственные функции оператора

$$Lu = [cu_{yy}]_{yy} \quad u = 2mu, \tag{1.11}$$

удовлетворяющие таким граничным условиям, которые соответствуют условиям закрепления балки. Если, например, оба конца балки свободны, то

$$u(l) = u(0) = 0, \quad u'(l) = u'(0) = 0. \tag{1.12}$$

Нетрудно проверить, что в случае краевых условий (1.12) оператор L — самосопряженный и собственные функции уравнения (1.11) обладают следующими свойствами: они ортонормированны с весом $m(y)$

$$\int_0^l m(y) u_n(y) u_m(y) dy = \begin{cases} 0 & (m \neq n) \\ 1 & (m = n), \end{cases} \tag{1.13}$$

и кроме того

$$\int_0^l c(y) u_n(y) u_m(y) dy = \int_0^l u_n u_m dy = \begin{cases} 0 & (m \neq n) \\ \delta_{nm} & (m = n), \end{cases} \tag{1.14}$$

где δ_{nm} — n -е собственное число оператора L .

Эти свойства позволяют сильно упростить систему (1.9), так как в этом случае

$$c_{nm} = \delta_{nm} \delta_{nm}^2, \quad \delta_{nm} = \delta_{nm} \quad (\delta_{nm} \text{ — символ Кронекера}). \tag{1.15}$$

Характеристическое уравнение (1.10) примет тогда вид:

$$\begin{vmatrix} \delta_{11}^2 & & & & & \dots \\ & \delta_{22}^2 & & & & \dots \\ & & \delta_{33}^2 & & & \dots \\ & & & \delta_{44}^2 & & \dots \\ & & & & \delta_{55}^2 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{vmatrix} = 0. \tag{1.16}$$

Бесконечная система (1.9) и определитель (1.10) или (1.16) всегда являются сходящимися. Этот факт следует из общей теории (см. § 3).

Если балка однородная и внешние силы равномерно распределены по размаху (функции $c(y)$, $\gamma(y)$ и $m(y)$ — постоянные), то построение системы $\{L_n\}$ элементарно. Если параметры балки меняются вдоль ее размаха, то построение системы координатных функций значительно усложняется.

4. Н.Е. Жуковский показал, что в динамическом отношении твердое тело, внутри которого находится полость, целиком (без свободной поверхности) заполненная идеальной и несжимаемой жидкостью, подобно некоторому твердому телу без жидкости. Масса такого «эквивалентного» твердого тела равна суммарной массе твердого тела и жидкости, а тензор инерции определяется плотностью жидкости и геометрией полости.

Интересно выяснить, в какой степени этот факт переносится на изучаемое явление: можно ли утверждать, что в динамическом отношении балка с жидкостью эквивалентна некоторой балке без жидкости, но с другим распределением масс по сечениям.

Для того чтобы ответить на этот вопрос, составим дифференциальные уравнения движения изучаемой системы.

Составим выражение действия по Гамильтону

$$L = L_1 + L_2,$$

где

$$L_1 = \frac{1}{2} \int_0^l \{mZ_t^2 + cZ_{yy}^2 + \gamma Z^2\} dy dt, \quad L_2 = \frac{1}{2} \int_0^l (N - Z_t)^2 dy dt.$$

Вариация первого функционала вычисляется стандартным способом; она равна

$$\delta L_1 = \int_0^l \{m\delta Z_t + (cZ_{yy})_{yy} - \gamma Z\} \delta Z dy dt. \quad (1.17)$$

Вычислим вариацию второго функционала

$$\delta L_2 = \frac{1}{2} \int_0^l (N - Z_t) (-\delta Z_t) dy dt.$$

Применим формулу Грина

$$L_2 = \int_0^t \frac{1}{n} \int_{y_1}^{y_2} Z_t \, dy \, dt.$$

Так как по определению оператора \mathbf{H} имеем $\mathbf{H}u/n = u$, то выражение L_2 может быть упрощено:

$$L_2 = \int_{y_1}^{y_2} \int_0^t ZF(y,t) \, dt - \int_{l_y} F(y,t) \, dl_y = (\mathbf{H}Z_t) \, dl_y. \quad (1.18)$$

Здесь через l_y обозначен периметр сечения полости, нормального оси y , ордината которого равна y , а y_1, y_2 — граничные координаты полости. Интегрируя (1.18) по частям по t и принимая во внимание изохронность вариации, получим

$$L_2 = \int_{y_1}^{y_2} \int_0^t (\mathbf{H}Z_{tt}) \, dl_y \, Z \, dy \, dt. \quad (1.19)$$

В этом выражении интегрирование по y мы можем считать происходящим в пределах от нуля до l . Для этого достаточно положить $\mathbf{H} = 0$ для значений y , лежащих вне интервала $[y_1, y_2]$.

Согласно принципу Гамильтона

$$L_1 - L_2 = 0.$$

Подставляя в это уравнение выражения для вариаций L_1 и L_2 и используя произвол величины Z , приходим к следующему интегродифференциальному уравнению изгибных колебаний балки, внутри которой находится жидкость:

$$(\mathbf{H}Z_{tt}) \, dl_y - mZ_{tt} - (cZ_{yy})_{yy} - Z = 0. \quad (1.20)$$

Первый интеграл учитывает инерцию жидкости. Если он равен нулю, то мы имеем уравнение изгибных колебаний балки без жидкости. Оно хорошо изучено. В этом случае ускорения в каком-либо сечении однозначно определяются величинами упругой силы и внешней силы в этом сечении. Что же касается первого члена, то его величина определяется характером ускорений всех сечений балки. Другими словами, гипотеза плоских сечений заведомо не верна для жидкости (частицы жидкости перемещаются вдоль оси y). Следовательно, никакой эквивалентной балки в общем случае ввести не удастся.

5. Упрощения, которые вносит в расчеты приведение задачи о балке с жидкостью к задаче об эквивалентной обычной балке, настолько значительны, что естественно выяснить те условия, когда такая замена не приводит к большим ошибкам. Подробное изложение этого вопроса требует довольно много места. Поэтому мы приведем здесь лишь окончательный результат.

Гипотеза плоских сечений может быть использована лишь в том случае, когда длина балки велика по сравнению с ее другими размерами. Кроме того, необходимо, чтобы поверхность полости мало отличалась от цилиндрической поверхности, образующая которой параллельна оси балки (оси y). Однако и в этом случае гипотеза плоских сечений может дать более или менее точный результат только для расчета первых собственных колебаний.

Предположим теперь, что гипотеза плоских сечений справедлива. Это значит, что жидкость движется только в плоскости сечений, нормальных к оси балки. Обозначим через $\varphi_y(x, z)$ потенциал скоростей течения жидкости в сечении, ордината которого равна y , через H_y — оператор Неймана для этого сечения (плоская фигура, ограниченная контуром l_y). Тогда

$$\varphi_y = H_y Z_t.$$

Здесь $\cos(\mathbf{n}^0 | \mathbf{z}^0)$ — направляющий косинус нормали к l_y в плоскости xz . Введем понятие присоединенной погонной массы жидкости

$$m_{ж}(y) = \int_{l_y} H_y dl.$$

Тогда кинетическая энергия системы балки + жидкость может быть записана в следующем виде:

$$T = \frac{1}{2} \int_0^l \{m(y) + m_{ж}(y)\} Z_t^2 dy. \quad (2.21)$$

Это выражение показывает, что в исследуемом случае задача сводится к изучению изгибных колебаний «эквивалентной балки», обладающей той же жесткостью на изгиб, но с измененной погонной массой

$$m^* = m + m_{ж}.$$

§ 2. Изгибные колебания балки в том случае, когда находящиеся внутри нее жидкие массы имеют свободную поверхность

1. В предыдущем параграфе было установлено, что если жидкость целиком заполняет полость, то она не добавляет новых степеней свободы. Движение жидкости приводит в этом случае только к некоторому изменению величины собственных частот и форм главных колебаний по сравнению с теми, которые имела бы балка, если бы жидкость в ней отсутствовала.

Если же жидкость имеет свободную поверхность, на которой могут образовываться волны, то движение жидкости приводит к появлению дополнительного спектра собственных частот и главных колебаний.

Итак, пусть жидкость имеет свободную поверхность. Введем еще одну «неподвижную» систему координат $x_1 y_1 z_1$ так, что плоскость $x_1 y_1$ совпадает со свободной поверхностью жидкости в положении равновесия. Ось $0z_1$ направим вертикально вверх (в общем случае по направлению массовых сил) так, как это показано на чертеже (фиг. 2).

Под областью теперь будем понимать объем, ограниченный смоченной поверхностью полости (поверхность S_1) и свободной поверхностью (плоская фигура S). В силу линейности задачи потенциал скоростей абсолютного движения жидкости можно записать так:

$$(P, t) = \int_{S_1} H(P, Q) v_n(Q, t) ds_Q - \int_S H(P, Q) \frac{1}{t} v_n ds_Q$$

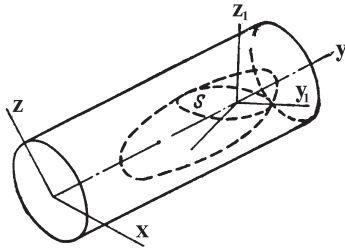
$$H v_n = H^* \frac{1}{t} v_n, \quad (2.1)$$

где $z_1 = z_1(t, x_1, y_1)$ — уравнение свободной поверхности, а скорость v_n определяется формулой $v_n = Z_t \cos(nz) - Z_t$.

Оператор H был определен. Оператор H^* определяет в некоторую функцию, нормальная производная которой на S_1 равна нулю, а на S равна $1/t v_n$.

Используя предположение о несжимаемости и гипотезу плоских сечений (из которой следует, что объем балки остается неизменным), установим, что

$$\int_S \frac{1}{t} v_n ds = 0.$$



Фиг. 2

образующихся на ее поверхности, Z — это перемещение точек плоскости S вследствие изгиба, вычисленное в координатах x_1, y_1, z_1 .

Кинетическая энергия системы балка + жидкость в исследуемом случае будет иметь вид:

$$T = \frac{1}{2} \int_0^l m Z_t^2 dy + \frac{1}{2} \int_0^l (\rho H Z_t)^2 dy + \frac{1}{2} \int_0^l (\rho H Z_t)(\rho^* U_t) dy + \frac{1}{2} \int_0^l (\rho^* U_t)^2 dy. \quad (2.2)$$

Потенциальная энергия этой системы равна сумме потенциальных энергий упругих и внешних сил и потенциальной энергии жидкости (которая определяется изменением аппликаты центра тяжести жидких масс):

$$\Pi = \frac{1}{2} \int_0^l c(Z_{yy})^2 dy + \frac{1}{2} \int_0^l Z^2 dy - g z_1 l.$$

В этой сумме вычислим последний интеграл:

$$z_1 l = \frac{1}{2} \int_S z_1^2 dx_1 dy_1 + \frac{1}{2} \int_S z_1^{*2} dx_1 dy_1.$$

Здесь $z_1 = z_1^*(x_1, y_1, t)$ уравнение поверхности в системе координат x_1, y_1, z_1 . Так как $U = Z$ то

$$z_1 l = \frac{1}{2} \int_S U^2 dx_1 dy_1 + \int_S U Z dx_1 dy_1 = J.$$

Последнее слагаемое этой суммы имеет следующую структуру:

$$J = \int_{y_1}^{y_2} Z^2(y, t) F(y) dy,$$

Отсюда следует, что функция $H^*(x_1, y_1, t, v_n)$ — гармоническая в x_1, y_1, t .

Для дальнейшего удобно ввести величину

$$U = Z.$$

U — это дополнительное перемещение жидкости за счет волн,

Z — это перемещение точек плоскости S вследствие изгиба, вычисленное в координатах x_1, y_1, z_1 .

где $F(y)$ — известная функция (которая от времени не зависит). Таким образом, это слагаемое имеет ту же структуру, что и потенциальная энергия внешних сил. Так оно и должно быть, ибо J определяет ту часть энергии силы тяжести, которая зависит только от прогиба. Поэтому величин J мы можем включить в выражение потенциальной энергии внешних сил, изменив надлежащим образом функцию $F(y)$. Таким образом, мы будем иметь

$$L = \frac{1}{2} \int_0^l c Z_{yy}^2 dy + \frac{1}{2} \int_0^l Z^2 dy - \frac{1}{2} \int_0^l g U^2 ds - g \int_0^l UZ ds. \quad (2.3)$$

2. Уравнения движения проще всего составить, если воспользоваться принципом Гамильтона

$$\delta L = 0, \quad L = L_1 + L_2, \quad (2.4)$$

где

$$L_1 = \frac{1}{2} \int_0^l m Z_t^2 dy dt - \frac{1}{2} \int_0^l (H - Z_t)^2 dt - \frac{1}{2} \int_0^l [c Z_{yy}^2 - Z^2] dy,$$

$$L_2 = \int_0^l (H - Z_t)(H^* U_t) dt - \frac{1}{2} \int_0^l (H^* U_t)^2 dt - \frac{1}{2} \int_0^l g U^2 ds dt - g \int_0^l UZ ds dt.$$

Вариация первого функционала была вычислена в предыдущем параграфе:

$$\delta L_1 = \int_0^l m Z_{tt} (c Z_{yy})_{yy} Z dy dt - \int_0^l (H - Z_t) \delta Z_t dt. \quad (2.5)$$

Вычислим вариацию второго функционала:

$$\delta L_2 = \int_0^l (H - Z_t)(H^* U_t) dt - \int_0^l (H - Z_t)(H^* U_t) dt + \int_0^l H^* U_t - H^* U_t dt - \int_0^l g U U ds dt - \int_0^l g U Z ds dt - \int_0^l g Z U ds dt. \quad (2.6)$$

Воспользуемся формулой Грина и тем, что

$$\frac{Hv}{n} = v(P, S), \quad \frac{H^* v}{n} = v(P, S).$$

Это позволяет представить выражение (2.6) в следующем виде:

$$L_2 = \int_0^t \int_S H Z_t U_t ds dt - \int_0^t \int_S H^* U_t Z_t ds dt - \int_0^t \int_S H^* U_t U_t ds dt - \int_0^t \int_S g U U ds dt - \int_0^t \int_S g U Z ds dt - \int_0^t \int_S g Z U ds dt.$$

Интегрируя по частям (по t) и пользуясь изохронностью вариаций, будем иметь

$$L_2 = \int_0^t \int_S \{H Z_{tt} - H^* U_{tt} - gU - gZ\} U ds dt - \int_0^t \int_S H^* U_{tt} Z ds dt - \int_0^t \int_S g U Z ds dt. \quad (2.7)$$

Подставляя (2.5) и (2.7) в уравнение (2.4) и используя независимость вариаций, получим следующие уравнения:

$$mZ_{tt} - H Z_{tt} dl - H^* U_{tt} dl - (cZ_{yy})_{yy} - Z - g U dl = 0, \\ H Z_{tt} - H^* U_{tt} - gU - gZ = 0. \quad (2.8)$$

Здесь через d обозначена линия пересечения фигуры S с плоскостью нормального сечения.

Таким образом, движение системы балка—жидкость в данном случае описывается системой двух интегро-дифференциальных уравнений.

3. Для определения собственных частот и форм главных колебаний воспользуемся методом Ритца. Положим в уравнении (2.4)

$$Z(y, t) = \cos \omega t \varphi(y) \quad \varphi(y) \text{ на } [0; l], \quad U(P, t) = \cos \omega t f(P) \quad P \text{ на } S. \quad (2.9)$$

Задача определения собственных колебаний сводится к минимизации следующего функционала:

$$L^* = \frac{1}{2} \int_0^l m \omega^2 dy - \frac{1}{2} \int_0^l (H \varphi'')^2 dy - \int_0^l (H \varphi'') (H^* f) dy - \frac{1}{2} \int_0^l (H^* f)''^2 dy - \frac{1}{2} \int_0^l [(c \varphi'')^2 - \omega^2] dy - \frac{1}{2} \int_S g f^2 ds - \int_S g f ds. \quad (2.10)$$

Для того чтобы решать эту вариационную задачу методом Ритца, необходимо выбрать две системы координатных функций $\varphi_n(y)$

и $\varphi_n(P)$, полных и ортонормированных на $[0, l]$ и S соответственно. Затем положим

$$a_n = \int_0^l f(x) \varphi_n(x) dx, \quad b_n = \int_S g(y) \varphi_n(y) dy.$$

После чего эта задача обычным путем сводится к системе однородных алгебраических уравнений с симметричной матрицей.

Таким образом, задача исследования свободных колебаний балки, внутри которой находятся жидкие массы, имеющие свободную поверхность, сводится к построению операторов N и N^* (для этого необходимо уметь численно решать задачу Неймана для области Ω), построению систем функций φ_n и ψ_n и составлению и решению однородной системы алгебраических уравнений.

В предыдущих параграфах был проведен формальный анализ проблемы. На примере простейшего «одномерного» движения была показана возможность редукции задачи к алгебраической. Ниже рассматривается общий случай малых колебаний балки, причем основное внимание сосредоточивается на проблеме существования главных колебаний и свойствах спектра.

§ 3. Произвольные колебания балки. Некоторые общие вопросы теории

1. Сведение к функциональному уравнению и некоторые упрощения. Введем обозначение прогиба в плоскости xOy — $X(y, t)$ и угла кручения — $Z(y, t)$. Кинетическую и потенциальную энергии балки без жидкости будем писать в виде

$$T_1 = \frac{1}{2} \int_0^l A_{11} X_t^2 + 2A_{12} X_t Z_t + 2A_{13} X_t + A_{22} Z_t^2 + 2A_{23} Z_t + A_{33} \dot{\varphi}_t^2 dy,$$

$$P_1 = \frac{1}{2} \int_0^l C_1 X_{yy}^2 + C_2 Z_{yy}^2 + C_3 \dot{\varphi}_y^2 + B_1 X^2 + B_2 Z^2 + B_3 \dot{\varphi}^2 dy. \quad (3.1)$$

Здесь C_i — жесткости на изгиб и кручение, B_i — функции, характеризующие внешние силы.

Нормальная составляющая скорости точек поверхности полости в нашем случае выражается формулой

$$v_n = X_t + Z_t + \dot{\varphi}_t,$$

где

$$x_1 = \cos(\mathbf{nx}^\circ), \quad x_2 = \cos(\mathbf{nz}^\circ), \quad x_3 = z \cos(\mathbf{nx}^\circ) - x \cos(\mathbf{nz}^\circ).$$

Введем упругое перемещение точек полости:

$$d_n = X_1 - Z_2 - x_3.$$

Аналогично тому, как это было сделано в предыдущем параграфе, введем функцию $U(P, t)$:

$$U(P, t) = (P, t) - d_n(P, t) - (P - S).$$

Повторяя затем рассуждения предыдущих параграфов, мы можем выписать выражения для энергий, функционала L и интегро-дифференциального уравнения движения

$$T = T_1 + \frac{1}{2} \int \{ H X_{t_1} - H Z_{t_2} - H_{t_3} - H^* U_t \}^2 d, \quad (3.2)$$

$$P = P_1 + \frac{1}{2} \int g U^2 ds - g \int U \{ X_1 - Z_2 - x_3 \} ds, \\ L = \int_0^t (T - P) dt, \quad (3.3)$$

$$A_{11} X_{tt} - \int_{l_y} H X_{tt} dl - A_{12} Z_{tt} - \int_{l_y} H Z_{tt} dl - A_{13} x_{tt} - \int_{l_y} H x_{tt} dl - \int_{l_y} H^* U_{tt} dl - (C_1 X_{yy})_{yy} - B_1 X - g_1 U dl = 0, \quad (3.4)$$

$$A_{12} X_{tt} - \int_{l_y} H X_{tt} dl - A_{22} Z_{tt} - \int_{l_y} H Z_{tt} dl - A_{23} x_{tt} - \int_{l_y} H x_{tt} dl - \int_{l_y} H^* U_{tt} dl - (C_2 Z_{yy})_{yy} - B_2 X - g_2 U dl = 0, \quad (3.5)$$

$$A_{13} X_{tt} - \int_{l_y} H X_{tt} dl - A_{23} Z_{tt} - \int_{l_y} H Z_{tt} dl - A_{33} x_{tt} - \int_{l_y} H x_{tt} dl - \int_{l_y} H^* U_{tt} dl - (C_3 x_y)_y - B_3 x - g_3 U dl = 0, \quad (3.6)$$

$$H X_{tt} - \int_{l_y} H X_{tt} dl - \int_{l_y} H Z_{tt} dl - \int_{l_y} H x_{tt} dl - \int_{l_y} H^* U_{tt} dl - g_1 X - g_2 Z - g_3 x - g U = 0. \quad (3.7)$$

Для исследования существования периодических решений системы (3.4)—(3.7) и изучения свойств спектра целесообразно записать эту систему в операторной форме.

Введем следующие функциональные пространства:

а) Пространства E_1 и E_2 функций $u_1(y)$ и $u_2(y)$ с суммируемым квадратом на $[0, l]$ и обладающих четвертыми обобщенными производными и со скалярным произведением

$$(u_i, v_i)_i = \int_0^l u_i v_i dy, \quad i = 1, 2. \quad (3.8)$$

б) Пространство E_3 функций $u_3(y)$ с суммируемым квадратом на $[0, l]$, обладающих вторыми обобщенными производными и со скалярным произведением типа (3.8).

в) Гильбертово пространство E_4 функций $u_4(P)$, $P \in S$, с суммируемым квадратом на S и со скалярным произведением вида:

$$(u_4, v_4)_4 = \int_S u_4 v_4 dx_1 dy_1. \quad (3.9)$$

г) Прямую сумму пространств $E = E_1 + E_2 + E_3 + E_4$ со скалярным произведением

$$(x, y) = \sum_{i=1}^4 (u_i, v_i)_i.$$

Здесь x — вектор с компонентами u_1, u_2, u_3, u_4 .

Функции u_i должны удовлетворять еще некоторым граничным условиям, которые определяются характером закрепления балки. Мы не будем детализировать эти условия. Будем считать их однородными и такими, чтобы они обеспечивали самосопряженность операторов, описывающих упругие колебания балки без жидкости.

Введем операторы L_{ij} и M_{ij} , действующие из E_j в E_i :

$$L_{ij}u_j = A_{ij}u_j + \int_{l_y}^i H_{ij}u_j dl \quad (i, j = 1, 2, 3), \quad L_{4j}u_j = H_{4j}u_j \quad (j = 1, 2, 3),$$

$$L_{i4}u_4 = \int_{l_y}^i H_{i4}^* u_4 dl \quad (i = 1, 2, 3), \quad L_{44}u_4 = H_{44}^* u_4,$$

$$M_{ii}u_i = (C_i u_{iyy})_{yy} + B_i u_i \quad (i = 1, 2), \quad M_{j4}u_4 = \int_d g_j u_4 dl \quad (j = 1, 2, 3),$$

$$M_{33}u_3 = (C_3 u_{3y})_y + B_3 u_3, \quad M_{44}u_4 = g u_4, \quad M_{4j}u_j = \int g_j u_j \quad (j = 1, 2, 3).$$

Так как H — интегральный оператор (1.2), ядро которого — функция Грина, то, для того чтобы операторы L_{ij} и M_{ij} действовали из F_j в F_i , необходимо наложить некоторые ограничения на функции A_{ij} , C_i и B_j . Мы не будем останавливаться на этих вопросах и примем раз навсегда, что всем нужным условиям эти функции удовлетворяют.

Введем еще операторы L и M следующими равенствами:

$$L \begin{pmatrix} L_{11} \cdots L_{14} \\ \cdots \cdots \cdots \\ L_{41} \cdots L_{44} \end{pmatrix}, \quad M \begin{pmatrix} M_{11} & 0 & 0 & M_{14} \\ 0 & M_{22} & 0 & M_{24} \\ 0 & 0 & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{pmatrix}.$$

В этих обозначениях система (3.4—3.7) примет вид:

$$Lx'' - Mx = 0. \quad (3.10)$$

Выясним некоторые свойства операторов L и M . Оператор L самосопряженный, т. е.

$$(Lx, y) = (x, Ly) \text{ или } \sum_{i,j=1}^4 (L_{ij}u_j, v_i)_i = \sum_{i,j=1}^4 (u_j, L_{ij}v_i)_i. \quad (3.11)$$

Справедливость такого рода равенств устанавливается простой проверкой.

Итак, оператор L — самосопряженный. Однако он не является вполне непрерывным; это обстоятельство усложняет доказательство.

Перейдем теперь к рассмотрению оператора M . Вычислим (Mx, x) :

$$\begin{aligned} (Mx, x) &= \int_0^l (C_1 u_{1,yy})_{yy} u_1 dy + 2 \int_0^l g u_{1,1} u_{4,d} dldy + \int_0^l B_1 u_1^2 dy \\ &+ \int_0^l (C_2 u_{2,yy})_{yy} u_2 dy + 2 \int_0^l g u_{2,2} u_{4,d} dldy + \int_0^l B_2 u_2^2 dy + \int_0^l (C_3 u_{3,y})_y u_3 dy \\ &+ 2 \int_0^l g u_{3,3} u_{4,d} dldy + \int_0^l B_3 u_3^2 dy + \int_S g u_4^2 ds. \end{aligned}$$

Сравнивая это выражение с равенством (3.2), имеем

$$(Mx, x) = 2\Pi. \quad (3.12)$$

Примечание. Для того чтобы убедиться в справедливости формулы (3.12), надо воспользоваться самосопряженностью граничных условий для функций u_i , $i = 1, 2, 3$.

Для того чтобы задача имела смысл, необходимо, чтобы функционал Π был положительно определенным. Вопросы об условиях, которым для этого должны удовлетворять функции C_i , i и B_i , не являются тривиальными. Однако это исследование сильно упрощается благодаря следующему факту.

Теорема. Для того чтобы функционал Π был положительно определенным, необходима и достаточна положительная определенность функционала Π^* , где

$$\begin{aligned} \Pi^* = & \int_1^2 g_{12} X Z ds + \int_1^3 g_{13} X ds + \int_2^3 g_{23} Z ds \\ & + \frac{1}{2} \int_1^2 g_{11} X^2 ds + \frac{1}{2} \int_2^3 g_{22} Z^2 ds + \frac{1}{2} \int_1^3 g_{33} ds, \end{aligned} \quad (3.13)$$

а функционал Π_1 определяется формулой (3.1).

Для того чтобы убедиться в справедливости этой теоремы, достаточно в выражении потенциальной энергии сделать замену:

$$U = v_1 X + v_2 Z + v_3. \quad (3.14)$$

После этого функционал Π примет вид:

$$\Pi = \Pi^* + \frac{1}{2} \int_1^3 g_{33} v^2 ds.$$

Эта теорема, которая является обобщением аналогичной теоремы теории движения твердого тела с жидкостью (см. [2]), носит принципиальный характер. Она показывает, что вопрос об устойчивости балки, внутри которой находятся жидкие массы, имеющие свободную поверхность, сводится к исследованию устойчивости той же балки без жидкости, но под действием другой системы внешних сил. Потенциальная энергия измененной системы сил отличается от потенциальной энергии исходной системы слагаемыми, которые однозначно определяются плотностью жидкости и геометрией полости.

Положительная определенность функционала Π^* будет, таким образом, означать, что в положении равновесия потенциальная

энергия системы балка + жидкость имеет минимум и, следовательно, это положение равновесия будет устойчивым. (Строго говоря, это утверждение еще должно быть доказано, и оно будет доказано ниже.)

В дальнейшем будем считать, что функционал Π^* , а следовательно, и оператор M — положительно определенные.

Оператор M — самосопряженный. Этот факт устанавливается непосредственной проверкой.

2. *Исследование квадратичного функционала Π^* .* Исследовать необходимые и достаточные условия для положительной определенности функционала Π^* в общем случае довольно сложно. Однако весьма просто составить только достаточные условия, при этом они будут иметь простой физический смысл.

Рассмотрим сначала простейший случай только крутильных колебаний балки.

Функционал Π^* будет иметь вид:

$$\dot{\Pi}^* = \frac{1}{2} \int_0^l \{C_3 \dot{y}^2 - B_3 \dot{b}_3^2\} dy - b_3(y) \int_d g \frac{2}{3} dl. \quad (3.15)$$

Если предположить, что C_3 нигде в нуль не обращается, то одно из достаточных условий можно выписать сразу (см., например, [4], с. 122):

$$B_3 > b_3. \quad (3.16)$$

Используя выражение $b_3 = z \cos(n, x) - x \cos(n, z)$, выражение (3.15) можно привести к виду

$$b_3(y) = g \{J_{zz}(y) \cos^2(n, z) - J_{xx} \cos^2(n, x) - 2L_{xz} \cos(n, x) \cos(n, z)\}. \quad (3.17)$$

Здесь через J_{xx} , J_{zz} , J_{xz} обозначены моменты инерции отрезка d .

Если ось балки горизонтальна, то $\mathbf{z}^\circ = \mathbf{n}^\circ$ и тогда

$$b_3(y) = gJ_{zz}. \quad (3.18)$$

Если ось балки вертикальна, то $\mathbf{y}^\circ = \mathbf{n}^\circ$ и тогда $b_3(y) = 0$, т. е. жидкость не влияет на знакоопределенность Π . Таким образом, для положительной определенности функционала Π^* (а следовательно, и для устойчивости положения равновесия) достаточно выполнения

условия, которое однозначно определяется геометрическими характеристиками свободной поверхности и плотностью жидкости.

Если балка совершает только изгибные колебания в одной из плоскостей, то подобные условия выводятся аналогично. В этом случае

$$\ddot{I}^* = \frac{1}{2} \int_0^l \{C_1 X_{yy}^2 - (B_1 - b_1) X^2\} dy - b_1(y) \int_0^l g^2 dl.$$

Так как $b_1 = \cos(n, x)$, то $b_1 = g d \cos^2(n, x)$. Для дальнейшего обозначим

$$J_{ik} = \int_0^l g_{ik} dl.$$

Критерий вида (3.16) легко получить и для общего случая. Для этого представим функционал Π^* в виде суммы:

$$\ddot{I}^* = \ddot{I}_1^* + \ddot{I}_2^*,$$

где

$$\begin{aligned} \ddot{I}_1^* &= \frac{1}{2} \int_0^l \{C_1 X_{yy}^2 - C_2 Z_{yy}^2 - C_3 \frac{\partial^2}{\partial y^2}\} dy, \\ \ddot{I}_2^* &= \frac{1}{2} \int_0^l \{B_1 - J_{11}\} X^2 dl - \frac{1}{2} \int_0^l \{B_2 - J_{22}\} Z^2 dl - \frac{1}{2} \int_0^l \{B_3 - J_{33}\} \frac{\partial^2}{\partial y^2} dl \\ &\quad - \int_0^l g_{12} X Z dl dy - \int_0^l g_{13} X \frac{\partial}{\partial y} dl dy - \int_0^l g_{23} Z \frac{\partial}{\partial y} dl dy. \end{aligned}$$

Для положительной определенности функционала Π_1^* достаточно, чтобы для любого $y \in [0, l]$ функции C_i удовлетворяли неравенствам $c_i > \epsilon_i$ где ϵ_i — произвольные сколь угодно малые положительные числа. Для положительной определенности функционала Π_2^* достаточно, например, выполнения следующих условий:

$$B_1 - J_{11} > \epsilon_1, (B_1 - J_{11})(B_2 - J_{22}) - g^2 J_{12} > \epsilon_2, \quad (3.19)$$

$$\begin{vmatrix} B_1 - J_{11} & J_{12} & J_{13} \\ J_{21} & B_2 - J_{22} & J_{23} \\ J_{31} & J_{32} & B_3 - J_{33} \end{vmatrix} > \epsilon_3, \quad (3.20)$$

где ϵ_3 — любые сколь угодно малые положительные числа.

Эти условия аналогичны неравенствам Сильвестра. Они накладывают ограничения на геометрические характеристики, и им можно придать наглядную форму. Покажем это на примере изгибных колебаний в двух взаимно-перпендикулярных плоскостях.

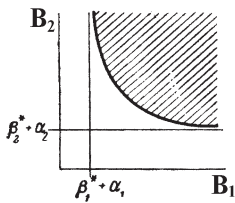
В этом случае условиям (3.20) можно придать следующий вид [1]

$$B_1 \beta_1^* \geq \alpha, \quad B_2 \beta_2^* \geq \alpha, \quad (B_1 \beta_1^*)(B_2 \beta_2^*) \geq \alpha^2 \quad (3.21)$$

Здесь

$$\beta_1^* = \frac{1}{d} \int_0^l \cos(n, x) dy, \quad \beta_2^* = \frac{1}{d} \int_0^l \cos(n, z) dy, \quad \alpha^2 = \frac{1}{d} \int_0^l g^2 dy \cos(n, x) \cos(n, z).$$

Следовательно, для выполнения этих условий величины B_1 и B_2 должны быть таковыми, чтобы для любого $y \in [0, l]$ точка (B_1, B_2) находилась в заштрихованной области плоскости $B_1 O B_2$ (фиг. 3).



Фиг. 3

Условия, которые только что обсуждались, очень грубые. Интуитивно совершенно ясно, что если жесткость положительна, то для обеспечения положительной определенности Π^* вовсе не необходимо требование, чтобы внешние силы носили восстанавливающий характер.

Рассмотрим колебания с «одной степенью свободы», например крутильные колебания. Имеем

$$\ddot{I}^* = C_3 \int_0^l \frac{1}{y^2} dy - r(y) \int_0^l \frac{1}{y^2} dy - r(y) B_3 - g \int_0^l \frac{1}{d} dy,$$

и пусть функция r — для y на $[0, l]$, где α — некоторое положительное число. Тогда

$$\ddot{I}^* = C_3 \int_0^l \frac{1}{y^2} dy - \alpha^2 \int_0^l \frac{1}{y^2} dy.$$

Оценим снизу первое слагаемое в правой части. Так как

$$\int_0^y \frac{1}{y} dy = \frac{1}{\alpha} \int_0^y \frac{1}{\sqrt{C_3}} dy,$$

то

$$\left(\int_0^l y \right)^2 \leq \int_0^l \frac{dy}{c_3} \int_0^l C_3 y^2 dy.$$

Но всегда можно считать, что

$$\int_0^l dy = 0,$$

поэтому предыдущее неравенство позволяет получить

$$\int_0^l C_3 y^2 dy \geq \left[\int_0^l y \right]^2 \int_0^l \frac{dy}{c_3},$$

$$\tilde{I}^* \geq \left[\int_0^l y \right]^2 \int_0^l \frac{dy}{c_3} \geq \left[\int_0^l y \right]^2.$$

Отсюда следует, что для положительной определенности функционала \tilde{I}^* достаточно, чтобы функция C_3 удовлетворяла неравенству

$$\frac{1}{l} \int_0^l \frac{dy}{C_3(y)} \leq 0. \quad (3.21)$$

Это условие налагает на жесткость C_3 ограничение снизу. Оно будет наверняка выполнено, если потребовать, чтобы функция C_3 удовлетворяла неравенству

$$C_3(y) \geq \mu, \quad (3.22)$$

где μ — надлежащим образом выбранное положительное число.

Условие вида (3.22) может быть получено и в случае произвольных малых колебаний. Может быть доказано следующее утверждение.

Для положительной определенности функционала \tilde{I}^* достаточно, чтобы функции C_i удовлетворяли неравенствам

$$\min C_i \geq \mu_i, \quad y \in [0, l], \quad i = 1, 2, 3, \quad (3.23)$$

где μ_i — некоторые вполне определенные положительные числа; они зависят от характера внешних сил и от геометрии жидких полостей и плотности жидкости.

3. *Доказательство существования главных колебаний.* Задача определения собственных колебаний сводится к исследованию спектра операторного уравнения

$$Lx = \lambda Mx,$$

где операторы L и M — самосопряженные, а M — положительно определенный. Оператор L не является вполне непрерывным. Это обстоятельство требует некоторых дополнительных выкладок.

В выражениях T и Π произведем замену (3.14) и составим новые уравнения движения. В простейшем случае чисто изгибных колебаний эта система имеет вид:

$$\begin{aligned} AX'' - \int_0^l H_1 X'' dl - \int_0^l H^* v'' dl - (C_1 X_{yy})_{yy} - \lambda X &= 0, \\ H_1 X'' - H^* v'' - gv &= 0. \end{aligned} \quad (3.24)$$

Здесь

$$H_1 = H - H^*, \quad H^* u = \int_S H(P, Q) u(Q) ds, \quad \lambda = \int_0^l B_1 - g^2 dl.$$

Напомним, что $H(P, Q)$ — функция Грина задачи Неймана для области Ω . Особенностью преобразованной системы (3.24) по сравнению с системой (3.3)—(3.7) является то, что в ее первых трех уравнениях не содержится членов, в которые входит v (в первые уравнения входят только производные этой функции по t).

Введем пространство $E^* = E_1 \oplus E_2 \oplus E_3$ и пусть $w^* \in E^*$.

Тогда систему уравнений движения в общем случае можно записать так:

$$Bw'' + Dw'' - Nw^* = 0, \quad H_1 \Gamma_1 w'' - H^* v'' - gv = 0. \quad (3.25)$$

Здесь

$$B = (L_{ij}^*), \quad Dv = \int_0^l \Gamma H^* v dl, \quad \Gamma = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \Gamma_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Операторы L_{ij}^* , $i, j = 1, 2, 3$ отличаются от операторов L_{ij} тем, что под знаком интеграла оператор N заменен на N_1 и интегрирование распространено по $l_y - d$.

Симметричный оператор N определяется матрицей $\| \|_{ij}$, где числа $_{ij}$, $i \ j$ определяются формулами

$$_{ij} = \int_d g_{ij} dl.$$

Нетрудно проверить, что оператор

$$\left\| \begin{array}{cc} B & D \\ N_1 \Gamma_1 & N^* \end{array} \right\|,$$

так же как и оператор L , самосопряженный. Для того чтобы в этом убедиться, достаточно проверить справедливость равенств

$$(w, w_1) = (Lx, x_1), \quad (w, w_1) = (x, Lx_1), \quad (w, E) = w \left\| \begin{array}{c} w^* \\ v \end{array} \right\|.$$

Элементы w и x связаны соотношением (3.14).

Оператор N — неограниченный, положительно определенный. Его обратный N^{-1} — вполне непрерывный интегральный оператор. Структуру этого оператора проще всего увидеть на каком-нибудь частном случае. Например, если рассматриваются чисто крутильные колебания, то уравнение $Nu = f$ имеет вид:

$$(C_3 y)_{,y} = f,$$

и, следовательно,

$$(P) = \int_0^l G(P, Q) f(Q) dQ,$$

где G — функция Грина, учитывающая соответствующие граничные условия и обладающая слабой особенностью.

Для отыскания периодических решений положим

$$w^* = w \cos n t, \quad v = v \cos n t. \tag{3.26}$$

Система (3.25) примет вид:

$$Bw = Dy - \frac{1}{2} Nw, \quad N_1 \Gamma_1 w = N^* y - \frac{1}{2} gy. \tag{3.27}$$

Сделаем здесь замену переменных $N^{1/2}w$ и $\sqrt{g}y$. Получим

$$N^{1/2}BN^{1/2} = N^{1/2}D\frac{1}{\sqrt{g}} = \frac{1}{2},$$

$$\frac{1}{\sqrt{g}}H_1\Gamma_1N^{1/2} = \frac{1}{\sqrt{g}}H^*\frac{1}{\sqrt{g}} = \frac{1}{2}. \quad (3.28)$$

Так как оператор B — ограниченный, то $N^{1/2}BN^{1/2}$ — вполне непрерывный (поскольку N^{-1} — вполне непрерывный). Операторы D , H_1 и H^* — вполне непрерывные как интегральные со слабой особенностью. Следовательно, оператор, определяющий левую часть системы (3.28), вполне непрерывный. Аналогично нетрудно проверить, что оператор

$$R = \begin{pmatrix} N^{1/2}BN^{1/2} & N^{1/2}D\frac{1}{\sqrt{g}} \\ \frac{1}{\sqrt{g}}H_1\Gamma_1N^{1/2} & \frac{1}{\sqrt{g}}H^*\frac{1}{\sqrt{g}} \end{pmatrix}$$

самосопряженный.

Таким образом, мы пришли к задаче о собственных значениях

$$Rf = \lambda f$$

для вполне непрерывного самосопряженного оператора R .

На основании известных теорем линейного анализа мы можем утверждать справедливость следующей основной теоремы.

Теорема. Если квадратичная форма Π — положительно определена, то система (3.3)—(3.7) имеет периодические решения (главные колебания) вида (3.26), где ω_n — положительные числа (собственные частоты), образующие счетную последовательность, такую, что $\omega_n \rightarrow \infty$ при $n \rightarrow \infty$.

Эта теорема содержит в себе как частный случай для данной задачи теорему Лагранжа о минимуме потенциальной энергии. Доказанная теорема означает, что если потенциальная энергия системы балка — жидкость в положении равновесия имеет минимум, то это положение равновесия устойчиво в том смысле, что все главные колебания этой системы ограничены.

4. *Доказательство полноты.* Система главных колебаний полна в E . Для того чтобы в этом убедиться, запишем основное операторное уравнение в виде

$$M^{-1}Lx = x.$$

Введем норму Фридрихса $(x, y)_F = (Mx, y)$. Для доказательства полноты в норме Фридрихса достаточно показать, что из условия $M^{-1}Lh = 0$ сразу следует $h = 0$. Рассмотрим

$$(M^{-1}Lh, h)_F = (Lh, h).$$

Скалярное произведение, стоящее справа, — это кинетическая энергия. Следовательно, если $h \neq 0$, то это выражение не может быть равным нулю. Следовательно, и $M^{-1}Lh = 0$.

Полнота системы главных колебаний позволяет разыскивать решение задачи Коши в виде ряда, составленного из главных колебаний.

На этом заканчивается доказательство существования полной системы главных колебаний изучаемой системы. Одновременно доказанные предложения обосновывают применение метода Ритца, схема которого подробно обсуждалась в первых параграфах этой работы.

Краткое резюме некоторых результатов этой статьи было опубликовано в Докладах АН СССР (см. [5]).

За ряд замечаний и советов автор выражает свою признательность А.А. Абрамову и М.А. Неймарку.

Литература

1. Жуковский Н.Е. О движении твердого тела, имеющего полости, наполненные однородной капельной жидкостью. Собр. соч., т. 2. — Гостехиздат, 1949.
2. Моисеев Н.Н. Задача о движении твердого тела, содержащего жидкие массы, имеющие свободную поверхность./Мат. сб., т. 32, вып. 1, 1953.
3. Моисеев Н.Н. Некоторые вопросы теории колебаний сосудов с жидкостью./Инж. сб., т. 19, 1953.
4. Михлин С.Г. Вариационные методы в математической физике. — ГИТТЛ, 1957.
5. Моисеев Н.Н. К теории упругих колебаний тел с жидкостью./ДАН СССР, т. 127, № 1, 1958.

1959 г.

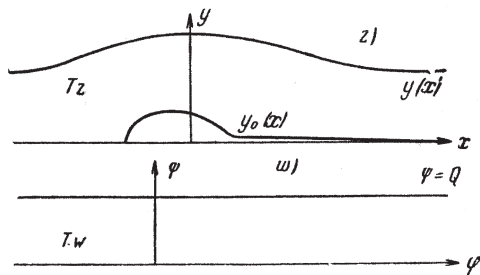
О неединственности возможных форм установившихся течений тяжелой жидкости при числах Фруда, близких к единице

Как известно, значение числа Фруда $F = 1$ является критическим. Характер установившегося потока тяжелой жидкости, скорость которой меньше критической, т. е. соответствующий $F < 1$, резко отличается от характера потока при $F > 1$. В этом смысле значение $F = 1$ является бифуркационным. Поэтому следует ожидать, что при числах Фруда, близких к единице, могут существовать различные формы установившегося потока.

Одним из примеров такой неоднозначности является известная задача об уединенной волне. Если заданы расход Q и глубина на h , причем число Фруда несколько более единицы, то возможны две формы реализации поступательного потока. Свободная поверхность либо может быть горизонтальной, либо может иметь форму уединенной волны. Точно так же следует ожидать, что обтекание потоком препятствия (или уступа) возможно не единственным образом. Так как задачи обтекания тяжелой жидкостью решались только в линейной постановке (см., например [1]), то эта неоднозначность не могла быть вскрыта. Явление неоднозначности обтекания носит существенно нелинейный характер. Кроме того, обычно применяемые методы перестают быть справедливыми при числах Фруда, близких к единице.

В настоящей работе используется теорема М.А. Лаврентьева для оценки модуля производной при отображении узких полос [2]. Они оказываются наиболее удобными как раз при числах Фруда, близких к единице. Это позволяет рассмотреть задачу в приближенной постановке, сохранив, однако, ее нелинейный характер. В работе показывается, что при числах Фруда, близких к единице, возможны во всяком случае две формы обтекания препятствия. Поскольку исследования основываются на теории приближенных конформных преобразований, они носят также приближенный характер: вопрос о теоремах существования не ставится. Попутно мы получаем новую приближенную теорию уединенной волны.

1. Рассмотрим задачу об обтекании препятствия установившимся безвихревым потоком тяжелой жидкости, имеющей свободную границу (фиг. 1). Задаче сводится к конформному отображению полосы T_z , ограниченной линией дна $y = y_0(x)$ и неизвестной свободной границей $y = y(x)$ на полосу T_w единичной ширины в плоскости w , если вдоль кривой $y = y(x)$ выполнено соотношение



Фиг. 1

$$\left| \frac{dw}{dz} \right|^2 = 2\nu c, \quad \nu = \frac{gh^3}{Q^2}. \quad (1.1)$$

Здесь потенциал w и координаты z, y взяты в безразмерном виде, h — характерный размер глубины, g — ускорение силы тяжести, Q — расход. Очевидно, что $\nu = 1/F^2$.

Если функции $y_0(x)$ и $y(x)$ удовлетворяют некоторым условиям гладкости, то для $|dw/dz|$ в точках верхней границы T_z имеет место оценка (см. [2]):

$$\left| \frac{dw}{dz} \right| = \frac{1}{y - y_0} \left(1 - \frac{y - y_0}{6} y_0' - \frac{y - y_0}{3} y'' \right). \quad (1.2)$$

Если положить $y_0 = 0$ и в равенстве (1.1) заменить значение $|dw/dz|$ оценкой (1.2) и, обозначая $y = 1 + \eta$, подобрать c так, чтобы 0 было интегралом уравнения (1.1), то из (1.2) получим

$$3(1 - \nu) = 3 + \nu \frac{1}{1 - \eta^2} f(x, \eta, \nu), \quad (1.3)$$

где

$$f(x, \eta, \nu) = 3 + y_0' \nu \frac{1}{2} + \frac{3}{2} y_0'' \nu \frac{1}{2} + y_0''' \nu \frac{1}{2}.$$

Таким образом, приближенно исследование свободной поверхности жидкости сводится к исследованию уравнения (1.3).

Любое соотношение h и Q может определять некоторый равномерный поток над гладким дном. Однако каждой комбинации h и

Q будет соответствовать свое значение постоянной c . Поэтому фиксация c означает выбор параметров равномерного потока, в окрестности которых проводится исследование.

В дальнейшем будем считать, что число ν близко к единице. Можно показать, что если разность $1 - \nu$ достаточно мала по абсолютной величине, то все условия указанной теоремы М.А. Лаврентьева выполнены.

2. Пусть вначале $\nu < 1$ (что означает, что $Q \sqrt{gh^3}$). Поставим задачу отыскания формы свободной границы, которая переходила бы в горизонтальную прямую при $\max |y_0(x)| \rightarrow 0$. Для этого положим

$$y_1 = \frac{1}{2} y_0^2, \quad y_2 = \frac{1}{2} y_0^3, \quad \dots, \quad y_n = \frac{1}{2} y_0^{n+1}. \quad (2.1)$$

Подставляя этот ряд в уравнение (1.3), найдем, что

$$y_1 = \frac{1}{2} (1 - \nu) y_1 F_1(x), \quad F_1 = \frac{1}{2} y_0. \quad (2.2)$$

Общее решение уравнения (2.2) имеет вид:

$$y_1 = \frac{1}{\sqrt{3(1-\nu)}} \exp(\sqrt{3(1-\nu)}x) \int_0^x F_1(\xi) \exp(-\sqrt{3(1-\nu)}\xi) d\xi + c_1 \exp(\sqrt{3(1-\nu)}x) + c_2 \int_0^x F_1(\xi) \exp(\sqrt{3(1-\nu)}\xi) d\xi. \quad (2.3)$$

Для упрощения дальнейших выкладок предположим*, что $F_1(\xi) \rightarrow 0$, если $|\xi| \rightarrow N$. Выберем постоянные c_1 и c_2 так, чтобы $\lim_{|x| \rightarrow N} y_1 = 0$. Очевидно, что для этого

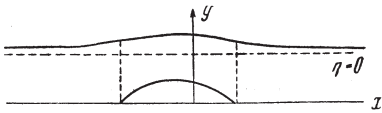
$$c_1 \int_0^N F_1(\xi) \exp(\sqrt{3(1-\nu)}\xi) d\xi + c_2 \int_0^N F_1(\xi) \exp(-\sqrt{3(1-\nu)}\xi) d\xi = 0$$

и, следовательно,

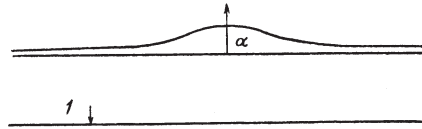
$$y_1 = \frac{1}{\sqrt{3(1-\nu)}} \exp(\sqrt{3(1-\nu)}x) \int_0^N F_1(\xi) \exp(-\sqrt{3(1-\nu)}\xi) d\xi + \exp(\sqrt{3(1-\nu)}x) \int_N^x F_1(\xi) \exp(\sqrt{3(1-\nu)}\xi) d\xi. \quad (2.4)$$

* Отказ от этого ограничения потребует наложить дополнительные ограничения на сходимость интегралов вида

$$F(\xi) \exp(\sqrt{3(1-\nu)}\xi) d\xi.$$



Фиг. 2



Фиг. 3

Форма свободной границы имеет вид, схематично указанный на чертеже (фиг. 2). При $\nu = 0$ (что соответствует Q или ν) форма свободной границы стремится к некоторому предельному значению. При $\nu = 1$ решение (2.4) перестает иметь смысл. Вычисление следующих приближений не представляет труда.

После того как свободная граница определена, потенциал скоростей можно найти, пользуясь приближенными формулами конформных отображений.

3. Уравнение (1.3) при $f = 0$ имеет также нетривиальные решения. Поэтому можно найти форму свободной границы, отличную от (2.4), которая переходит в некоторую линию, отличную от горизонтальной прямой при $\max |y_0(x)| = 0$. Для определения такого течения положим $1 - \nu$, где — некоторая малая величина порядка . Перепишем (1.3) в виде

$$3 \quad \frac{1}{2} \epsilon^2 f^*(x, \nu), \quad f^* = f - \epsilon^2. \quad (3.1)$$

Будем искать решение (2.4) в виде ряда $\epsilon^0 + \epsilon^1 + \dots$. Для ϵ^0 получаем

$$3 \quad \frac{1}{2} \epsilon^2. \quad (3.2)$$

Уравнение (3.2) приближенно описывает уединенную волну. В самом деле, оно имеет однопараметрическое семейство частных решений $\phi_0(x, x_0)$, которые при $x \rightarrow \pm\infty$ обращаются в нуль (фиг. 3):

$$\overline{\text{ch}^2 \left\{ \frac{1}{2} \sqrt{3} (x - x_0) \right\}}, \quad (3.3)$$

где x_0 может быть любым. Формула (3.3) с точностью до малых высшего порядка совпадает с известной формулой Бусинеска [3]

$$\text{sch}^2 \sqrt{\frac{3}{4(1 - \epsilon^2)}} (x - x_0). \quad (3.4)$$

Если ввести понятие скорости распространения волны $v = Q/h$, то $v = gh/v^2$. Так как $1/v$, то зависимость скорости распространения от амплитуды будет

$$v = \sqrt{\frac{gh}{1 - \frac{gh}{h^2}}}, \quad (3.5)$$

где h — амплитуда. Разлагая подкоренное выражение в ряд и удерживая первые два слагаемых, получаем снова формулу Бусинеска для скорости распространения уединенной волны:

$$v = \sqrt{g(h - \frac{1}{2}h^2)}.$$

Отметим здесь еще физический смысл h : h есть асимптотическая глубина — глубина слоя жидкости на ∞ . Волновая поверхность расположена целиком над линией $y = h$.

Для η получим такое уравнение:

$$\eta_{xx} + 3\eta_x + 9\eta = f_1^*(x, x_0, v), \quad f_1^* = \frac{1}{2}[f(x, x_0, v) - \eta^2]. \quad (3.6)$$

Рассмотрим однородную часть уравнения (3.6). Нетрудно проверить, что одно его частное решение будет

$$y_1(x - x_0) = \frac{\text{th}[\frac{1}{2}\sqrt{3}(x - x_2)]}{\text{ch}^2[\frac{1}{2}\sqrt{3}(x - x_0)]}.$$

Тогда второе частное решение можно найти, пользуясь теоремой Лиувилля:

$$y_2(x - x_0) = \frac{\text{th}[\frac{1}{2}\sqrt{3}(x - x_0)]}{\text{ch}^2[\frac{1}{2}\sqrt{3}(x - x_0)]} \int_0^x \frac{\text{ch}^4[\frac{1}{2}\sqrt{3}(x - x_0)]}{\text{th}^2[\frac{1}{2}\sqrt{2}(x - x_0)]} f_1^*(x, x_0, v) y_1(x - x_0) dx + c_1$$

Следовательно, общее решение уравнения (3.6) будет

$$\eta = \frac{1}{y_2 y_1 - y_1 y_2} \int_0^x y_1(x - x_0) f_1^*(x, x_0, v) y_2(x - x_0) dx + c_1 y_1(x - x_0) + c_2 y_2(x - x_0). \quad (3.7)$$

Так как $y_2(x)$ неограниченно при $x \rightarrow \infty$, то для существования ограниченного решения необходимо, чтобы

$$\int_{-\infty}^{\infty} f_1^*(x, x_1, \dots, x_0) y_1(x) dx = 0. \quad (3.8)$$

Величина x_0 должна считаться заданной. Она определяет при данном значении асимптотической глубины h величину расхода Q , или, что то же самое, скорость распространения волны. Поэтому уравнение (3.8) — это уравнение для определения параметра x_0 . (Следует заметить, что можно, наоборот, задать x_0 , но тогда уравнение (3.8) определит значение Q .)

Так как частное решение y_1 обращается в нуль при $x = x_0$, то решение (3.7) при любом значении постоянной c_1 , если имеет место (3.8), удовлетворяет требованию $\lim_{x \rightarrow \infty} y = 0$ при $|x| \rightarrow \infty$.

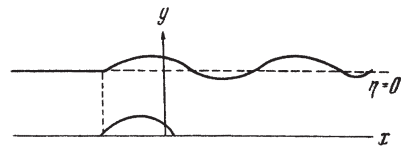
Значение c_1 должно быть определено из уравнения второго приближения.

Для определения следующих членов разложения f_k мы получаем уравнения, аналогичные (3.6), только функции f_i^* будут отличаться от f_1^* . Поэтому можно повторить изложенную выше процедуру. Уравнение (3.8) будет иметь вид:

$$\int_{-\infty}^{\infty} f_i^*(x, x_0, c_1, \dots, c_{i-1}) y_1(x) dx = 0. \quad (3.9)$$

Величины x_0, c_1, \dots, c_{i-1} будут известными, c_{i-1} надо определить из (3.9).

Таким образом, мы нашли поток, обтекающий препятствие, свободная граница которого при $\max|y_0| \rightarrow 0$ переходит в уединенную волну. При заданном расходе Q вершина этой волны вполне определенным образом расположена относительно препятствия. (Значение x_0 определено.)



Фиг. 4

Итак, если $v < 1$ и достаточно мало от нее отличается, то можно приближенно определить во всяком случае два режима обтекания. Совершенно аналогично обстоит дело, если $v > 1$.

4. Пусть теперь $v > 1$ (что означает, что $Q \sim \sqrt{gh^3}$). Найдем поток, свободная граница которого переходила бы в горизонтальную

прямую при $\max|y_0| = 0$. Для этого снова представим φ в виде ряда (2.1). Для φ_1 мы будем иметь уравнение (2.2). Так как на этот раз $\nu > 1$, то представим это уравнение в виде

$$\varphi_1 = 3(\nu - 1) F_1(x). \quad (4.1)$$

Его общим решением будет функция

$$\varphi_1(x) = \frac{1}{\sqrt{3(\nu - 1)}} \sin\sqrt{3(\nu - 1)x} \int_0^x F_1(\xi) \cos\sqrt{3(\nu - 1)} d\xi + c_0 + \cos\sqrt{3(\nu - 1)x} \int_0^x F_1(\xi) \sin\sqrt{3(\nu - 1)} d\xi + c_2.$$

Если слева на φ нет волн, т. е. если $\varphi = 0$ при $x = 0$, то значения постоянных c_1 и c_2 надо выбрать так, чтобы функция $\varphi_1(x)$ имела вид:

$$\varphi_1(x) = \frac{1}{\sqrt{3(\nu - 1)}} \sin\sqrt{3(\nu - 1)x} \int_N^x F_1(\xi) \cos\sqrt{3(\nu - 1)} d\xi + \cos\sqrt{3(\nu - 1)x} \int_N^x F_1(\xi) \sin\sqrt{3(\nu - 1)} d\xi. \quad (4.2)$$

Форма свободной поверхности в этом случае имеет вид (фиг. 4). Эта форма существенно отличается от той, которая описывается функцией (2.4). В самом деле, если $\nu < 1$, то, если слева на φ (вверх по потоку) нет волн, то их нет и справа на φ . Если же $\nu > 1$, то при отсутствии волн слева на φ справа (вниз по потоку) поверхность жидкости обязательно будет волновой —

$$\varphi(x) = \frac{1}{\sqrt{3(\nu - 1)}} \{A \sin\sqrt{3(\nu - 1)x} + B \cos\sqrt{3(\nu - 1)x}\},$$

где A и B — постоянные:

$$A = \int_N^N F_1(x) \cos\sqrt{3(\nu - 1)x} dx, \quad B = \int_N^N F_2(x) \sin\sqrt{3(\nu - 1)x} dx.$$

Когда $\nu > 1$, то обтекание препятствия потоком также неоднозначно. В частности, можно построить течение, свободная граница которого при $\max|y_0(x)| \rightarrow 0$ переходит в непериодическую кривую, имеющую горизонтальные асимптоты. Обозначим $\nu - 1 = \epsilon$ и, отбрасывая малые третьего порядка, получим

$$3 \frac{\partial^2}{\partial x^2} f^*(x, y, \nu). \quad (4.3)$$

Решение уравнения (4.3) будем искать в виде ряда $y_0 = y_0^{(0)} + \epsilon y_0^{(1)} + \dots$, где $y_0^{(0)}$ удовлетворяет уравнению

$$3 \frac{\partial^2}{\partial x^2} y_0^{(0)} = 0. \quad (4.4)$$

Заметим, что уравнение (4.4) преобразованием $x = \frac{2}{3} \xi$ переходит в уравнение (3.2); последнее имеет решение (3.3). Следовательно, уравнение (4.4) имеет частное решение:

$$\frac{1}{ch^2 \left\{ \frac{1}{2} \sqrt{3} (x - x_0) \right\}} = \frac{2}{3}. \quad (4.5)$$

Это решение будет также описывать уединенную волну, но асимптотическая глубина которой

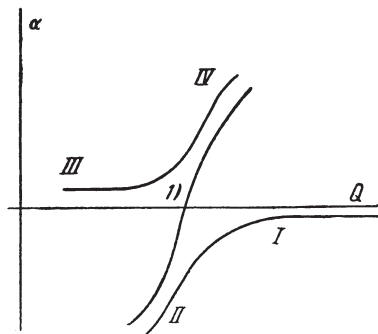
$$h^* = h[1 - 2(\nu - 1)].$$

Скорость распространения этой волны будет

$$v = \sqrt{\frac{gh}{1}}.$$

Если в качестве функции $y_0^{(0)}$ мы возьмем решение (4.5), то вычисления $y_0^{(1)}, y_0^{(2)}, \dots$ и т. д. ничем не будут отличаться от тех, которые были проделаны в конце п. 3.

Таким образом, для $\nu > 1$ мы нашли поток, обтекающий препятствие, свободная граница которого при $\max|y_0| \rightarrow 0$ переходит в некоторую уединенную волну, асимптоты и положение горба которой вполне определены заданием расхода Q .



Фиг. 5

Итак, если $\nu > 1$ и достаточно мало отличается от единицы, то можно приближенно определить во всяком случае два режима обтекания.

5. В заключение представим себе качественно возможный характер потока в плоскости Q (амплитуда — расход), если Q изменяется в окрестности значения $\sqrt{gh^3}$ [схематично эти зависимости изображены на чертеже (фиг. 5)].

Пусть $Q = \sqrt{gh^3}$ ($\nu = 1$). Тогда возможна такая форма обтекания препятствия, при которой свободная граница описывается формулой (2.4) (кривая *I*). При уменьшении расхода эта форма равновесия свободной поверхности деформируется и при $Q = \sqrt{gh^3}$ может начать изменяться по закону, описываемому кривой *II*. Если наибольшее возвышение препятствия $h = 0$, то свободная граница в этом случае переходит в уединенную волну, асимптота которой определяется равенством

$$h^* = h + \frac{2}{3} \frac{gh^3}{Q^2} = 1.$$

Одновременно возможна форма равновесия свободной границы, которая близка к горизонтальной прямой (кривая *III*) и описывается формулой (4.2). Эта форма при увеличении расхода Q будет как-то изменяться, и если Q станет больше $\sqrt{gh^3}$, то может перейти в форму, близкую к уединенной волне с асимптотой $h^* = h$.

Изучение вопроса неоднозначности, которое изложено в этой заметке, никак не претендует на полноту. Ряд обстоятельств просто остался невыясненным, например, вопрос о возможности существования других форм равновесия, отличных от рассмотренных. К тому же рассмотрение является приближенным.

Литература

1. Кочин Н.Е. О движении тяжелой жидкости в канале с дном, имеющим уступ. Собр. соч., т. II. — М., Изд-во АН СССР, 1949.
2. Лаврентьев М.А. и Шабат Б.В. Методы теории функций комплексного переменного. — ГИТТЛ, 1951.
3. Ламб. Гидродинамика. — ГИТТЛ, 1947.

1957 г.

Теорема существования и неединственности вихревых волн периодического типа

1. Определение установившихся гравитационных волн на поверхности завихренной жидкости сводится к следующей краевой задаче:

$F(\cdot)$ в области T

$$(x, 0) = 0, \quad (x, t) = 1$$

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2vf = 0 \text{ при } y = f, \quad v = \frac{gh^3}{Q^2}, \quad h = \frac{Q}{c}. \quad (1)$$

Здесь $y = f(x)$ — уравнение неизвестной свободной границы (см. рис.), v — функция тока, h — «глубина» жидкости, Q — расход, c — скорость.

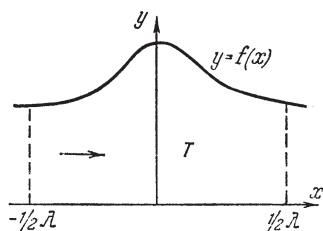
В соотношениях (1) все неизвестные предполагаются безразмерными — в качестве характерных размеров приняты h и Q .

Постановка и первые результаты задачи (1) принадлежат Дюбрей-Жакотен [1]. К настоящему времени наиболее общие результаты в этой области опубликованы Гуйоном [2, 3]. Его основные предположения следующие: $F(\cdot)$ — непрерывная функция, а поток близок к равномерному потоку $v = u$.

2. Если в качестве независимых переменных выбрать x и y , а в качестве искомого, принять $u = u(x, y)$ и $v = v(x, y)$, то задача, сводится к следующей [2]:

$$\begin{aligned} uv_x - vu_x - u_x v_y + v_x u_y &= F(x, y) \\ v_y = 0 \text{ при } y = f(x), \\ u_x - v v_x - v u_y / u &= 0 \text{ при } y = 0. \end{aligned} \quad (2)$$

Задача (2) всегда допускает решение



$$v = 0, \quad u = z \int_0^{\frac{1}{2}} (2F - 1) dz.$$

Предположим, что функция $F(z)$ такова, что $z'(z) > 0$, где d — некоторая положительная постоянная.

Введем новые переменные

$$u = ze^{\cos \theta}, \quad v = ze^{\sin \theta} \quad \frac{1}{2} \ln \frac{u^2 - v^2}{z^2}.$$

Здесь θ — угол наклона вектора скорости. В этих переменных задача (2) эквивалентна следующей:

$$z = x - \theta_1(\theta), \quad z = x - \theta_2(\theta), \quad (3)$$

$$0 \text{ при } \theta, \quad (4)$$

$$\frac{d}{dx} \left(v e^{2\theta} \frac{d\theta}{dz} \right). \quad (5)$$

Звездочка означает, что значения θ_1 и θ_2 — граничные. Φ_1 и Φ_2 — нелинейные операторы

$$\theta_1(\theta) = z(e^{-\theta} - 1) - x(\cos \theta - 1) - x \sin \theta,$$

$$\theta_2(\theta) = \frac{F(\theta)}{z(\theta)}(e^{-\theta} - e^{-\theta_0}) - z(e^{-\theta} - 1) - x(\cos \theta - 1) - x \sin \theta.$$

Задача (3)—(5) допускает тривиальное решение $\theta = 0$. Поставим задачу отыскания периодических по x решений задачи (3)—(5) периода 2π (безразмерная длина волны), добавив к перечисленным условиям условие симметрии

$$\theta(\pi/2) = \theta(3\pi/2) = 0. \quad (6)$$

3. Рассмотрим вспомогательную задачу, в которой условие (5), мы заменим следующим:

$$(x, 1) \frac{d}{dx} \left(v e^{2\theta} \frac{d\theta}{dz} \right) = f_1(x). \quad (7)$$

Положим $\theta_1 = \theta_2$; $\theta_1 = \theta_2$, где θ_1 и θ_2 — решение краевой задачи (4), (6) и (7) для системы (задача А)

$$z = 0, \quad z = x = 0, \quad (8)$$

где z_1 и z_2 решают однородную краевую задачу для систем (3), в которой выражения правых частей заменены следующими (задача В):

$$i(z_1 - z_2, z_1 - z_2).$$

Лемма 1. Решение задачи А имеет вид

$$z_1 = A_1(z), \quad z_2 = B_1(z) + \int_0^x \frac{d^*}{dx} dz. \quad (9)$$

Здесь операторы A_1 и B_1 — линейные интегральные операторы со слабой особенностью и $z_0 = z^*(0)$.

Лемма 2. Задача В эквивалентна следующей системе уравнений:

$$\begin{aligned} z_2(x, z) &= \int_0^{1/2} K_{11}(x, z; x, z) z_1(z_1 - z_2, z_1 - z_2) dx dz \\ z_1(x, z) &= \int_0^{1/2} K_{12}(z_1(x, z), z_2(x, z); z_1 - z_2) dx dz \\ z_2(x, z) &= \int_0^{1/2} K_{21}(z_1) dx dz + \int_0^{1/2} K_{22}(z_2) dx dz, \end{aligned} \quad (10)$$

где

$$K_{11} = \frac{m}{n, m} \frac{2 \sin n x \sin n x}{(n)^2} \frac{g_m(z) g_m(z)}{z(z)},$$

$$K_{12} = \frac{n}{n, m} \frac{2 \sin n x \cos n x}{(n)^2} \frac{g_m(z) g_m(z)}{z(z)},$$

$$K_{21} = \frac{n}{n, m} \frac{2 \cos n x \sin n x}{(n)^2} \frac{g_m(z) g_m(z)}{z(z)},$$

$$K_{22} = \frac{m}{n, m} \frac{2 \cos n x \cos n x}{(n)^2} \frac{g_m(z) g_m(z)}{z(z)}.$$

Здесь $z = 2 / \dots$, g_n и n — собственные функции и числа оператора:

$$\frac{d}{dz} z \frac{dg_n}{dz} = \frac{2}{z} g_n, \quad g(1) = \frac{dg}{dz} = 0, \quad \frac{z}{n} \frac{dq_n}{dz}.$$

Функции g_n и x_n нормированы с весом $1/z$.

Лемма 3. Для любых дифференцируемых g_1 и x_1 , удовлетворяющих условиям $|g_1| < \epsilon$, $|x_1| < \epsilon, \dots, |g_n| < \epsilon$, где ϵ — некоторое достаточно малое положительное число, система (10) имеет единственное решение

$$g_2 = A_2(g_1, x_1), \quad x_2 = B_2(g_1, x_1), \quad (11)$$

где A_1 и B_2 — интегро-степенные ряды своих переменных; они сходятся равномерно в прямоугольнике T

$$|x| < \frac{1}{2}, \quad |g| < \frac{1}{2}.$$

Справедливость этой леммы следует из теории интегро-дифференциальных уравнений, развитой в [5], частным случаем которых является система (10). Кроме того, доказательство опирается на тот факт, что однородная краевая задача (4), (6), (7) для системы (3) имеет только тривиальное решение.

Обозначим через A_i^* операторы A_i при условии, что $g_i = 1$. Тогда

$$A_2^*(A_1, x_1) = D. \quad (12)$$

На основании сказанного D — интегро-степенной ряд, который сходится равномерно на $|x| < \frac{1}{2}$, $|g| < \frac{1}{2}$, если величина ϵ достаточно мала. Подставляя полученные выражения в уравнение (5), получим

$$\frac{v}{z^2} \exp(-2g_0) \exp \int_0^x dx. \quad (13)$$

Функционал $\exp(-2g_0)$ — аналитический. Обозначим

$$k = v \exp(-2g_0).$$

Тогда задача сведется к уравнению вида

$$kR = \dots \quad (14)$$

Согласно предыдущему, оператор R — это оператор Ляпунова и, следовательно, в силу общей теории уравнение (14) имеет нетривиальные решения с малой нормой в окрестности однократных собственных значений соответствующей линейной задачи. Отсюда, принимая во внимание структуру функционала k , мы приходим к следующей основной теореме.

Теорема. Если функция $F(\cdot)$ такова, что существует интеграл

$$\int_0^1 F(\cdot) d\tau \quad (z, d = 0),$$

то для достаточно малого $\epsilon > 0$ задача (3)—(5) при фиксированном значении периода имеет однопараметрическое семейство решений, если только

$$v_n \neq v, \quad (n = 1, 2, \dots)$$

где v_n — собственные числа линеаризованной задачи.

Примечания. 1. Для расчета параметров волн нет необходимости строить эффективно оператор R . Можно показать, что решение задачи (3)—(5) является аналитической функцией параметра $\sqrt{v_n - v}$. Поэтому проще всего непосредственно искать решение исходной краевой задачи в виде рядов по степеням этого параметра.

2. Полученное решение будет близко не к равномерному потоку, как это было в предыдущих исследованиях, а к некоторому вихревому потоку. При $F = 0$ из полученных результатов следуют классические результаты Некрасова — Леви-Чивита.

3. Для определения функционала J_0 может быть использовано равенство

$$\int_0^1 (u_x)_{x=0} d\tau = \int_0^1 z(\tau) \exp(\theta, \tau) d\tau, \quad (15)$$

где

$$\theta = B_1(\cdot) - B_2(A_1(\cdot)), \quad \theta_0 = B_1(\cdot).$$

Операторы A_1 в B_1 имеют следующий вид:

$$A_1(x) = \int_0^x \frac{2 \cos nx \sin nx}{n} dx \frac{z(1)g_n^*(1)}{n} dx$$

$$B_1(x) = \int_0^x \frac{2 \cos nx \sin nx}{n} g_n^*(1) dx.$$

Здесь g_n — функции, удовлетворяющие уравнению

$$z \frac{d}{dz} (z g_n^*) - n^2 z g_n^* = 0$$

и условию $g_n(0) = 0$. Функции g_n^* имеют вид

$$g_n^*(z) = \operatorname{ch} n \frac{d}{z(z)}.$$

5. Функции g_n , введенные в лемме 2, имеют вид

$$g_n(z) = \cos n \frac{d}{z(z)} - n \frac{z}{z(z)},$$

где — нормирующий множитель.

Следовательно, все вычисления могут быть проведены эффективно при любом задании функции $F(z)$.

В заключение считаю своим приятным долгом выразить признательность Ю.А. Кравченко (Франция, Гренобль), обратившему мое внимание на эти задачи, за дискуссии, которые послужили причиной этого исследования.

Литература

1. *Dubreil-Jacotin M.L.* Sur la determination rogoreuse des ondes permanents periodiques d'ampleur finie. J. Math. Pures et appl. 13, 1934.
2. *Gouyon R.* Sur les houles planes en profondeur infinie. C.R., 1958, t. 247, 33–35.
3. *Gouyon R.* Sur les houles planes en profondeur finie. C.R., 1958, t. 247, 180–182.
4. *Моисеев Н.Н.* О течении тяжелой жидкости над волнистым дном./ ПММ, 1957, т. 21, вып. 1.
5. *Lichtenstein* Vorlesungen ueber hichtlinearen Integral und Integrodifferenzial gleichungen. Berlin, 1931.

1960 г.

К теории волн в завихренной жидкости

В теории волн, распространяющихся на поверхности завихренной тяжелой жидкости, есть два вопроса, имеющие определенный теоретический интерес.

1. В потенциальном потоке тяжелой жидкости существует понятие критической скорости $c = \sqrt{gh}$, где h — глубина жидкости. Свойства потока существенным образом зависят от того, является ли поток докритическим или сверхкритическим. Например, если сверхкритический поток обтекает какое-либо препятствие, то за препятствием волн не образуется, и волновое сопротивление равно нулю. Если обтекающий поток — докритический, то за телом образуются волны, и волновое сопротивление отлично от нуля. Имеет ли место аналогичное явление в потоке завихренной жидкости и как изменяется критическая скорость в зависимости от интенсивности вихря?

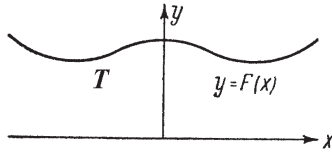
2. М.А. Лаврентьев установил следующий факт, имеющий место в теории потенциальных волн [1]. Если увеличивать длину волны λ , сохраняя при этом ее амплитуду a , то волна при $\lambda \rightarrow \infty$ переходит в уединенную, а горизонтальная прямая, проходящая через нижние точки впадин, вырождается в асимптоту ее поверхности. Сохраняется ли этот факт в теории вихревых волн? Существуют ли вообще на поверхности завихренной жидкости уединенные волны?

В работе приводится исследование этих вопросов, которое опирается на асимптотические методы, являющиеся обобщением известного метода приближенных конформных отображений узких полос М.А. Лаврентьева [2, 3].

1. Постановка задачи

Функция тока ψ плоского течения удовлетворяет уравнению

$$(1.1)$$



Фиг. 1

Здесь — вихрь; если движение установившееся, то линии тока являются изовихревыми линиями и ().

Если функция () задана, приходим к следующей формулировке задачи об установившихся волнах на поверхности завихренной жидкости (фиг. 1).

Определить функцию (x, y), удовлетворяющую внутри T уравнению

$$(x, y) = F(x), \quad (1.2)$$

и функцию $y = f(x)$ по следующим условиям:

$$0 \text{ при } y = 0, \quad 1 \text{ при } y = f(x), \quad (1.3)$$

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2vf = C \text{ при } y = f(x) \quad (v = gh^3/Q^2). \quad (1.4)$$

Здесь C — постоянная энергия. Эти выражения записаны в безразмерном виде. В качестве основных характерных размеров приняты Q — расход и h — глубина жидкости во впадине волны. Задача, так сформулированная для случая периодических волн, в строгой постановке впервые изучалась Дюбрей-Жакотен [4].

Первое из условий (1.3) — это условие непротекания, второе означает, что свободная поверхность есть линия тока, условие (1.4) устанавливает постоянство давления на свободной границе.

Для эффективного решения задачи и упрощения анализа рассмотрим два частных вида задания функции $F(x, y)$:

$$F(x, y) = 2kx, \quad F(x, y) = kx^2. \quad (1.5)$$

Первый случай соответствует линейному закону распределения скоростей в потоке, второй — квадратичному.

2. Вспомогательная задача

1°. Рассмотрим вспомогательную задачу определения функции, удовлетворяющей уравнению (1.2) и обоим условиям (1.3), в предположении, что функция $y = f(x)$, граница области T , задана. В работах [5, 6] была изложена процедура построения формального решения подобной задачи.

В уравнении (1.2) положим $x = y^2$. Решение ищем в виде ряда

$$u_0 + u_2 y^2 + \dots,$$

подчиняя функции u_n следующим краевым условиям:

$$u_n = 0 \text{ при } y = 0, \quad u_0 = 1, \quad u_2 = u_3 = \dots = 0 \text{ при } y = 1, \quad f'(0) = f'(1).$$

Вычисление функций u_n не представляет труда. Возвращаясь затем к старым переменным, получим

1. Если $F(x) = 2k$, то

$$(x, y) = \frac{1}{f} k f y + k y^2 - \frac{f f(1 - k f^2) - 2f}{6f^3} y(y^2 - f^2) + \dots \quad (2.1)$$

2. Если $F(x) = k$, то

$$(x, y) = \frac{\text{sh}\sqrt{k}y}{\text{sh}\sqrt{k}f} - \frac{Bf}{2\sqrt{k}} \text{cth}\sqrt{k}y - \frac{By}{2k} \text{ch}\sqrt{k}y, \quad (2.2)$$

где

$$B = \frac{\frac{1}{2}\sqrt{k}f \text{sh}2\sqrt{k}f - kf^2(1 - \text{ch}^2\sqrt{k}f)}{\text{sh}^3\sqrt{k}f}.$$

В общем случае ряды (2.1) и (2.2) расходятся.

Функцию (x, y) будем называть r -асимптотическим решением задачи (1.2) и (1.3), если ее решение можно представить в виде

$$(x, y) = (x, y) + O(r),$$

где r — некоторый малый параметр.

Теорема. Для того чтобы функция

$$u_n = 0 + \dots + u_n$$

была r -асимптотическим решением задачи (1.2), (1.3), необходимо и достаточно, чтобы производные $f^{(m)}(x)$ удовлетворили условиям

$$f^{(m)}(x) = O(r^m) \quad (m = 1, \dots, 2n - 2),$$

где l_m — произвольные положительные числа, а степень (число r) определяется порядком функции $n-2$

$$O(n-2) = O(r).$$

Доказательство этой теоремы проводится аналогично доказательству такой же теоремы для уравнения Лапласа [6].

Таким образом, если граница f достаточно гладкая, то конечные отрезки рядов (2.1) или (2.2) можно использовать в качестве приближенных решений. Если, например, ограничиться функцией $n-2$, то достаточно потребовать малости первых трех производных.

В дальнейшем будем ограничиваться учетом членов, содержащих первые и вторые производные.

Вычислим с этой точностью граничные значения производных x и y (при $y = f(x)$); в частности, нетрудно установить, что

$$v^2 = \frac{2}{x} \frac{2}{y} p f + q f^2 + r. \quad (2.3)$$

Здесь p , q и r — функции f ; если $F(\cdot) = 2k$, то

$$p = \frac{2}{3f} + \frac{8}{15} k f^2, \quad q = \frac{1}{3f^2} + \frac{8}{15} k f^2, \quad r = \frac{1}{f^2} + \frac{2}{3} k. \quad (2.4)$$

Если $F(\cdot) = k$, то

$$p = \frac{\sqrt{k} \operatorname{ch}^2 \sqrt{k} f}{\operatorname{sh}^3 \sqrt{k} f} - \frac{\sqrt{k} f \operatorname{ch}^2 \sqrt{k} f}{\operatorname{sh} \sqrt{k} f} - \frac{\sqrt{k} f \operatorname{sh} \sqrt{k} f}{\operatorname{ch} \sqrt{k} f}, \quad r = \frac{\operatorname{ch}^2 \sqrt{k} f}{\operatorname{sh}^2 \sqrt{k} f},$$

$$q = \frac{k \operatorname{ch}^2 \sqrt{k} f}{\operatorname{sh}^2 \sqrt{k} f} + \frac{1}{\operatorname{sh}^2 \sqrt{k} f \operatorname{ch} \sqrt{k} f} - \frac{\sqrt{k} f \operatorname{ch}^2 \sqrt{k} f}{\operatorname{sh} \sqrt{k} f} - \frac{\sqrt{k} f \operatorname{sh} \sqrt{k} f}{\operatorname{ch} \sqrt{k} f}.$$

Приближенное решение вспомогательной задачи описывает вихревое движение идеальной жидкости в криволинейной трубе.

Примечание. Ограничения, наложенные на вид функции $F(\cdot)$, не носят принципиального характера, изложенный метод позволяет построить решение в квадратурах для произвольного задания $F(\cdot)$. В самом деле, функция φ_0 будет удовлетворять уравнению

$$0,yy \quad F(0).$$

Решение этого уравнения выражается в квадратурах

$$y \quad A_2 \int_0^z F(z) dz \quad A_1 \int_0^z dz.$$

Здесь постоянные A_1 и A_2 находятся из условий

$$y = 0 \text{ при } z = 0, \quad y = f(x) \text{ при } z = 1.$$

3. Уравнения длинных волн

Предположим априори, что поверхность изучаемой волны достаточно гладкая и ее производные малы. Заменим выражение квадрата скорости в (1.5) его приближенным выражением (2.3). Это позволит свести задачу к обыкновенному дифференциальному уравнению относительно свободной границы

$$f \quad p(f) \quad f^2 q(f) \quad r(f) \quad 2vf \quad C. \quad (3.1)$$

Задача состоит в определении трехпараметрического семейства решений. В качестве определяющих параметров целесообразно принять постоянную энергию C , число v и характеристику завихренности k . Тогда амплитуда волны и ее длина будут однозначно определяться заданием этих трех величин. Уравнение (3.1) допускает первый интеграл

$$u \quad \int_1^f \frac{2r(\xi)}{p(\xi)} d\xi \quad 4v \frac{f}{p(f)} \quad \frac{2C}{p(f)} \exp \int_1^f \frac{p(t)}{q(t)} dt \quad d. \quad (3.2)$$

Здесь $u = \int_1^f f^2 \exp \int_1^f \frac{p(t)}{q(t)} dt$. Постоянная интегрирования выбрана

из условия, что $u = 0$ при $f = 1$ (в нижней точке впадины скорость горизонтальна). Так как в вершине волны скорость также горизонтальна, то постоянную энергию легко выразить через высоту горба f^* .

Исследуем уравнение (3.2) качественно в плоскости uf , предполагая характеристику завихренности k малой.

4. Преобразование уравнения (3.2)

Введем вместо f переменную $y = f - 1$ и разложим функции, входящие в (3.2), в ряд по k .

1) Если $F(\nu) = 2k$, то будем иметь

$$\begin{aligned}
 u & y \{ 3 \ 6\nu \ 3C \ k(12\nu \ 6C) \ k^2(\dots) \ \dots \} \\
 & y^2 \{ 9\nu \ 3C \ k(6 \ 42\nu \ 18C) \ k^2(\dots) \ \dots \} \\
 & y^3 \{ 3\nu \ k(6 \ 54\nu \ 18C) \ k^2(\dots) \ \dots \} \\
 & y^4 \{ k(30\nu \ 6C) \ k^2(\dots) \ \dots \} \ y^5 \{ 6k\nu \ k^2(\dots) \ \dots \} \\
 & y^6 \{ k^2(\dots) \ \dots \} \ \dots
 \end{aligned} \tag{4.1}$$

2) Если $F(\nu) = k$, то

$$\begin{aligned}
 u & y \ 3 \ 6\nu \ 3C \ k \ \frac{16}{5}\nu \ \frac{8}{5}C \ \frac{2}{5} \ k^2(\dots) \ \dots \\
 & y^2 \ 9\nu \ 3C \ k \ \frac{6}{5} \ \frac{56}{5}\nu \ \frac{24}{5}C \ k^2(\dots) \ \dots \\
 & y^3 \ 3\nu \ k \ \frac{8}{5} \ \frac{72}{5}\nu \ \frac{24}{5}C \ k^2(\dots) \ \dots \\
 & y^4 \ k \ 8\nu \ \frac{8}{5}C \ k^2(\dots) \ \dots \\
 & y^5 \ k \ \frac{8}{5}\nu \ k^2(\dots) \ \dots \ y^6 [k^2(\dots) \ \dots] \ \dots
 \end{aligned} \tag{4.2}$$

Для дальнейших вычислений удобно ввести новую постоянную M так, чтобы при $M=0$ обращался в нуль коэффициент при первой степени y . Положим

$$3M = 3 \ 6\nu \ 3C \ k(12\nu \ 6C) \ \dots \tag{4.3}$$

Тогда уравнение (4.1) преобразуется к следующему

$$\begin{aligned}
 u & 3My \ y^2 \{ 3M [1 \ 4k \ k^2(\dots) \ \dots] \ \dots \} \\
 & y^3 \{ 3\nu \ 6k [3(M \ \nu) \ 2] \ k^2(\dots) \ \dots \} \\
 & 6y^4 [k(3\nu \ M \ 2) \ k^2(\dots) \ \dots] \\
 & y^5 [6\nu k \ k^2(\dots) \ \dots] \ \dots \ P(y, k, \nu, M).
 \end{aligned} \tag{4.4}$$

Параметр ν в этом уравнении определяется формулой

$$(1 - \nu) \frac{3}{(1 - k)^2} - \frac{3k^2}{2(1 - k)^2} (1 - \nu)(3 - 6k) - k^2 \frac{15}{2} - 9\nu \dots \quad (4.5)$$

3) Если $F(\nu) = k$ то будем иметь

$$\begin{aligned} u &= 3M\nu - y^2 - 3M - 1 - \frac{16}{15}k - k^2(\dots) \dots \\ y^3 &= 3\nu - 2k - \frac{12}{5}(M - \nu) - \frac{8}{5} - k^2(\dots) \dots \\ y^4 &= \frac{8}{5}k(3\nu - M - 1) - k^2(\dots) \dots - y^5 - \frac{8}{5}\nu k - k^2(\dots) \dots \dots \\ &\dots P(y, k, \nu, M). \end{aligned} \quad (4.6)$$

Параметр ν в этом уравнении определяется формулой

$$(1 - \nu) - 3 - \frac{8}{5}k - k^2(\dots) \dots \quad (4.7)$$

5. Качественное исследование уравнения (4.4)

1°. Уравнение (4.4) описывает семейство кривых, зависящих от трех параметров: постоянной M (которая однозначно определяется амплитудой), параметром ν (зависящим от скорости распространения c) и характеристикой завихренности k . Эту совокупность кривых легче всего изучить качественными методами. Для этого надо знать распределение нулей полинома, стоящего в правой части (4.4). Для малых значений k характер этого распределения может быть обнаружен на основании следующих лемм.

Лемма 5.1. Три наименьшие по абсолютной величине корня y_1, y_2, y_3 уравнения $P(y, k, \nu, M) = 0$ при $k \rightarrow 0$ будут стремиться к корням уравнения

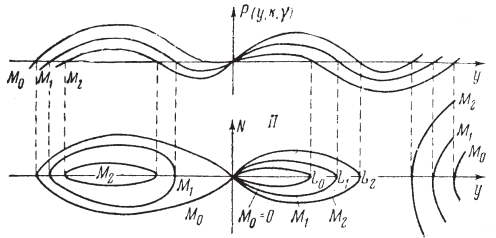
$$P(y, 0, \nu, M) = 0.$$

Лемма 5.2. Остальные корни этого уравнения будут неограниченно возрастать при $k \rightarrow 0$, причем

$$y_{4,5} = \frac{1}{2\sqrt{k}} + O\left(\frac{1}{k}\right) \text{ и т. д.}$$

Доказательство обеих этих лемм следует из того факта, что $P(y, k, v, M) = 0$ при $k \rightarrow 0$ стремится равномерно к $P(y, 0, v, M) = 0$ в любой ограниченной области плоскости комплексного переменного k .

2°. Рассмотрим сначала случай $M = 0$. Пользуясь приведенными леммами, легко убеждаемся, что в этом случае характер кривой $u = P(y, k, v, M)$ будет таким, каким он изображен на фиг. 2, где

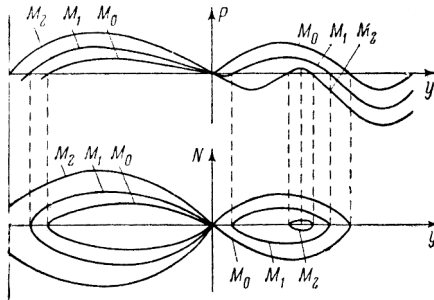


Фиг. 2

$$M_2 > M_1 > M_0 > 0, \\ N \text{ и } dy/dx \text{ } df/dx.$$

Рассмотрим теперь структуру фазовой плоскости, представленной на фиг. 2. Так как по смыслу величина y положительна и $u(0) = 0$ (касательная во впадине горизонтальна), то физический смысл имеют лишь те кривые, которые проходят через начало координат и лежат в правой полуплоскости (зона II).

Топологическая структура фазовой плоскости определяется параметрами k и M , которые при данных k и v зависят от амплитуды M . С ростом параметра M увеличивается k . Если амплитуда M будет уменьшаться, то параметр k также будет уменьшаться и при некотором значении $k = 0$ он обратится в



Фиг. 3

ноль; это предельное решение на фиг. 2 обозначено через M_0 . Дальнейшее уменьшение M приведет к отрицательным значениям постоянной M . Нетрудно убедиться, что в этом случае уравнение (4.4) не содержит решений, имеющих физический смысл. Для этого достаточно составить диаграмму, аналогичную рассмотренной (фиг. 3). В самом деле, если $M > 0$ и $M = 0$, то один из корней, лежащих в левой полуплоскости, стремится к нулю. При $M = 0$ полином P имеет двойной нуль. При $M < 0$ этот корень переходит в правую

полуплоскость, а наклон касательной к P в нуле становится отрицательным.

Следовательно, при $\gamma = 0$ уравнение (4.4) не содержит ограниченных решений, лежащих в правой полуплоскости, проходящих через начало координат и вдоль которых значение постоянной M — отрицательно.

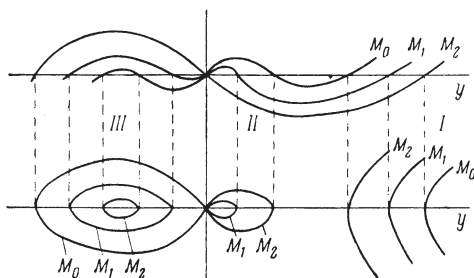
Из этого можно заключить, что при условии $\gamma = 0$ на поверхности завихренной жидкости могут существовать лишь такие волны, амплитуды которых превосходят некоторое число M_0 . Таким образом, обнаружено существование однопараметрического (при фиксированных ω и k) семейства решений уравнения (4.4), которое при $\gamma \rightarrow 0$ стремится к особому решению $M_0(\omega)$.

Если $\gamma > 0$, то особое решение M_0 стягивается в точку.

Для случая $\gamma = 0$ фазовая диаграмма изображена на фиг. 4.

Анализ этой диаграммы нам показывает, что нулевое значение параметра бифуркационное: при $\gamma = 0$ вся топологическая структура фазовой плоскости становится иной, исчезает особое решение и ограничение снизу для амплитуды волн.

При $\gamma = 0$ на поверхности завихренной жидкости могут существовать волны любой сколь угодно малой амплитуды. Решения, принадлежащие зоне II (фиг. 4), описывают волны, близкие к обычным периодическим волнам, которые изучались Дюбрей-Жакотен и ее последователями [7, 8].



Фиг. 4

Так же, как и в предыдущих случаях, отрицательные значения M приводят к решениям, не имеющим физического смысла.

3°. Итак, мы установили, что значение $\gamma = 0$ — бифуркационное. Математически это означает, что изменение знака γ приводит к изменению числа корней в правой полуплоскости. Следовательно, смысл равенства $\gamma = 0$ совершенно такой же, как и равенства $\nu - 1 = 0$ в теории потенциальных течений. Из этой теории известно, что свойства потока при $\nu - 1 > 0$ (докритическое течение) существенно отличаются от свойств потока при $\nu - 1 < 0$ (сверхкритическое течение).

Рассмотрим формулу (4.5). Полагая $\epsilon = 0$, найдем, что

$$v_* = 1 + \frac{1}{2}k^2. \quad (5.1)$$

Таким образом, критическая величина параметра $v = v_*$ — четная функция характеристики завихренности k и всегда больше 1.

Переходя к размерным величинам, мы найдем, что

$$c^{*2} = \frac{qh}{1 + \frac{1}{2}K^2}. \quad (5.2)$$

Здесь c^* — критическая скорость распространения волны. Так как в случае потенциального потока

$$c^* = \sqrt{gh}, \quad (5.3)$$

то из равенства (5.2) следует утверждение: значение критической скорости в вихревом потоке всегда меньше критической скорости в потенциальном потоке. Этот факт можно объяснить так: равенство (5.3) означает, что критическое значение скорости соответствует равенству потенциальной энергии удвоенному значению кинетической. Аналогичный смысл имеет равенство (5.2). Так как кинетическая энергия вихревого потока при прочих равных условиях больше кинетической энергии потенциального потока, а потенциальная энергия определяется только геометрией свободной поверхности, то равенство потенциальной энергии потока удвоенной кинетической наступает при меньших значениях средней скорости потока. Отсюда и следует, что критическая скорость в вихревом потоке всегда будет меньше критического значения c в потенциальном.

6. Исследование особого решения. Коноидальные волны в завихренной жидкости

1°. Уравнения (4.4) и (4.6) могут быть проинтегрированы в квадратурах. Однако результат получается громоздким. Поэтому при исследовании ограничиваемся тем случаем, когда решение близко к особому (M мало) и режим близок к критическому (ϵ мало). Обозначим

$$M = \epsilon^2 m, \quad a = 3 + 6k + 12kv, \quad b = 6k(3v - 2), \quad d = 6vk.$$

Если v близко к 1, а k невелико, то $a > 0$. Уравнение (4.4) будет

$$\frac{dy}{dx} = 3^2 my \{3(1-4k)^2 m\} y^2 (a - 18km^2) y^3 + (b - 6m^2) y^4 - cy^5 \dots \quad (6.1)$$

Можно пытаться разыскать решения этого уравнения в виде ряда

$$y = \sum_n y_n |x|^{nq}, \quad (6.2)$$

где p и q — некоторые действительные числа. Уравнение (6.1) всегда допускает тривиальное решение. Поэтому необходим предварительный анализ, который должен установить, при каких соотношениях между числами p и q ряд (6.2) может представить нетривиальное решение. Оказывается, что единственными значениями p и q , которые для этого могут быть использованы, будут $p = 1/2$ и $q = 1$.

Подставляя ряд (6.2) в (6.1), можно обычным способом построить процесс вычисления y_n . К тому же самому результату мы придем, если вместо уравнения (6.1) рассмотрим уравнение

$$\frac{dy}{dx} = 3^2 My - y^2 \{3(1-4k)M\} + y^3(a - 18kM^2) - y^4(b - 2M) - cy^5 \dots \quad (6.3)$$

и будем разыскивать решение в виде ряда

$$y = y_1 + y_2 + \dots \quad (6.4)$$

и в окончательном результате положим $c = 1$. Для вычисления такая процедура более удобна. Для функций y_i получим следующие уравнения:

$$\frac{dy_1}{dx} = \sqrt{3My_1^2 - y_1^2 - ay_1^3} - \sqrt{aP(y_1)},$$

$$\frac{dy_2}{dx} = \frac{1}{\sqrt{aP(y_1)}} - y_2 \frac{dP(y_1)}{dy_1} - y_1^2 3M(1-4k) - y_1^3 - y_1^4 b \text{ и т. д.} \quad (6.5)$$

Первое уравнение этой системы нелинейное, остальные — линейные неоднородные.

2°. Рассмотрим уравнение первого приближения. Переменные в этом уравнении разделяются, и мы имеем

$$y \frac{dy}{\sqrt{P(y)}} = \sqrt{ax}. \quad (6.6)$$

Здесь y — максимальное значение y . Оно соответствует началу координат (абсцисса горба волны равна нулю) и определяет максимальное значение корня полинома $P(y)$.

Нас интересуют только ограниченные решения системы (6.5). Поэтому мы должны рассматривать лишь такие значения параметров M и ν , при которых все три корня этого полинома действительные. Один из них всегда равен нулю, один положительный ($\nu > 0$) и один всегда отрицательный ($M > 0$). Перепишем уравнение (6.6)

$$y \frac{dy}{\sqrt{y(y - \nu)(y - M)}} = \sqrt{ax}. \quad (6.7)$$

Левая часть этого уравнения имеет стандартную форму эллиптических интегралов. Поэтому

$$\sqrt{ax} = \frac{2}{\sqrt{1 - k^2}} F(\phi, k), \quad (6.8)$$

Здесь $F(\phi, k)$ — эллиптический интеграл первого рода

$$F(\phi, k) = \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}, \quad \arccos \sqrt{\frac{y}{M}}, \quad k^2 = \frac{M - \nu}{M}. \quad (6.9)$$

Используя известные формулы обращения эллиптических интегралов, можно записать решение уравнения через эллиптические функции

$$\text{am} \sqrt{a \left(\frac{x}{2} \right)}, \quad (6.10)$$

$$\cos \text{am} \sqrt{a \left(\frac{x}{2} \right)} = \text{cn} \sqrt{a \left(\frac{x}{2} \right)}, \quad (6.11)$$

$$y = \cos^2 \text{am} \sqrt{a \left(\frac{x}{2} \right)}. \quad (6.12)$$

Здесь через cn обозначена эллиптическая функция Якоби: ее модуль определяется (6.9). Поэтому волны, определенные формулой (6.12), будем называть коноидальными, по аналогии с коноидальными волнами в незавихренной жидкости, полученные Кортвегом и де Фрисом [9].

3°. Найдем связь между длиной волны λ , амплитудой A , числом Фруда и завихренностью k . Так как $y(0) = A$ (максимальное возвышение), а $y(\lambda/2) = 0$, то из формулы (6.8) находим

$$\frac{4}{\sqrt{a(2 - A^2/a)}} J(\lambda/2). \quad (6.13)$$

Здесь $J(\lambda/2)$ — полный эллиптический интеграл первого рода. Для того чтобы найти требуемую зависимость, надо в формуле (6.13) выразить λ через A . Для этого используем связь λ/a коэффициентов полинома $P(y)$ с корнями y_1 и y_2 . В результате получим

$$\frac{4}{\sqrt{a(2 - A^2/a)}} J(\lambda/2). \quad (6.14)$$

Формула (6.14) является аналогом формулы теории длинных безвихревых волн [6]. Модуль эллиптического интеграла в (6.16) определяется выражением

$$J(\lambda/2) = \int_0^{\lambda/2} \frac{dy}{\sqrt{a(2 - y^2/a)}}. \quad (6.15)$$

4°. Рассмотрим предельный случай малых значений A , т. е. решения, близкие к особому. Отбрасывая в формуле (6.15) слагаемые, содержащие A^2 , получим

$$y = \text{cn}^2 \sqrt{a} \frac{x}{2}. \quad (6.16)$$

В этом случае $\lambda \sim 1$, следовательно, величину $J(\lambda/2)$ можно заметить ее асимптотическим представлением

$$J(\lambda/2) \sim \ln \frac{4}{\sqrt{1 - A^2/a}}.$$

Пользуясь этим выражением, а также выражением (6.15), из (6.13) получим приближенную зависимость длины волны от амплитуды, скорости и завихренности

$$\frac{2}{\sqrt{a}} \ln \frac{16}{a}. \quad (6.17)$$

Эта простая формула дает аналитическое выражение того факта, который был установлен качественными методами. При фиксированном $a > 0$, с уменьшением амплитуды длина волны увеличивается: $\lambda \propto 1/a$. Волн, амплитуда которых меньше a^* , в этом случае не существует.

Замечание. Необходимо еще рассмотреть два вопроса: найти форму предельного решения и показать возможность вычисления последующих приближений (сходимость использованных рядов при малых значениях a следует из общей теории).

Ответ на указанные вопросы нетрудно получить непосредственно из результатов этого параграфа. Однако для наглядности (чтобы не оперировать с эллиптическими функциями) вопрос о предельном решении рассмотрим самостоятельно. Для этого частного случая вопрос о возможности вычисления последующих приближений решается особенно просто.

7. Предельное решение

Предельное решение соответствует нулевому значению постоянной энергии M . Уравнение (6.3) в этом случае имеет вид

$$\frac{dy}{dx} = y^2 - y^3 - y^4 b - y^5 c. \quad (7.1)$$

Система, аналогичная системе (6.5), будет

$$\frac{dy_1}{dx} = \sqrt{y_1^2 (2 - ay_1)}, \quad (7.2)$$

$$\frac{dy_2}{dx} = \frac{1}{\sqrt{y_1^2 (2 - ay_1)}} \{y_2 [2 - y_1 - 3ay_1^2] - y_1^3 - y_1^4 b\} \text{ и т. д.} \quad (7.3)$$

Задача состоит в определении ограниченных решений. Постоянные, появляющиеся при интегрировании этой системы, должны выбираться из условия того, чтобы начало координат соответствовало вершине горба волны, где касательная горизонтальна. Система уравнений (7.2)—(7.3) интегрируется в замкнутом виде

$$y = \frac{1}{\operatorname{ch}^2 \sqrt{1/4} x} - \frac{2}{a} \frac{1}{a^2} + \frac{2b}{a} \dots$$

$$\frac{1}{\operatorname{ch}^4 \sqrt{1/4} x} - \frac{2}{a^3} \dots - \frac{1}{\operatorname{ch}^6 \sqrt{1/4} x} \dots \dots \quad (7.4)$$

Отсюда $\lim_{x \rightarrow \infty} y(x) = 0$ при $x \rightarrow \infty$. Это значит, что особое решение имеет вид уединенной волны. При $k = 0$ ряд (7.4) содержит одно первое слагаемое и описывает форму обычной уединенной волны. В этом случае формула (7.4) совпадает с точностью до величин порядка $(1 - \nu)^2$ с известными приближенными формулами Релея и Буссинеска [10].

Будем называть «амплитудой» уединенной волны максимальное значение функции $y(x)$. Тогда

$$\frac{2}{a} - \frac{2}{a^2} \dots \quad (7.5)$$

Для малых

$$\sim \frac{1}{a} (1 - \nu) \frac{3}{(1 - k)^2} - \frac{3k^2}{2(1 - k)^2} \dots \quad (7.6)$$

Если k отрицательно, то амплитуда будет всегда больше в вихревом случае, чем амплитуда уединенной волны в безвихревом потоке. Уединенная волна может существовать и в том случае, если $1 - \nu = 0$, и даже когда $1 - \nu$ отрицательно для всех значений ν , для которых $\nu > 0$. Этот результат согласуется с результатами качественного исследования.

Таким образом: 1) в вихревом потоке существует критическое значение скорости и это значение всегда меньше критической скорости потенциального потока; 2) на поверхности завихренной жидкости могут существовать уединенные волны (7.4).

Литература

1. Лаврентьев М.А. До теорії довгих хвиль./Зб. Праці Ін-ту матем. АН УРСР, 1946, № 8.
2. Лаврентьев М.А. О некоторых свойствах однолистных функций с приложениями к теории струй./Мат. сб., 1938, т. 4, № 3.
3. Лаврентьев М.А. и Шабат Б.В. Методы теории функций комплексного переменного. — Гостехиздат, 1959.

4. *Dubreil-Jacotin M.L.* Sur la détermination rigoureuse des ondes permanentes périodiques d'amplitude finie. *J. Math.*, 1934, № 13.

5. *Иванов Ю.П., Моисеев Н.Н., Тер-Крикоров А.М.* Об асимптотическом характере формул М.А. Лаврентьева. *ДАН СССР*, 1958, т. 123, № 2.

6. *Моисеев Н.Н., Тер-Крикоров А.М.* Исследование движения тяжелой жидкости при скоростях, близких к критической. *Труды МФТИ*, 1959, № 3.

7. *Gouyon R.* Sur les houles planes en profondeur infinie. *C.R.*, 1958, t. 247, p. 33–35.

8. *Gouyon R.* Sur les houles planes en profondeur finie. *C.R.*, 1958, t. 247, p. 266–269.

9. *Kortweg G., De Vries.* On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary waves. *Philos. Mag.*, 1895, vol. XXXIX, № 5.

10. *Ламб.* Гидродинамика. — Гостехиздат, 1947.

1960 г.

О некоторых особенностях течения жидкости, подверженной действию сил поверхностного натяжения

Рассматриваются некоторые особенности течения идеальной жидкости, подверженной силам поверхностного натяжения: возможность вырождения волн в уединенные, существование уединенных волн типа впадин, существование осесимметричных струй волнообразной формы и возможность потери устойчивости струи, вызванной действием сил поверхностного натяжения. Исследование проводится в рамках асимптотики узких полос (см. [1, 2]).

§ 1. Длинные волны в жидкости, подверженной действию сил поверхностного натяжения

1°. Задача определения установившихся волн на поверхности тяжелой жидкости, подверженной силам поверхностного натяжения, сводится к определению функции $y = f(x)$, гармонической в полосе $0 < y < h$, (фиг. 1) и функции $f(x)$ по условиям

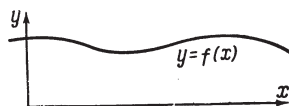
$$y = f(x), \quad y = 0, \quad (1.1)$$

$$x^2 + y^2 + 2\nu f' + K = C \text{ при } y = f(x). \quad (1.2)$$

Задача (1.1), (1.2) сформулирована в безразмерном виде. В качестве характерных параметров приняты глубина жидкости h в некоторой точке, положение которой будет указано ниже, и расход жидкости Q . В соотношении (1.2) безразмерный параметр ν , значение средней кривизны K , безразмерный параметр C , характеризующий действие капиллярных сил, определяются следующими формулами:

$$\nu = \frac{gh^3}{Q^2}, \quad \frac{gh}{V}, \quad K = \frac{f}{2(1 - f'^2)^{3/2}},$$

$$\frac{4h}{Q^2}, \quad \frac{4}{V^2 h}, \quad V = \frac{Q}{h}. \quad (1.3)$$



Фиг. 1

Здесь V — характерная скорость, σ — коэффициент поверхностного натяжения.

2°. Функция ρ , гармоническая внутри полосы $0 < y < f(x)$ и ее производные имеют представления (см. [1], с. 184)

$$(x, y) \sim \frac{y}{f} \frac{2f^2 - f^2}{6f} y + \frac{y^2}{f^2} \dots \quad (1.4)$$

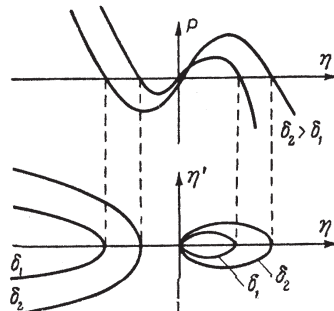
$$v_x = \frac{dy}{dy} \sim \frac{1}{f} \frac{2f^2 - f^2}{6f} 3 \frac{y^2}{f^2} + \dots, \quad v_y = \frac{df}{x} \sim \frac{yf}{f^2} \dots$$

Так как граница $y = f(x)$ заранее неизвестна, то необходимо задать некоторыми априорными оценками о производных функции $y = f(x)$ (имеется в виду изучение длинных волн и условий их вырождения в уединенную). Поэтому естественно принять ограничения М.А. Лаврентьева [1]

$$f = O(\frac{1}{2}), \quad f' = O(\frac{1}{2}), \quad f'' = O(\frac{1}{2}) \text{ и т. д.}, \quad (1.5)$$

имеющие место для уединенных волн в идеальной жидкости.

При условиях (1.5) из (1.4) имеем асимптотическое представление для модуля скорости на кривой $y = f(x)$



$$(v_x^2 + v_y^2)_{y=f(x)} = \frac{1}{f^2} + \frac{2}{3} f f'' = O(\frac{1}{2}). \quad (1.6)$$

Отбрасывая величины порядка $O(\frac{1}{2})$, уравнение (1.2) запишем в виде

$$f + \frac{3}{2} f' + 3 f'' + \frac{3}{2} f''' + c f = 0. \quad (1.7)$$

Фиг. 2

Уравнение (1.7) имеет первый интеграл, и его фазовая плоскость может быть легко изучена. Однако для дальнейшего имеет значение та ее часть, для которой выполнены условия (1.5). Эти условия заведомо выполнены при $f=1$. Поэтому можно получить те же качественные результаты, без громоздких вычислений, если

положим $f = 1$ и удержим в (1.7) несколько первых членов разложения. Ограничиваясь членами порядка $O(\epsilon^2)$, приведем (1.7) к виду

$$(1 - \frac{3}{2}\epsilon) A + B^2 = D, \quad (1.8)$$

$$A = 3\nu(2 - \epsilon) - \frac{3}{2}(1 - \epsilon) - \frac{3}{2}C(1 - \epsilon),$$

$$B = 3\nu(1 - 2\epsilon^2) - \frac{3}{2}(1 - \epsilon^2) - \frac{3}{2}C(1 - \epsilon^2), \quad (1.9)$$

$$D = \frac{3}{2}C - 3\nu - \frac{3}{2}, \quad \frac{3}{2} / (1 - \frac{3}{2}\epsilon).$$

Для некоторого упрощения записи введем вместо C новую постоянную γ : $C = 2 - \gamma$. Тогда формулы (1.9) примут вид

$$A = 3(\nu - 1) - \frac{3}{2}(1 - \epsilon),$$

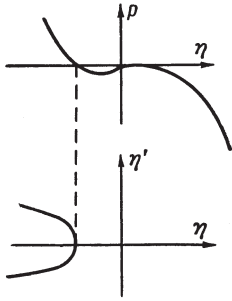
$$B = 3\nu - \frac{3}{2} - 3(\nu - 1) - \frac{3}{2}(1 - \epsilon), \quad D = \frac{3}{2}. \quad (1.10)$$

Уравнение (1.8) допускает первый интеграл

$$\frac{1}{2}(1 - \frac{3}{2}\epsilon)^2 F + D = \frac{1}{2}A^2 + \frac{1}{3}B^3 + P(\epsilon). \quad (1.11)$$

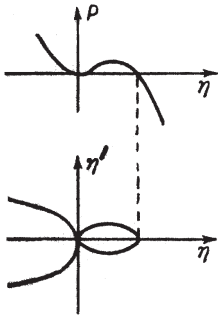
Здесь F — новая постоянная; можем ее принять равной нулю. Это эквивалентно тому, что в качестве линейного размера выбирается глубина жидкости в вершине бугра или во впадине волны. Тогда значение $\epsilon = 0$ отвечает $\gamma = 0$. Таким образом, поведение фазовых кривых определяется тремя параметрами: ν , ϵ и γ .

3°. Рассмотрим сначала тот случай, когда $\epsilon = 0$, $\frac{2}{3}$ и $B = 0$. Последнее условие выполняется, например, если ν близко к 1, а ϵ мало. На фиг. 2 изображены для этого случая кривые $P(\epsilon)$ и фазовая плоскость уравнения (1.8). Ветви кривых, лежащие в левой полуплоскости, физического смысла не имеют, поскольку они описывают неограниченные решения уравнения (1.8). Ветви, лежащие в правой полуплоскости, описывают периодические решения, амплитуда которых тем меньше, чем меньше значение ϵ . Таким образом, при $\epsilon > 0$ по свободной поверхности распространяются волны, длина которых определяется квадратурой:



Фиг. 3

Таким образом, периодическое решение в этом случае выродилось в точку $\eta = 0$, т. е. течение вырождается в плоскопараллельный поток. Совсем иная ситуация будет при $v < 1$. Для этого случая фазовый портрет течения изображен на фиг. 4. Видим, что при $\eta = 0$ максимальное возвышение («амплитуда») ρ^* не стремится к нулю. Свободная поверхность в этом случае представляет уединенную волну.



Фиг. 4

Итак, если $v > 1$, то при $\eta = 0$ волны, которые распространяются с данной скоростью (v — фиксировано), вырождаются в равномерный поток, амплитуда этих волн $\rho^* \rightarrow 0$. Эти волны при $\eta = 0$ были изучены Литменом [3]. Если $v < 1$, то $\lim_{\eta \rightarrow 0} \rho^* \neq 0$ при $\eta = 0$, предельное решение представляет уединенную волну, которая среди волн, распространяющихся с данной скоростью, имеет наименьшую амплитуду $\rho^*(v)$, причем амплитуда уединенной волны стремится к нулю при $v \rightarrow 1$. Эти волны были изучены Кортевегом и де Фрисом [4], соответствующая теорема существования доказана Я.И. Секерж-Зеньковичем [5]. Значение $v = 1$ является бифуркационным. При $v > 1$ существует единственная форма движения бесконечной длины волны — равномерный поток. При $v < 1$ и осуществляются две предельные формы: равномерный поток и уединенная волна. Исследуемые здесь волны аналогичны обычным гравитационным (коноидальным) волнам, в которые они и переходят при $\eta = 0$. Случаи таких волн могут быть легко изучены аналитическими методами. При $v > 1$ волновые движения могут быть исследованы квазилинейными методами (методом Ляпунова—Пуанкаре или

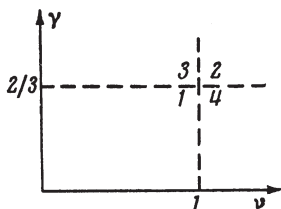
$$\sqrt{2} \frac{d}{\sqrt{(D - \frac{1}{2}A^2 - \frac{1}{3}B^3)(1 - \frac{3}{2}d^2)}}, \quad (1.12)$$

в которой интеграл распространен по всей фазовой траектории. Из (1.12) видим, что $\lim_{\eta \rightarrow 0} \rho^* = 0$. Поэтому естественно поставить вопрос о характере предельного движения при $\eta = 0$. Если $v > 1$ (т. е. если скорость распространения волны относительно покоящейся жидкости $V < \sqrt{gh}$), то при $\eta = 0$ фазовый портрет течения будет таким, каким он изображен на фиг. 3.

методом Крылова—Боголюбова), поскольку уже в первом приближении уравнение (1.8) допускает периодическое решение

$$* \cos \frac{3(1-\nu)^{1/2}}{1-\nu^{3/2}} x \quad (\text{const}). \quad (1.13)$$

Если $\nu < 1$, то уравнение первого приближения не содержит периодических решений, и, следовательно, такие волны нельзя изучать в линейном приближении*. В этом случае решение может быть получено в эллиптических функциях. Решение становится особенно простым $\epsilon = 0$. Легко проверить, что в этом случае уравнение (1.8) имеет решение



Фиг. 5

$$\frac{3|A|}{2B \operatorname{ch}^2 \frac{1}{2} \sqrt{|A|x}} \quad \frac{*}{\operatorname{ch}^2 \frac{1}{2} \sqrt{Ax}}, \quad (1.14)$$

которое стремится к нулю при $|x| \rightarrow \infty$.

Отсюда видно, что с увеличением поверхностного натяжения амплитуда уединенной волны

$$* \frac{3(1-\nu)}{2\nu-1-3(1-\nu)}$$

растет.

4°. Случай $\nu < 0$, как легко проверить, физического смысла не имеет.

Действительно, рассматривается асимптотическая теория, причем заранее предполагается, что решение удовлетворяет условиям (1.5). Из (1.13) и (1.14) следует, что сделанные априорные оценки имеют место, если ν близко к 1. Таким образом, в качестве ϵ надо принять величину $\sqrt{|1-\nu|}$, и приведенное рассмотрение имеет смысл только в окрестности точки $\nu = 1$.

* В строгой (не асимптотической постановке) задача о волнах может быть сведена к операторному уравнению с оператором, у которого производная Фреше обращается в нуль при $\nu = 1$ и $\epsilon = 0$.

5°. Выше рассмотрен случай только малых значений ν . Силы поверхностного натяжения только изменяют параметры волн, но не приводят к появлению новых форм движения. Предположим теперь, что $\nu > \frac{2}{3}$. При переходе параметра ν через свое бифуркационное значение $\frac{2}{3}$ изменяются знаки у коэффициентов A , B и D в уравнении (1.8). Проводя исследование фазовой плоскости аналогично тому, как было сделано выше, придем к следующим результатам.

а) Если $\nu < \frac{2}{3}$ и $\nu < 1$, то при $\nu = \frac{2}{3}$ волны вырождаются в равномерный поток, т. е. при $\nu = \frac{2}{3}$ амплитуда волн обращается в нуль.

б) Если $\nu > \frac{2}{3}$ и $\nu > 1$, то при $\nu = \frac{2}{3}$ амплитуда имеет конечный предел, это предельное течение представляет собой уединенную волну. Однако эти волны имеют форму не горба, а впадины.

Итак, если параметры ν и ν принадлежат областям 3 и 4 (фиг. 5), то при $\nu = \frac{2}{3}$ существует единственная предельная форма движения — равномерный поток; если параметры ν и ν принадлежат областям 1 и 2, то при $\nu = \frac{2}{3}$ существуют две предельные формы движения, кроме равномерного потока существует еще движение, при котором свободная поверхность имеет форму волны, но для значений параметров, принадлежащих области 1, уединенная волна имеет форму горба, а для области 2 — впадины.

Примечание. Приведенный асимптотический анализ имеет смысл лишь в случае, когда $|A| \left| 1 - \frac{2}{3}\nu \right|^1$ — малая величина. Рассматривался тот случай, когда малость этой величины обеспечивалась близостью ν к 1 и ν к нулю. Однако эта величина может быть малой и в том случае, если велик знаменатель. Это дает новый тип движения жидкости, который может быть изучен асимптотическими методами для любого значения ν . Пусть, например, $\nu = 0$, т. е. рассматриваются чисто капиллярные волны.

Тогда

$$\frac{A}{1 - \frac{2}{3}\nu} = \frac{2}{3(1 - \frac{2}{3}\nu)^2}.$$

Уравнение (1.8) допускает решение типа уединенной волны, если $D = 0$, т. е. если $\nu = 0$. Но в этом случае и $A = 0$, однако уравнение (1.9) здесь (при $\nu = \frac{2}{3}$) просто не имеет ограниченных решений, кроме тривиального, т. е. чисто капиллярные течения не содержат уединенной волны.

§ 2. Осесимметричные струи

1°. Будем рассматривать осесимметричную струю, подверженную действию сил поверхностного натяжения; течение будем считать установившимся. Цель рассмотрения — изучить возможные формы струи. Уравнение границы струи будем задавать в виде $r = f(z)$. Течение струи предполагается потенциальным. Поскольку струя предполагается осесимметричной, то задача сводится к отысканию функции тока $\psi(r, z)$, удовлетворяющей уравнению

$$\Delta \psi = -\frac{1}{r} \frac{\partial \psi}{\partial r} = 0, \quad (2.1)$$

и функции $f(z)$ по следующим условиям:

$$\psi = 1 \text{ при } r = f(z), \quad \psi = 0 \text{ при } r = 0, \quad (2.2)$$

$$f''^2 \left[\frac{\partial^2 \psi}{\partial r^2} + \frac{\partial^2 \psi}{\partial z^2} \right] = 2K - C \text{ при } r = f(z). \quad (2.3)$$

Условие (2.2) — это кинематическое условие (в данном случае — условие постоянства расхода), (2.3) — динамическое. Задача записана в безразмерных переменных. В качестве характерных параметров принят радиус струи R в некотором сечении, положение которого будет установлено ниже, и значение функции тока на поверхности струи Q , C — постоянная энергия, которая является функционалом, K — средняя кривизна поверхности струи

$$K = \frac{1}{2} \frac{1}{f \cos} - \frac{f}{(1 - f'^2)^{3/2}} = \frac{1}{2} \frac{(1 - f'^2)^{1/2}}{f} - \frac{f}{(1 - f'^2)^{3/2}}. \quad (2.4)$$

Здесь γ — безразмерный параметр, характеризующий действие сил поверхностного натяжения

$$\frac{\gamma}{RV^2}, \text{ или } \frac{R^3}{Q^2} = V \frac{Q}{R^2}. \quad (2.5)$$

Отсюда видно, что роль поверхностного натяжения будет тем больше, чем больше коэффициент поверхностного натяжения γ и чем меньше толщина струи и характерная скорость частиц жидкости.

2°. Функция v_r , удовлетворяющая (2.1) внутри полосы $0 < r < f(z)$ и обращающаяся в единицу на границе $r = f(z)$ и в нуль на прямой $r = 0$, имеет следующее асимптотическое представление (см. [1], с. 185):

$$v_r \sim \frac{r^2}{f^2} - \frac{3f^2}{4f^2} \frac{f}{f} r^2 + \frac{r^2}{f^2} \dots \quad (2.6)$$

и аналогично

$$v_z = \frac{1}{r} \frac{r}{z} \sim \frac{2rf}{f^3} \dots, \quad v_z = \frac{1}{r} \frac{r}{r} \sim \frac{2}{f^2} - \frac{3f^2}{2f^2} \frac{f}{f} + \frac{2r^2}{f^2} \dots \quad (2.7)$$

и далее

$$v_r^2 = v_z^2 = \frac{1}{f^2} \left(\frac{r}{r} - \frac{z}{z} \right) \sim \frac{4r^2 f^2}{f^6} - \frac{4}{f^4} + \frac{6f^2}{f^4} - \frac{2ff}{f^4} + \frac{2r^2}{f^2} \dots \quad (2.8)$$

Так как граница неизвестна, то следует задаться некоторыми ее априорными оценками. Имея в виду те же цели, что и в предыдущем параграфе, примем снова оценки М.А. Лаврентьева

$$f = O\left(\frac{1}{2}\right), \quad f = O\left(\frac{1}{2}\right) \text{ и т. д.} \quad (2.9)$$

Справедливость этих оценок, а также смысл параметра α установим ниже. Используя оценки (2.9), перепишем формулы (2.7) и (2.8) в виде

$$\begin{aligned} \frac{r^2}{f^2} - \frac{f}{4f^2} r^2 + \frac{r^2}{f^2} &= O\left(\frac{1}{2}\right), \\ (v_r^2 - v_z^2)_{r,f} &= \frac{4}{f^2} + \frac{1}{2} ff = O\left(\frac{1}{2}\right). \end{aligned} \quad (2.10)$$

В соответствии с оценками (2.9) преобразуем (2.4)

$$K = \frac{1}{2}(f^{-1} - f) = O\left(\frac{1}{3}\right). \quad (2.11)$$

Подставляя (2.10) и (2.11) в динамическое условие (2.3), получим обыкновенное дифференциальное уравнение второго порядка относительно свободной границы — функции f

$$(2 - f^3)f' - 4f^{-1} - f^2 = Cf^3 = 0. \quad (2.12)$$

Заметим, что уравнение (2.12) допускает первый интеграл, и его свойства легко изучить методами фазовой плоскости.

Для наших целей достаточно рассмотреть квадратичное приближение. Для этого положим $f = 1$ и удержим в (2.12) члены порядка $O(\epsilon^2)$.

Уравнение (2.12) примет следующий вид:

$$(2.13) \quad A \dot{z} + B z + B^2 z^2 + D z^3 = 0,$$

где постоянные A , B и D выражаются формулами

$$(2.14) \quad \begin{aligned} A &= D - 4\epsilon^2 - 3C, & D &= C - 4, \\ B &= D - 2C - (3 - 4C)\epsilon^2 D, & &= 3/(2 - \epsilon^2). \end{aligned}$$

3°. В случае равномерного потока и уединенной волны имеем $\epsilon = 0$ и $\epsilon = 0$ при $|z| \ll 1$. Следовательно, в этом случае $D = 0$ и $C = 4 - \epsilon^2$. Рассмотрим течения, близкие к этим. Положим $C = 4 - \epsilon^2 + \delta$, а $D = \delta$ будем считать малой величиной. В этом случае

$$A = 16 - (3 - \delta)\epsilon^2.$$

Так как излагаемая теория имеет смысл при малых A , то положим $16 - (3 - \delta)\epsilon^2 = \delta$, тогда $\delta = \frac{24}{7} - \frac{3}{98}\epsilon^2 = O(\epsilon^2)$ и далее

$$A = \frac{3}{7} + O(\max(\epsilon^2, \delta^2)), \quad B = 13\frac{3}{7} + O(\max(|\delta|, |\epsilon^2|)).$$

Таким образом, для малых ϵ и δ имеем $b > 0$, а знак A определяется соотношением δ и ϵ^2 . Замечая также, что $2 - \epsilon^2 < 0$ для малых ϵ , и повторяя рассуждения предыдущего параграфа, можем легко получить следующие результаты.

а) Если $\delta > 0$, т. е. если постоянная энергия $C = 12 - \delta$, то уравнение (2.13) не имеет ограниченных решений. Это значит, что струя, простирающаяся в бесконечность в обе стороны, неограниченно расширяется хотя бы в одном из направлений.

б) Если $\delta < 0$, то уравнение (2.13) допускает решения, ограниченные на всей прямой $z \in \mathbb{R}$. Если $\delta < 0$, то эти решения представляют собой нелинейные волны, амплитуды которых уменьшаются вместе с ϵ .

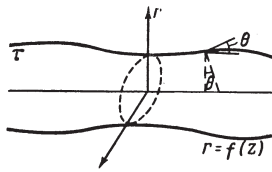
в) Характер предельных решений при $\epsilon \rightarrow 0$ (т. е. при $\epsilon \rightarrow 0$) будет существенно зависеть от ϵ . Если $\epsilon > 0$ (т. е. $\epsilon > 16$), то

$$\lim_{\epsilon \rightarrow 0} \frac{A}{2} = 0 \text{ при } \epsilon > 0$$

и амплитуда предельного решения не равна нулю. Предельное решение представляет собой уединенную волну типа впадины. Если $\epsilon < 0$, то предельное решение представляет собой равномерный поток.

§ 3. Нестационарные задачи теории струй

1°. Составим уравнения длинных осесимметричных волн, распространяющихся вдоль цилиндрической струи идеальной невесомой жидкости, подверженной действию сил капиллярности. Потенциал скоростей ϕ и форма свободной поверхности $r = f(z, t)$ определяется как решение следующей задачи (фиг. 6):



Фиг. 6

$$\frac{1}{r} \frac{\partial \phi}{\partial r} = \frac{1}{z} \frac{\partial \phi}{\partial z} = 0, \quad V_r = \frac{\partial \phi}{\partial r} = 0 \text{ при } r = 0, \quad (3.1)$$

$$\frac{\partial \phi}{\partial t} = \frac{1}{r} \frac{\partial \phi}{\partial r} \frac{df}{dz} \text{ при } r = f, \quad (3.2)$$

$$\frac{\partial \phi}{\partial t} = \frac{1}{2} \frac{\partial \phi}{\partial r} \frac{df}{dz} = \frac{1}{2} \frac{\partial \phi}{\partial z} \frac{df}{dz} = 2K C \text{ при } r = f. \quad (3.3)$$

Безразмерные переменные здесь введены так же, как и в предыдущем параграфе.

2°. Рассмотрим сначала вспомогательную задачу определения потенциала скоростей, удовлетворяющего уравнению (3.1), условию (3.2) и условию

$$\phi(f, z, t) = a(z, t). \quad (3.4)$$

Обозначая через ϵ малый параметр, сделаем в (3.1) замену переменного $z = \epsilon \zeta$.

Тогда уравнение (3.1) примет вид

$$\frac{1}{r} \frac{d}{dz} \left(r \frac{d\psi}{dz} \right) + \psi = 0. \quad (3.5)$$

Решение (3.5) будем искать в виде ряда

$$\psi = \psi_0 + \psi_1 + \dots, \quad (3.6)$$

где ψ_i удовлетворяют уравнениям

$$\frac{1}{r} \frac{d}{dz} \left(r \frac{d\psi_i}{dz} \right) + \psi_i = 0, \quad \psi_0 = \frac{1}{r}, \quad \psi_1 = \frac{0}{2}, \dots \quad (3.7)$$

Функции ψ_i удовлетворяют следующим краевым условиям:

$$\psi_i / r = 0 \text{ при } r = 0, \quad \psi_0 = a \text{ при } r = f, \quad \psi_i = 0 \text{ (} i > 0 \text{) при } r = f.$$

Решая эту последовательность краевых задач и возвращаясь снова к переменной z , получим следующее представление для ψ через его граничные значения:

$$\psi \sim a(z, t) \frac{1}{4} a_{zz} (f^2 - r^2) \dots \quad (3.8)$$

Для того чтобы (3.8) имело смысл, необходимо и достаточно, чтобы производные потенциала a и функции f достаточно быстро убывали вместе с их номером. Это условие имеет место, если волны достаточно длинные. Используя представление потенциала (3.8), можем преобразовать кинематическое и динамическое соотношения (3.2) и (3.3)

$$\frac{f}{t} = \frac{1}{4} a_{zz} f - f_z a_z \dots, \quad \frac{a}{t} = \frac{1}{2} a_z^2 - \frac{1}{f} f_{zz} \dots = C. \quad (3.9)$$

Если в равенствах (3.9) ограничиться определенным числом членов, то получим уравнения, которые дают приближенное описание распространению осесимметричных нелинейных волн вдоль струи.

3°. Уравнения (3.9) могут быть использованы для изучения разнообразных явлений, возникающих в теории струй. В частности, легко может быть изучена устойчивость струи относительно длинных возмущений. Плато [6], по-видимому, был первым, кто обратил внимание на то, что силы поверхностного натяжения могут служить причиной неустойчивости струи. Подробное исследование этого

обстоятельства было проведено Релеем [7], который показал, что цилиндрическая форма струи неустойчива относительно достаточно длинных возмущений. Этот факт следует из системы (3.9).

Положим $a = V_z$, $b = f - 1$ и линеаризуем систему (3.9), в которой удержим производные, порядок которых не выше второго

$$i \frac{1}{4} b_{zz} V_z, \quad b_i = 2b_z V \quad (\quad_{zz}) = 0. \quad (3.10)$$

Будем разыскивать решение типа бегущих волн. Для этого положим

$$Ae^{i z - t}, \quad b = Be^{i z - t}. \quad (3.11)$$

Для определения A и B имеем уравнения

$$A(-iV) - \frac{1}{4} b^2 B = 0, \quad A(1 - b^2) - B(2i - V) = 0. \quad (3.12)$$

Заметим, что если $b = 0$, т. е. поверхностное натяжение отсутствует, то второе уравнение (3.12) немедленно дает $b = 2i - V$, т. е. при отсутствии сил поверхностного натяжения струя находится в безразличном равновесии. Если добавить силы поверхностного натяжения, то струя потеряет устойчивость. В самом деле, характеристическое уравнение системы (3.12) имеет вид

$$b^2 - 3i - V \quad (\quad \frac{1}{4} b^2 (1 - b^2) = 0). \quad (3.13)$$

Отсюда сразу следует, что один из корней (3.13) имеет положительную действительную часть. Таким образом, силы поверхностного натяжения разрывают тонкую струю. Интегрируя систему (3.9), можно проследить процесс разрыва струи на капли. Изложенная теория имеет смысл только для длинных возмущений. Итак, тривиальная форма струи невесомой жидкости, подверженная действию сил поверхностного натяжения, неустойчива. Однако, как видно из предыдущего параграфа, возможны и другие формы осесимметричной струи. В частности, могут существовать волнообразные формы. Поэтому естественным образом возникает вопрос: не будут ли существовать среди этих возможных форм струйных течений устойчивые течения?

При подготовке рукописи автор воспользовался рядом замечаний Л.Н. Сретенского и Ф.Л. Черноусько, которых он искренне благодарит.

Литература

1. *Моисеев Н.Н.* Асимптотические методы типа узких полос./ В кн. «Некоторые проблемы математики и механики», Новосибирск, 1961.

2. *Моисеев Н.Н.* и *Тер-Крикоров А.М.* Исследование движения тяжелой жидкости при скоростях, близких к критической./Труды Моск. физико-техн. ин-та, 1959, № 3.

3. *Litman W.* On the existence of periodic waves near critical speed. *Communs Pure and Appl. Math.*, 1957, vol. 10, p. 241.

4. *Kortweg G., De Vries.* On the change of form of long waves advancing in a rectangular canal on a new type of long solitary waves. *Philos. Mag., Ser. 5*, 1895, 39, p. 422.

5. *Секерж-Зенькович Я.И.* К теории капиллярно-гравитационной уединенной волны./Сб. аннотаций II Всес. съезда по теор. и прикл. мех., 1964.

6. *Plateau.* *Statique experimentale et thйorique des liquids soumis aux seules forces moleculaires.* Paris, 1873.

7. *Reley.* On the instability of jets. *Proc. Math. Soc.*, 1878, vol. 10, p. 7.

1965 г.

Асимптотика быстрых вращений

§ 1. Введение

1. Предположим, что колебательная система описывается дифференциальным уравнением второго порядка,

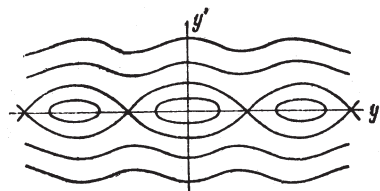
$$y'' + f(y, \epsilon) = F(y, y', t), \quad (1.1)$$

где ϵ — малый параметр, t — медленное время.

Уравнение

$$y'' + f(y, \epsilon) = 0 \quad (1.2)$$

будем называть порождающим, если ϵ — некоторое фиксированное число. Функцию $f(y, \epsilon)$ будем предполагать такой, что уравнение



(1.2) имеет периодическую фазовую плоскость периода 2. На рисунке изображена фазовая плоскость, причем внутри одного периода имеется только одно положение равновесия. Для исследования движений с большой энергией такое предположение не является существенным.

Из периодичности фазовой скорости следует, что $f(y, \epsilon)$ периодическая по y периода 2 и что при любом

$$\bar{f}(\epsilon) = \frac{1}{2} \int_{-y}^{y} f(x, \epsilon) dx > 0.$$

Фазовые траектории, заключенные внутри петли сепаратрисы, соответствуют колебательным движениям системы и определяются малыми начальными энергиями. Фазовые траектории, расположенные вне сепаратрисы, соответствуют вращательным движениям системы и определяются большими начальными энергиями.

Совокупность решений уравнения (1.1) естественно рассматривать зависящей от двух параметров, один из которых — параметр ϵ , а другой параметр — это величина начальной энергии. Подобная точка зрения использовалась И.Г. Малкиным при исследовании систем, близких к системам Ляпунова [1]. Предполагая малость начальной энергии, он строил общий интеграл порождающего уравнения методом Ляпунова. Опираясь на построенное решение порождающего уравнения, И.Г. Малкин развил специальный вариант метода Пуанкаре, который позволяет исследовать для малых ϵ уравнение (1.1). Позднее подобная идея использовалась Г.Е. Кузмаком и Ю.А. Митропольским при исследовании систем с медленным временем [2], [3]. В этих работах предполагалось сочетать методы построения решения порождающего уравнения в виде ряда с асимптотическим исследованием по ϵ .

Поскольку асимптотические методы осреднения в практических задачах представляются наиболее простыми и экономичными, то исследование системы (1.1) в случае малых начальных энергий целесообразно проводить по следующей схеме:

- 1) рассматривая порождающее уравнение как квазилинейное, построить его общий интеграл;
- 2) опираясь на это решение, привести систему (1.1) к стандартному виду:

$$\dot{x} = X(x, y, t, \epsilon), \quad (1.3)$$

$$\dot{y} = Y_0(x, t) + \epsilon Y_1(x, y, t, \epsilon)$$

и известным способом построить ее асимптотическое решение.

Следуя указанной схеме, можно исследовать почти всю область колебательных движений, за исключением области, примыкающей изнутри к сепаратрисе, поскольку в этой области период одного колебания становится большим и стремится к бесконечности, когда начальная энергия стремится к некоторому критическому значению, соответствующему движению по сепаратрисе.

Примечание. В прикладных работах при изучении малых колебаний системы (1.1) обычно линеаризуют левую часть уравнения, относя нелинейные слагаемые к функции F , с последующим использованием (в той или иной форме) метода малого параметра. Таким

способом можно получить только те решения, которые при $\omega = 0$ переходят в тривиальное решение уравнения (1.2). Рассматривая уравнение (1.1) как двухпараметрическое, мы можем отыскивать решения системы, близкие к нетривиальным решениям уравнения (1.2). Нетрудно привести примеры, когда уравнение (1.2) не содержит, например, периодических решений 1-го типа, но имеет периодические решения 2-го типа. Таким образом, рассмотрение системы (1.1) как системы, зависящей от двух параметров, позволяет исследовать ее свойства с несравненно большей полнотой.

2. Настоящая работа посвящена исследованию другого предельного случая — случая больших начальных энергий. Оказывается, что, как и в случае малых энергий, здесь также удастся построить асимптотические решения порождающего уравнения. Если решение для малых энергий можно представить в виде отрезка ряда, расположенного по положительным степеням корня из энергии, то для больших энергий решение представляется в виде ряда по отрицательным степеням корня из энергии.

Как только решение порождающего уравнения построено, мы имеем возможность привести систему (1.1) к виду (1.3) и использовать схему, которая была изложена выше.

Таким образом, в работе показывается возможность эффективного исследования всей фазовой плоскости за исключением полосы, примыкающей к сепаратрисе.

§ 2. Построение асимптотики порождающего решения

1. Будем рассматривать неограниченные движения системы (1.2). Положим

$$y = \omega z, \quad (2.1)$$

где ω — значение величины y в начальный момент, z — новое переменное. После замены уравнение (1.2) примет вид

$$\dot{z} = f(y). \quad (2.2)$$

Система (2.1)—(2.2) эквивалентна уравнению (1.2); так как оно не содержит времени, то, не ограничивая общности, рассмотрим для системы (2.1)—(2.2) следующую задачу Коши:

$$\begin{aligned} y(0) &= 0, \\ z(0) &= 0. \end{aligned}$$

Будем считать величину ϵ большой, и введем обозначение $1/\epsilon$. После замены независимого переменного $t = s/\epsilon$ система уравнений (2.1)—(2.2) будет приведена к виду

$$\begin{aligned} \dot{z} &= f(y), \\ \dot{y} &= 1 - z, \end{aligned} \quad \dot{z} = \frac{dz}{ds}. \quad (2.3)$$

Система уравнений (2.3) имеет стандартную форму. Следуя общей схеме, сделаем замену переменных:

$$\begin{aligned} z &= u_1(s, \epsilon) + \epsilon^2 u_2(s, \epsilon) + \dots, \\ y &= v_1(s, \epsilon) + \epsilon^2 v_2(s, \epsilon) + \dots, \end{aligned} \quad (2.4)$$

где новые неизвестные u_i и v_i удовлетворяют следующим дифференциальным уравнениям:

$$\begin{aligned} \dot{u}_1 &= A_1(s) + \epsilon^2 A_2(s) + \dots, \\ \dot{v}_1 &= B_1(s) + \epsilon^2 B_2(s) + \dots \end{aligned} \quad (2.5)$$

Потребуем, чтобы u_i и v_i были ограниченными функциями s , и кроме того пусть $u_i(s, 0) = v_i(s, 0) = 0$.

Перейдем к определению неизвестных, входящих в выражения (2.4) и (2.5). Используя (2.5), продифференцируем (2.4). Подставляя затем полученные ряды в систему (2.3) и сравнивая коэффициенты при одинаковых степенях ϵ , мы придем к следующим уравнениям:

$$A_1 u_1 = f(s), \quad (2.6)$$

$$B_1 v_1 = 1, \quad (2.7)$$

$$A_2 u_1 + A_1 u_2 + B_1 v_2 = f(s) v_1, \quad (2.8)$$

$$B_2 v_1 + A_1 v_2 + B_1 v_3 = u_1. \quad (2.9)$$

Рассмотрим уравнение (2.6). Для того чтобы u_1 было ограничено при $x \rightarrow \infty$, необходимо

$$A_1 \bar{f} = 0.$$

Таким образом,

$$u_1(x, y) = \int_0^x f(\xi) d\xi + C(x),$$

где $C(x)$ — произвольная функция. Так как $u_1(x, 0) = 0$, то

$$C(x) = 0.$$

Введем обозначения

$$\int_0^x f(\xi) d\xi = F(x), \tag{2.10}$$

$$\int_0^x g(\xi) d\xi = G(x).$$

Тогда $u_1(x, y) = F(x)$.

Рассмотрим теперь уравнение (2.7). На основании того же требования ограниченности

$$B_1$$

и, следовательно, $v_1 = 0$, т. е. $v_1 = 0$.

Таким образом, уравнение (2.8) перепишется так:

$$u_2 = f(x) + A_2.$$

Отсюда

$$A_2 \bar{f} = 0, \\ u_2 = f(x) + A_2.$$

Аналогично из уравнения (2.9) получаем

$$B_2 = \dots,$$

$$v_2 = \dots$$

и т. д.

Если ограничиться в разложениях (2.4) и (2.5) членами порядка n , то

$$z^n = 0,$$

откуда, в силу выбора начальных условий,

$$z = 0.$$

Таким образом,

$$z = \dots \quad (2.11)$$

Аналогично,

$$y = \dots,$$

т. е.

$$s^2 = \dots, \quad (2.12)$$

$$y = \dots$$

Возвращаясь к переменному t и полагая $s = 1/t$, получим

$$z = \frac{1}{t} = \frac{1}{t}, \quad (2.13)$$

$$y = \frac{1}{t} = \frac{1}{t} = \frac{1}{t} = \frac{1}{t}.$$

Продолжив эту процедуру, мы можем получить решение с любой точностью относительно $1/t$.

Так как $\bar{\omega}$ — это постоянная, не зависящая от t , то удобно ввести новую постоянную:

$$\frac{1}{\bar{\omega}}.$$

Тогда асимптотическое представление общего решения уравнения для больших энергий имеет вид

$$y \approx (t - t_0)^{-\frac{1}{2}} \left[\bar{\omega}^{-1} (t - t_0)^{\frac{1}{2}} \{ (t - t_0) \} \right] O \left(\frac{1}{3} \right). \quad (2.14)$$

Решение (2.14) содержит две произвольные постоянные t_0 и $\bar{\omega}$.

Примечание. Формула (2.14) сохраняет свою силу и в том случае, когда функция $f(y)$ не имеет производных. Она может быть получена методом последовательных приближений. Для этого вместо (2.4) и (2.5) положим

$$\begin{aligned} z &= u(t, y), \\ y &= v(t, y), \end{aligned} \quad (2.4)$$

$$\begin{aligned} A &= A(t, y), \\ 1 &= B(t, y). \end{aligned} \quad (2.5)$$

Подставляя разложения (2.4) и (2.5) в исходные уравнения, получим

$$A \frac{du}{dt} + A \frac{dv}{dt} + (1 - B) \frac{dv}{dt} = f(v),$$

$$B \frac{dv}{dt} + A \frac{dv}{dt} + (1 - B) \frac{dv}{dt} = u.$$

Эту систему будем решать по следующей итерационной схеме:

$$u^{(k)} = f(v^{(k-1)}) - u^{(k-1)} A^{(k-1)} - u^{(k-1)} B^{(k-1)} A^{(k)},$$

$$u^{(k)} = u^{(k-1)} - v^{(k-1)} A^{(k-1)} - v^{(k-1)} B^{(k-1)} B^{(k)}.$$

Нетрудно убедиться, что, используя начальные условия, условия ограниченности функций u и v и проведя два шага итерационного процесса, мы придем к формуле (2.14).

2. Рассмотрим теперь некоторые примеры:

$f(y) = \sin y$; тогда $(y) = 1 - \cos y$,

$y = 0$; $(y) = \sin y$.

Поэтому

$$y = t - \frac{1}{2} \sin t. \quad (2.15)$$

В рассматриваемом случае решение уравнения (1.2) выражается через эллиптические функции, и асимптотику (2.15) можно построить непосредственно.

3. Пусть

$$f(y) = \begin{cases} 1 - y & [0, 1], \\ 1 - y & [1, 0]. \end{cases}$$

Тогда

$$(y) = \begin{cases} y, & [0, 1], \\ 1 - y, & [1, 0], \end{cases} \quad - \frac{1}{2};$$

$$(y) = \begin{cases} \frac{1}{2} - y, & [0, 1], \\ \frac{1}{2} - y, & [1, 0]. \end{cases}$$

Таким образом,

$$y = \begin{cases} t - \frac{1}{2} - \frac{1}{2} t - \frac{1}{2} t^2, & t = 0, -1, \\ t - \frac{1}{2} - \frac{1}{2} t - \frac{1}{2} t^2, & t = -1, 0. \end{cases}$$

Примечание. Построение приближенного решения в случае больших энергий может быть осуществлено и другими способами,

отличными от изложенного. Например, положив $y = t$, $t = s$, найдем, что функция удовлетворяет уравнению

$$\frac{d^2}{ds^2} f(s) = f(s);$$

так как правая часть зависит аналитически от параметра s , то решение можно искать в виде ряда. Например, в том случае, когда $f(y) = \sin y$, решение (2.1) может быть представлено рядом

$$y = t \sin t - \frac{1}{2} \frac{5}{16} t^6 \dots - \sin 2 t - \frac{1}{8} t^4 \dots - \sin 3 t - \frac{1}{48} t^6 \dots$$

§ 3. Некоторые обобщения

Методика, изложенная в предыдущем параграфе, без каких-либо изменений может быть использована для асимптотического интегрирования уравнения

$$y' = f(x, y) = 0, \tag{3.1}$$

где $x = (x_1, \dots, x_n)$ — вектор, изменение которого описывается системой дифференциальных уравнений

$$\dot{x}_i = X_i(x_1, \dots, x_n, y), \quad i = 1, 2, \dots, n.$$

По-прежнему предполагается, что $f(x, y)$ периодична по y , $\bar{f}(x) = 0$ для любых x и $y(0) = \bar{y}$, где \bar{y} — большое число.

Систему (3.1)—(3.2) для достаточно больших \bar{y} и будем называть системой с вращающимся звеном. После замены $t = s$, $1/\bar{y}$, $y = z$ система (3.2) будет приведена к виду

$$\begin{aligned} \dot{x}_i &= X_i, \\ \dot{z} &= f, \\ \dot{y} &= 1 - z. \end{aligned} \tag{3.3}$$

Рассмотрим для этой системы задачу Коши:

$$\begin{aligned}x_i(0) &= x_i, \\z(0) &= 0, \\y(0) &= y_0.\end{aligned}$$

Сделаем стандартную замену переменных:

$$\begin{aligned}x_i &= x_i + \int_0^t \sum_{k=1}^n w_{ik}^k(x_1, \dots, x_n, z, y) dt, \\z &= \int_0^t \sum_{k=1}^n u_k^k(x_1, \dots, x_n, z, y) dt, \\y &= \int_0^t \sum_{k=1}^n v_k^k(x_1, \dots, x_n, z, y) dt, \\C_{ik} &= \int_0^t \sum_{k=1}^n C_{ik}^k(x_1, \dots, x_n, z, y) dt, \\A_{ik} &= \int_0^t \sum_{k=1}^n A_{ik}^k(x_1, \dots, x_n, z, y) dt, \\B_k &= \int_0^t \sum_{k=1}^n B_k^k(x_1, \dots, x_n, z, y) dt.\end{aligned}$$

Накладывая дополнительные требования

$$\begin{aligned}u_k(x_1, \dots, x_n, 0, y) &= 0, \\v_k(x_1, \dots, x_n, 0, y) &= 0, \\w_{ik}(x_1, \dots, x_n, 0, y) &= 0\end{aligned}$$

и повторяя дословно все рассуждения предыдущего параграфа, мы получим

$$\begin{aligned}C_{i1} &= \bar{X}_i, \quad w_{i1} = \int_0^t (X_i - \bar{X}) d, \\A_1 &= 0, \quad u_1 = \sum_{j=1}^n (x_j - \bar{X}_j), \\B_1 &= 0, \quad v_1 = 0\end{aligned} \tag{3.4}$$

$$C_{i2} = \frac{X_i}{j} w_{j1} - \frac{w_{i1}}{j} \bar{X}_j,$$

$$u_{i2} = \frac{X_i w_{j1} - w_{i1} \bar{X}_j}{j} - (X_i - \bar{X}_i) C_{i2} d, \quad (3.5)$$

$$A_2 = \frac{f}{j} w_{j1} - \bar{X}_j,$$

$$u_2 = \frac{f}{j} w_{j1} - X_j - t A_2 d,$$

$$B_2 = \bar{\quad}, \quad v_2 = \bar{\quad}$$

и т. д.

Разумеется, в общем случае квадратуры в полученных формулах уже не берутся.

Частным случаем рассмотренной системы является уравнение

$$y' = f(y, t) = 0, \quad (3.6)$$

которое может быть приведено к виду (3.1) заменой

$t = x$.

Система (3.3) в этом случае имеет вид

$$\frac{dx}{ds} = \quad,$$

$$\frac{dz}{ds} = f(x, y),$$

$$\frac{dy}{ds} = 1 - z.$$

Формулы (3.4) и (3.5), соответственно, примут вид

$$C_1 = 1, \quad w_1 = 0, \quad A_1 = 0; \quad u_1 = \quad (\quad, \quad),$$

$$B_1 = \quad, \quad v_1 = 0, \quad C_2 = 0, \quad w_2 = 0,$$

$$A_2 = \quad, \quad u_2 = \quad \quad \quad, \quad B_2 = \quad, \quad v_2 = \quad.$$

Возвращаясь к старым переменным, получим

$$z(t) = \frac{1}{2} (t - t_0) \quad t, \quad t = \frac{1}{2} (t_0) t \quad ,$$

$$y(t) = \frac{1}{2} (t_0) t - \frac{1}{2} (t) t - \frac{1}{2} (t_0)$$

$$, \quad \frac{1}{2} (t_0) t \quad .$$

§ 4. Приведение системы (1.1) к стандартной форме

1. Предположим, что общий интеграл порождающего уравнения известен:

$$y = Q(c, \dots), \quad (4.1)$$

где $(c)(t - t_0)$, c, t_0 — произвольные постоянные. Тогда уравнение (1.1) может быть приведено к нормальной форме (1.3), если в качестве новых переменных принять величины $c(t)$ и (t) . Один из способов такой редукции приведен в [3], для общих гамильтоновых систем 2-го порядка этот вопрос рассмотрен в [4]. Поэтому здесь необходимые вычисления мы проведем очень кратко, имея в виду вывести только те формулы, которые будут использованы в дальнейшем.

Будем рассматривать неограниченные движения; имея в виду представление (2.14), положим

$$y = cZ(c, \dots), \quad (4.2)$$

где Z — периодическая функция периода 2. Кроме того, положим

$$y = (c, \dots) Q(c, \dots). \quad (4.3)$$

Дифференцируя (4.1) и приравнявая (4.3), получим

$$Q_c c = Q - Q = 0. \quad (4.4)$$

Дифференцируем (4.3) и подставляем в уравнение (1.1):

$$(cQ - Q_c)c - Q \quad (Q - Q) f(Q,) = F. \quad (4.5)$$

Разрешая систему (4.4)–(4.5) относительно c и Q и учитывая, что

$$2Q - f(Q,) = 0,$$

мы приходим к следующей системе двух уравнений относительно c и Q :

$$\begin{aligned} c - \{FQ - Q^2 - (Q - Q - Q - Q)\} &= X(c, Q,), \\ (c, Q) - \{FQ_c - Q - Q_c - cQ - Q\} &= Y(c, Q,), \end{aligned} \quad (4.6)$$

где

$$(Q_cQ - Q - cQ) = cQ^2. \quad (4.7)$$

Таким образом, уравнение (1.1) приведено к нормальной форме.

Величина c не зависит от Q (см. [3] и [4]).

2. Если в выражении (4.2) функцию $Z(Q, c)$ заменить асимптотическим представлением (2.14) и положить

$$t, \quad c = \frac{1}{2},$$

то

$$y = cZ(Q, c).$$

Асимптотическое представление величины c :

$$\sim \frac{1}{2} - \frac{1}{2} [ZZ - Z^2] - \frac{1}{2\sqrt{c^3}} \sqrt{c} [ZZ - Z^2], \quad (4.8)$$

откуда

$$\frac{1}{2} - \frac{1}{2} - O\left(\frac{1}{2}\right), \quad (4.9)$$

или, возвращаясь к переменному c ,

$$\frac{1}{2} \frac{1}{\sqrt{c^3}}. \tag{4.10}$$

3. Систему уравнений (4.6) можно исследовать асимптотическими методами. Делая замену (2.4):

$$c \quad \begin{matrix} u_1 \dots, \\ v_1 \dots, \end{matrix} \tag{4.11}$$

находим, что u и v удовлетворяют уравнениям

$$\begin{aligned} \frac{1}{2} \frac{d^2 X}{dt^2} &= \dots, \\ (c, \dots) \frac{1}{2} \frac{d^2 Y}{dt^2} &= O(\dots). \end{aligned} \tag{4.12}$$

В системах с медленно меняющимися параметрами рассматривается величина

$$J(c, \dots) = \frac{1}{2} \int_0^2 (c, \dots) Q^2 dt. \tag{4.13}$$

При $F=0$ она является адиабатическим инвариантом.

Вычислим $dJ/dt = J_c c + J_{\dots}$:

$$\frac{dJ}{dt} = \frac{1}{2} \int_0^2 \{Q^2 Q_c + 2 Q Q_c\} dt + c \frac{1}{2} \int_0^2 \{Q Q^2 + 2 Q Q_c\} dt. \tag{4.14}$$

Заметим, что

$$\int_0^2 2 Q Q_c dt = \int_0^2 [Q Q_c - Q_c Q] dt.$$

Поэтому первый интеграл в выражении (4.14) равен — c .

Преобразуя аналогично второй интеграл в выражении (4.14), мы замечаем, что первое из уравнений системы (4.12) можно переписать в виде

$$\frac{dJ}{dt} = \frac{1}{2} \int_0^2 FQ dt. \tag{4.15}$$

4. Приведенные вычисления позволяют установить связь между J и c :

$$J_c. \quad (4.16)$$

Отсюда сразу находим асимптотическое представление для величины J :

$$J \sim \frac{1}{\sqrt{c}} \sim O\left(\frac{1}{\sqrt{c}}\right). \quad (4.17)$$

§ 5. Некоторые примеры

1. Рассмотрим уравнение

$$y'' + g(y) \sin y = 0. \quad (5.1)$$

Порождающее решение имеет вид

$$y = t - \frac{g(t)}{2} \sin t. \quad (5.2)$$

Следовательно,

$$y'' = -c g \sin. \quad (5.3)$$

Из уравнения (4.15) и представления (4.17) следует, что $c = \text{const}$. Таким образом, если через a мы обозначим амплитуду — максимальное отклонение от равномерного вращения, то в рассматриваемом случае

$$a = \text{const} \cdot g(t). \quad (5.4)$$

Равенство (5.4) позволяет наглядно показать различие свойств движения с малой и большой энергиями. Если начальная энергия мала, то

$$J \sim a^2 = \text{const}, \quad (5.5)$$

т. е.

$$a \sim \frac{1}{\sqrt[4]{g(t)}}.$$

Таким образом, в случае малых колебаний с ростом возвращающей силы колебания около положения равновесия будут происходить со все уменьшающейся амплитудой. В случае вращательных движений амплитуда отклонений от режима установившегося вращения увеличивается вместе с увеличением возвращающей силы. Заметим, что угловая скорость вращения с точностью до членов порядка $O(1/3)$ не изменяется при изменении величины возвращающей силы.

2. Рассмотрим аналог уравнения Ван-дер-Поля:

$$\ddot{y} + \sin y = (1 - b \sin^2 y)y, \quad b > 0. \quad (5.6)$$

Уравнение для c :

$$c = \frac{1}{2} \int_0^{2\pi} (1 - b \sin^2 Q) (c) Q^2 dQ. \quad (5.7)$$

В нашем случае

$$\frac{1}{2} \int_0^{2\pi} 2\sqrt{c^3}, \quad Q = c \sin, \quad Q = 1 - c \cos.$$

В выражении для Q и Q удержаны члены самого низшего порядка по отношению к c . Поступая так же при вычислении интеграла в уравнении (4.7), получим

$$c = 1 - \frac{b}{2} - \frac{c^2 b}{2}. \quad (5.8)$$

Уравнение (5.8) позволяет вычислить стационарные амплитуды

$$c^2 = \frac{2 - b}{b}. \quad (5.9)$$

Так как при выводе формулы (5.9) мы считали амплитуду малой, то все рассуждения верны лишь при условии, что разность $2 - b$ мала.

Исследуя стандартным способом устойчивость решения (5.9) уравнения (5.8), убеждаемся в его неустойчивости. Это обстоятельство качественно отличает свойства решений уравнения (5.6) от его квазилинейного аналога

$$\ddot{y} + y = (1 - by^2)y, \quad (5.10)$$

так как известно, что его периодические решения всегда устойчивы.

Это различие — следствие того, что демпфирование в неограниченных движениях играет иную роль, чем в вибрационных движениях, происходящих в окрестности равновесия. В последнем случае демпфирование стремится вернуть систему в исходное состояние — состояние покоя. Что касается случая неограниченных движений, то демпфирование стремится разрушить исходное состояние — состояние равномерного вращения. Поэтому для построения правильной аналогии мы должны компенсировать ту составляющую демпфирования, которая определяет постоянное уменьшение угловой скорости. Это значит, что мы должны заменить (5.6) уравнением

$$\dot{y} + \sin y + (1 - b \sin^2 y)y = d, \quad (5.11)$$

где d — некоторое положительное число. Уравнение (5.11) уже будет обладать известными свойствами обычного уравнения Ван-дер-Поля; его периодическое решение будет устойчивым.

§ 6. Резонансные явления

1. Рассмотрим тот случай, когда функция F в уравнении (1.1) является периодической функцией времени периода T . Рассмотрим случай, когда

$$F = (t) 2v,$$

где v — некоторая постоянная, а (t) подобрано так, что функция (t) имеет период 2 по переменному ξ . Кроме того, примем, что

$$v > 0, \quad Z > 0.$$

Ограничиваясь случаем главного резонанса*, положим $t = h$. Тогда система уравнений (4.6) принимает вид

$$\begin{aligned} c &= -\{ (t) 2v + (1 - cZ) \} (1 - cZ), \\ h &= -\{ (t) 2v + (1 - cZ) \} Z. \end{aligned} \quad (6.1)$$

Обозначим через c_0 решение уравнения

* Более сложные задачи рассматриваются аналогично.

(c_0)

и будем изучать решения уравнения (6.1), предполагая, что

$$c = c_0 + O(\epsilon^k).$$

Тогда

$$(c_0)(c - c_0) = O(\epsilon^k), \quad k \geq 1. \quad (6.2)$$

Следовательно, вблизи значений $c = c_0$ функции c и h будут меняться медленно. На этом основании, используя принцип осреднения, перепишем систему (6.1) в виде

$$\begin{aligned} \dot{c} &= \frac{1}{2\pi} \int_0^{2\pi} \{c_0 + \epsilon(t)Z - 2\nu(c_0)(1 - c_0^2 Z^2)\} dt = O(\epsilon^2), \\ \dot{h} &= (c_0)(c - c_0) - \frac{1}{2\pi} \int_0^{2\pi} (t)Z dt = O(\epsilon^2). \end{aligned} \quad (6.3)$$

Система (6.3) получается из системы (6.1) осреднением по t на интервале $(0; 2\pi)$, поскольку правые части системы (6.1) являются периодическими функциями t периода 2π . При ее выводе мы используем условия $\bar{c} = 0, \bar{f} = 0, Z = 0$; кроме того, было принято во внимание, что

$$\frac{d}{dt} = O(\epsilon) \frac{d}{dt}.$$

Отбрасывая члены порядка ϵ^2 , мы приходим к «укороченной» системе. Поскольку мы рассматриваем движения с большей энергией (ϵ велико), то в первом из уравнений (6.3) мы отбросим слагаемое $c_0^2 Z^2$, поскольку оно имеет порядок $O(\epsilon^4)$ и мало по сравнению с единицей. Кроме того, заменим величину ϵ асимптотическим выражением (4.10). Теперь укороченные уравнения мы будем писать в виде

$$\begin{aligned} \dot{c} &= 2\sqrt{c_0^3} - c_0 \frac{Z}{\sqrt{c_0}}, \\ \dot{h} &= \frac{c - c_0}{2\sqrt{c_0^3}} - 2\sqrt{c_0^3} \frac{Z}{c_0}. \end{aligned} \quad (6.4)$$

В этих уравнениях принято обозначение

$$- \frac{2}{2} \int_0^2 (s) ds.$$

2. Возможные стационарные режимы определяются системой уравнений

$$\overline{Z} = \frac{2v}{\sqrt{c_0^3}}, \quad 2 \overline{Z} = \frac{c}{2c_0^3}. \quad (6.5)$$

Из первого уравнения мы находим h . Это уравнение является трансцендентным и может иметь много корней (или ни одного). Второе уравнение линейное и позволяет определить c .

Обозначим через c^* и h^* какое-либо решение системы (6.5) и изучим устойчивость этого решения.

Для этого составим уравнение в вариациях:

$$c = 2 \sqrt{c_0^5} \overline{Z} = h, \quad (6.6)$$

$$h = \frac{c}{2\sqrt{c_0^3}} = 2\sqrt{c_0^3} \overline{Z} = h. \quad (6.7)$$

Используя первое из уравнений (6.5), преобразуем (6.7):

$$h = \frac{c}{2\sqrt{c_0^3}} = 4 v h. \quad (6.8)$$

Составим характеристическое уравнение системы (6.6) и (6.8):

$$\lambda^2 - 4 v \sqrt{c_0} \overline{Z} = 0. \quad (6.9)$$

Мы видим, что коэффициент при первой степени λ не зависит от вида функций f и g и для случая демпфирования (γ и ν — одного знака) всегда отрицателен. Таким образом, любой резонансный режим является неустойчивым, если только в системе есть затухание.

Таким образом, мы пришли к результату, который был обнаружен Ф.Л. Черноусько в случае колебания и вращения маятника. Использование асимптотики больших энергий показывает его универсальность.

3. В качестве примера рассмотрим случай, когда

$$f(y) = \sin y, \quad g(t) = \sin t.$$

Тогда

$$\overline{Z} = \sin h, \quad \overline{Z} = \cos h, \quad \overline{Z} = \sin h$$

и далее

$$\overline{Z} = \frac{1}{2} \cosh h, \quad \overline{Z} = \frac{1}{2} \sinh h, \quad \overline{Z} = \frac{1}{2} \cosh h.$$

Поэтому система (6.5) будет иметь вид

$$\begin{aligned} \sin h &= \frac{4\nu}{\sqrt{c_0^3}}, \\ \cos h &= \frac{c - c_0}{2c_0^3}. \end{aligned} \tag{6.10}$$

Отсюда следует, что существуют только два резонансных режима c_1 и c_2 , причем в одном случае $c_1 - c_0 > 0$, а в другом $c_1 - c_0 < 0$. Используя второе уравнение (6.10), уравнение (6.9) можно переписать в виде

$$2 - 4\nu \frac{c - c_0}{4c_0^2} = 0. \tag{6.11}$$

Таким образом, в случае маятника колебание с амплитудой c_1 будет колебательно неустойчиво, а с амплитудой c_2 — аperiodически неустойчиво.

Заклучение

1. Основное содержание работы состоит в изложении теории, позволяющей изучить случай больших энергий с той же полнотой и столь же элементарно, как и случай колебаний, близких к квазилинейным. Изложенная теория может быть распространена на более общий случай колебания систем, содержащих вращающиеся звенья (см. [5]).

2. Изложение тех разделов, которые посвящены нерезонансным решениям, опираются на многочисленные работы, содержащие обоснование метода осреднения (см. [6], где приведена соответствующая библиография). Что касается рассуждений, связанных с исследованием резонансных явлений, то они носят интуитивный характер и требуют обоснования.

В.М. Волосов и Ф.Л. Черноусько прочли рукопись и сделали ряд замечаний. Ф.Л. Черноусько, кроме того, проверил все вычисления. Автор пользуется случаем выразить им свою искреннюю благодарность.

Литература

1. *Малкин И.Г.* Методы Ляпунова и Пуанкаре в теории нелинейных колебаний. – М.-Л., Гостехиздат, 1948.

2. *Кузмак Г.Е.* Асимптотические решения нелинейных дифференциальных уравнений второго порядка с переменными коэффициентами./Прикл. матем. и механ., 1959, 23, вып. 3, 515–526.

3. *Митропольский Ю.А.* Нестационарные процессы в нелинейных колебательных системах. – Киев, АН УССР, 1955.

4. *Черноусько Ф.Л.* О резонансе в существенно нелинейной системе./Ж. вычисл. матем. и матем. физ., 1963, 3, № 1, 131–144.

5. *Moiseyev N.N.* Methods of non-linear mechanics in the problems of the dynamics of satellites. Repts of the XIII International Congress of Astronautics, Varna, 1962.

6. *Волосов В.М.* Некоторые виды расчетов в теории нелинейных колебаний, связанные с усреднением./Ж. вычисл. матем. и матем. физ., 1963, 3, № 1, 3–53.

1962 г.

Часть 2

**Оптимизация, исследование операций
и теория управления**

Методы динамического программирования в теории оптимальных управлений

I. Системы, допускающие использование шкалы управлений

В послевоенные годы вариационное исчисление обогатилось рядом исследований, посвященных неклассическим задачам (см., например, [1]—[6]). Имеется также ряд работ, в которых предлагаются численные методы решения таких задач (см., например, [7], [8]). Использование вычислительных машин позволило решить много конкретных задач, возникших в технике. Для эффективного построения решений вариационную задачу обычно редуцируют (при помощи необходимых условий экстремума) к некоторой краевой задаче.

Данная работа посвящена прямому решению вариационных задач оптимального управления, без использования редукции к крайвым задачам. Термин «динамическое программирование» трактуется автором весьма произвольно, и обсуждаемая процедура непосредственной связи с процедурой Беллмана не имеет; она связана непосредственно с методами организованного перебора, используемыми при решении экономических задач (см. [9]), является их дальнейшим развитием и адаптацией к задачам механики управляемых систем.

§ 1. Задача о минимуме функционала

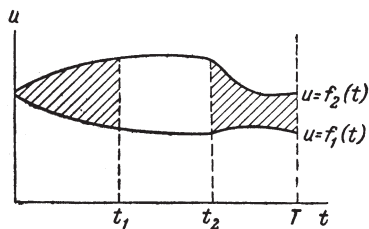
1. Рассмотрим задачу отыскания минимума функционала

$$I(u) = \int_0^T F(t, Lu) dt, \quad (1.1)$$

где L — некоторый оператор, u — функция t , которую мы условимся называть «управлением», и G — множество допустимых значений функции u . Оно может быть задано, например, следующим образом (см. фиг. 1): если $t \in [0, t_1)$ или $t \in (t_2, t]$ то значения $u(t)$ должны принадлежать либо заштрихованной области, либо ее границам f_1 и f_2 . Отрезок $[t_1, t_2]$ «неуправляемый», т. е. значения функции $u(t)$ однозначно определяются ее значением $u(t_1)$. Например, если

$f_1(t)$ и $f_2(t)$ то u , где — заданная постоянная. Если прямая $u(t_1) = t$ в некоторой точке достигает границы, то затем значения $u(t)$ равны граничным.

Построим множество точек $\{u_{ij}\}$ по следующему рецепту. Разобьем интервал $(0, t_1)$ на N частей точками $t = h, 2h, \dots$. На каждой из указанных прямых в области определения G отметим точки u_{ij} . Первый индекс будет обозначать номер вертикальной прямой. Вто-



Фиг. 1

рой индекс — номер точки на этой прямой. Систему точек $\{u_{ij}\}$ условимся называть, следуя Беллману, шкалой управлений.

Операцию, которая двум точкам $u_{ij}, u_{i-1,k}$ ставит в соответствие функцию $u_{ij}^{i-1,k}(t)$, будем называть A -операцией. В данной работе мы будем использовать полигональную

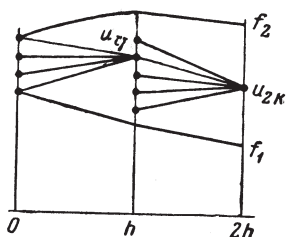
аппроксимацию, поэтому A -операция в нашем случае указанным точкам ставит в соответствие отрезок прямой

$$A(u_{ij}, u_{i-1,k}) = u_{ij}^{i-1,k} = u_{ij} + \frac{u_{ij} - u_{i-1,k}}{h}(t - ih). \quad (1.2)$$

Определим, кроме того, число

$$I_{ij}^{i-1,k} = \int_{ih}^{(i-1)h} F(t, Lu_{ij}^{i-1,k}) dt.$$

2. Перейдем к описанию алгоритма динамического программирования для случая ограничений описанного выше типа. На первом шаге фиксируем точку u_{1j} и при помощи A -операции соединим ее со всеми точками u_{0s} (см. фиг. 2). Для каждой из полученных функций u_{0s}^{1j} вычисляем I_{0s}^{1j} . Точке u_{1j} поставим в соответствие число



Фиг. 2

и управление, на котором достигается этот минимум. Повторим эту процедуру для каждой из точек множества $\{u_{1j}\}$.

На втором шаге отметим точку u_{2k} , лежащую на прямой $t = 2h$; при помощи A -операции построим функции u_{1j}^{2k} и вычислим числа I_{1j}^{2k} .

Составим выражение $I_{0,1j}^{2k} I_0^{1j} I_{1j}^{2k}$. Это число является функцией индекса j . Пусть

$$I_0^{2k} = \min_j I_{0,1j}^{2k}. \quad (1.3)$$

Точке u_{2k} для любого k поставим в соответствие число I_0^{2k} и управление u , на котором достигается* минимум (1.3).

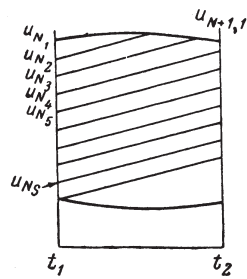
Примечание. Предположим теперь, что речь идет о минимизации функционала

$$I_{2h} = \int_0^{2h} F(t, Lu) dt.$$

Тогда для определения приближенного решения нам было бы достаточно найти число

$$I_{2h} = \min_k I_0^{2k}.$$

Используя описанную процедуру, каждой точке шкалы управлений u_{Ni} , лежащей на прямой $t = t_1$, мы поставили в соответствие число I_0^{Ni} и управление, на котором реализуется это значение функционала. Следующий отрезок времени является «неуправляемым». Каждая точка u_{Ni} определяет однозначно точку шкалы управлений, лежащую на прямой $t = t_2$, т. е. точку $u_{N-1,i}$ (см. фиг. 3). Например, если прямая $u = u_{Ns}$ ($t = t_1$) не пересекает границу, то эта точка определяет точку



Фиг. 3

$$u_{N-1,s} = u_{Ns}(t_2, t_1). \quad (1.4)$$

Каждой такой точке мы поставим в соответствие число

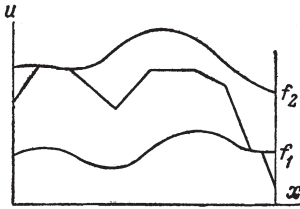
$$I_0^{N-1,s} = I_0^{Nr} \int_{t_1}^{t_2} F(t, L(u_{nr}(t, t_1))) dt.$$

* Идея такой организации перебора, насколько известно автору, впервые встречается в работе [6].

Вследствие условий неуправляемого участка, как это видно из фиг. 3, точка $f_1(t_2)$ может соответствовать нескольким точкам u_{Nr} и, следовательно, этой точке может соответствовать несколько значений функционала $I_0^{N-1,r}$. Из этих значений запоминаем наименьшее и то управление, на котором оно достигается.

Итак, на прямой $t = t_2$ мы получили некоторый набор точек $u_{N-1,r}$, каждой из которых поставлено в соответствие число $I_0^{N-1,r}$.

На участке (t_2, T) снова строим шкалу управлений и повторяем всю ту процедуру, которая была описана для интервала $(0, t_1)$. В результате этого мы получим на прямой $t = T$ точки u_{mi} , каждой из которых поставлено в соответствие значение функционала I_0^{Mi} . Искомое значение функционала мы получим из условия



Фиг. 4

$$I = \min_i I_0^{Mi}.$$

Функцию $u(t)$ мы получили в форме полигональной кривой.

3. Точность решения задачи определяется шагом h и числом точек шкалы управлений. Из-за ограниченности памяти машины мы не всегда сможем взять достаточное число точек в шкале управлений, и решение, которое мы получим, будет недостаточно точным. Покажем, что решение всегда может быть уточнено.

Обозначим через $u^{(1)}$ функцию, которую мы нашли (см. фиг. 4). Погрузим функцию $u^{(1)}$ в некоторую область, ограниченную кривыми $f_1^{(1)}$ и $f_2^{(1)}$. Область, ограниченная этими кривыми, должна целиком принадлежать области определения G . Так как значения функции $u^{(1)}$ могут совпадать с граничными, то и функции $f_1^{(1)}$ и $f_2^{(1)}$ могут иметь с функцией $u^{(1)}$ общие точки, как это показано на фиг. 4. Возьмем шаг h_1 меньший h , и построим шкалу управлений для более «узкой» области, сохранив в новой шкале управлений точки пересечения кривой $u^{(1)}$ с прямыми $t = nh_1$. Для этой шкалы управлений повторим описанную выше процедуру перебора. В результате мы получим новое значение функционала $I^{(2)}$ и новое решение $u^{(2)}$. Так как функция $u^{(1)}$ содержится среди ломаных, которые сравнивались, то $I^{(2)} \leq I^{(1)}$.

Этот процесс мы можем повторить любое число раз и получить последовательность значений функционала $I^{(k)}$ и управлений

$u^{(k)}(t)$, причем $I^{(k-1)} = I^{(k)}$. Таким образом, если функционал I ограничен снизу, то последовательность $\{I^k\}$ сходится. Следовательно, последовательность $\{u^{(k)}\}$ сходится в смысле слабой сходимости к функции u , которая может и не быть непрерывной.

Примечания. 1. Чем шире область определения функции u , тем более трудоемким становится процесс отыскания функции u . Наоборот, чем больше ограничений накладывается на область допустимых значений, тем экономнее становится процедура расчета. Задача становится особенно простой, когда управление носит релейный характер. В этом случае на каждом шаге шкала управлений состоит из двух точек и к сравнению допускаются только два управления. Это обстоятельство является важной особенностью данного метода.

2. Изложенная схема может быть распространена и на многомерные задачи. Пусть в выражении (1.1) функция u заменена конечномерным вектором. Этот случай отличается от рассмотренного только тем, что шкала управлений становится многомерной; все рассуждения сохраняют свою силу. Значительно более сложным является тот случай, когда многомерной становится область интегрирования.

3. В этой работе условия, которым должен удовлетворять оператор L , не обсуждались. Условимся, например, что функция F в выражении (1.1) имеет вид $F(t, u)$ или $F(t, u, \dot{u})$.

В более сложных случаях приходится вводить более сложную структуру A -операции.

В рамках излагаемой схемы может рассматриваться функционал вида

$$I(u) = \int_0^T F(Mx, Lu) dt,$$

где x — некоторый вектор, M — оператор. Однако если выражение Mx зависит также и от u , то рассуждения остаются справедливыми только в том случае, если выражение $M(x, u)$ не зависит от истории. Например, если функция x определена дифференциальным уравнением $\dot{x} = u$, то изложенная процедура смысла не имеет. В самом деле, сравнение при $u = u_{ij}$ различных вариантов управления невозможно, поскольку им будут соответствовать различные значения x .

§ 2. Краевые задачи для уравнения Лагранжа

1. Рассмотрим систему уравнений Лагранжа:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, 2, \dots, n, \quad (2.1)$$

где $I = L(q, \dot{q}, t)$ — обобщенная функция Лагранжа.

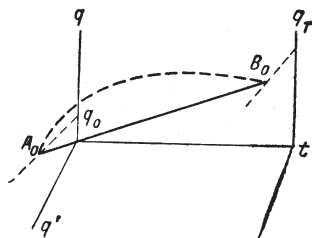
Пусть для уравнения (2.1) сформулирована следующая краевая задача: определить решение системы (2.1) — вектор $q(t)$ — по условиям

$$t = t_0, \quad q = q^{(0)}; \quad t = T, \quad q = q^{(T)}. \quad (2.2)$$

Из принципа Гамильтона—Остроградского следует, что краевая задача (2.2) эквивалентна следующей задаче вариационного исчисления: определить «управление» — вектор $q(t)$, который обращает в минимум функционал

$$W(q) = \int_0^T L(q, \dot{q}, t) dt. \quad (2.3)$$

Эта задача является частным случаем задачи, рассмотренной в § 1 данной работы, и ее решение может быть получено при помощи процедуры, которая там была описана.



Фиг. 5

2. Особенность рассматриваемой задачи состоит в том, что на управление — вектор q не накладывается никаких ограничений, кроме краевых условий. Поэтому в этой задаче трудно избежать итерационного процесса. Рассмотрим один из возможных вариантов на примере одного уравнения. Итак, пусть в системе (2.1) $n = 1$. Тогда задачу можно сформулировать так (см. фиг. 5):

в пространстве q, \dot{q}, t найти кривую, соединяющую прямые $t = 0, q = q^{(0)}$ и $t = T, q = q^{(T)}$, на которой функционал (2.3) достигает минимального значения.

Возьмем две произвольные точки A_0 и B_0 , лежащие на этих прямых, и соединим их прямой. Эту прямую мы примем в качестве нулевого приближения q_0 . Пусть $W_0 = W(q_0)$. Зададим шаг $h_1 = T/N_1$ и число N_1 . Шкалу управлений расположим на «прямых» $t = kh$, $k = 1, 2, \dots, N_1$. Множество $\{q_{ki}\}$ будет состоять из трех точек

$q_{k0}\{q_0(kh)\}$, $q_{k1}\{q_0(kh)\}$ и $q_{k2}\{q_0(kh)\}$. Таким образом, шкала управлений содержит точки пересечения прямой q_0 с прямыми $t = kh$. Под A -операцией будем понимать операцию проведения прямых (1.2), тогда

$$W_{ij}^{i-1,k} \underset{ih_1}{L} q_{ij} \frac{q_{i-1,k} q_{ij}}{h_1}(t_1 - ih_1); \frac{q_{i-1,k} q_{ij}}{h_1}; t dt.$$

Используя описанную процедуру перебора, мы найдем первое приближение — функцию $q_1(t)$ и $W_1 = W(q_1)$. Задавая новые значения $h = h_2$ и t_2 , построим новую шкалу управлений, следуя тому же способу, приняв в качестве точек q_{k0} точки с координатами $t = kh_2$, $q = q_1(kh_2)$.

Если число степеней свободы $n > 1$, то шкалу управлений мы будем размещать на гиперсферах радиуса r , центры которых расположены в точках предыдущего приближения.

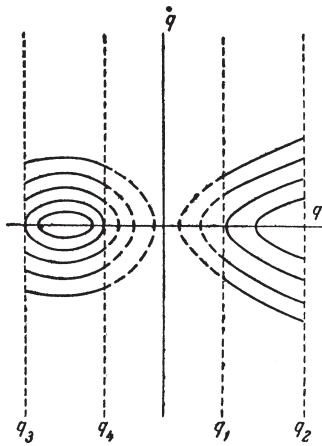
3. Изложенная схема расчета может быть использована в различных задачах механики, например при проектировании орбит импульсных перелетов в сложных гравитационных полях. Рассмотрим, например, задачу о перелете космического аппарата за заданное время T из точки $A(x_0, y_0, z_0)$ в точку $B(x_T, y_T, z_T)$ в пространстве, связанном с центром Земли; кроме тяготения Земли учитывается еще сила притяжения Луны. Задача состоит в отыскании начального импульса, реализующего этот перелет.

Уравнения движения запишем в следующей форме:

$$m_1 \ddot{\mathbf{r}}_1 = f \frac{m_1 m_0}{r_1^3} \mathbf{r}_1 - \frac{m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} (\mathbf{r}_1 - \mathbf{r}_2), \quad (2.4)$$

$$m_2 \ddot{\mathbf{r}}_2 = f \frac{m_2 m_0}{r_2^3} \mathbf{r}_2 - \frac{m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} (\mathbf{r}_2 - \mathbf{r}_1).$$

Здесь m_1 — масса Луны, m_2 — масса аппарата, m_0 — масса Земли, f — гравитационная постоянная, \mathbf{r}_1 и \mathbf{r}_2 — радиусы-векторы Луны и аппарата относительно центра неподвижной Земли. Второе слагаемое в правой части первого уравнения мало и его



Фиг. 6

следует отбросить; тогда первое из уравнений системы (2.4) можно проинтегрировать независимо от второго.

Систему (2.4) перепишем так:

$$m_1 \ddot{r}_1 = f(r_1) U, \quad m_2 \ddot{r}_2 = f(r_2) U, \quad (2.5)$$

где

$$U = \frac{1}{2} \frac{m_0 m_1}{r_1} - \frac{m_0 m_2}{r_2} - \frac{m_1 m_2}{|r_1 - r_2|}.$$

Функционал W для системы (2.5) имеет вид

$$W = \int_0^T m_1 \frac{\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2}{2} + m_2 \frac{\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2}{2} - f U dt. \quad (2.6)$$

Размерность шкалы управлений, если строго следовать изложенной схеме, равна в этом случае шести. Однако, как мы это заметили, первое из уравнений системы (2.5) может быть проинтегрировано независимо от второго. Следовательно, величины x_1 , y_1 , z_1 , x_2 , y_2 и z_2 следует считать заданными функциями времени. Поэтому в конкретных расчетах размерность шкалы управлений будет равна трем.

Примечание. Может оказаться, что поставленная задача не имеет решения. В этом случае предлагаемый процесс также не может быть реализован. Поясним это на примере.

Предположим, что функция Лагранжа имеет вид

$$L = \frac{\dot{q}^2}{2} - \frac{1}{q} - \frac{q^2}{2}.$$

Фазовая плоскость этого уравнения изображена на фиг. 6. Мы видим, что любые две прямые, проведенные в одной полуплоскости $q > 0$ или $q < 0$, можно соединить бесчисленным множеством фазовых траекторий (например, прямые q_1 и q_2 и прямые q_3 и q_4). Каждой из этих фазовых траекторий будет отвечать свое время

движения. Перебирая фазовые траектории, которые соединяют две указанные прямые, мы сможем найти ту, вдоль которой время перехода равно заданному. Как это следует из фиг. 6, не существует фазовой траектории, соединяющей две прямые, лежащие в разных полуплоскостях, например прямые q_1 и q_4 . Если мы предпримем попытку реализовать изложенный алгоритм, то уже на первом шаге получим расходящийся интеграл, поскольку, каков бы ни был выбор кривой, соединяющей прямые q_1 и q_4 , она будет пересекать плоскость $q = 0$.

§ 3. Задачи, для которых справедлив принцип максимума

Л.С. Понтрягина

1. Рассмотрим следующую задачу оптимального управления: определить вектор-функцию $u(t)$, дающую минимум функционала

$$F(x, T) = \int_0^T f(x, u) dt, \quad (3.1)$$

при условии, что вектор-функция x с компонентами x_1, \dots, x_n удовлетворяет дифференциальному уравнению

$$\dot{x} = f(x, u) \quad (3.2)$$

и граничным условиям

$$x(0) = x_0, \quad x(T) = x_T. \quad (3.3)$$

Кроме ограничений (3.2) и (3.3) потребуем, чтобы $u \in G$, где G может быть и замкнутым множеством. Вектор-функцию $f(x, u)$ и функцию $F(t)$ будем предполагать непрерывно дифференцируемыми.

Если вектор-функции x и u решают поставленную вариационную задачу, то, согласно теореме Л.С. Понтрягина, они должны быть решением следующей краевой задачи: определить вектор-функции x и p , которые удовлетворяют следующей системе дифференциальных уравнений:

$$\dot{x} = \tilde{p} H, \quad \dot{p} = -\tilde{x} H, \quad (3.4)$$

где

$$H = H^* p_{n-1}, \quad H^* = (fp). \quad (3.5)$$

Функция $F(t)$ определяется равенствами $F = \Phi$, $F(0) = 0$. Функция p_{n+1} удовлетворяет уравнению

$$p_{n-1} = \frac{H}{F} = 0 \quad (3.6)$$

при следующем условии на правом конце:

$$p_{n-1}(T) = 1. \quad (3.7)$$

Символы \tilde{x} и \tilde{p} означают, что градиенты вычисляются в предположении, что $u(x, p, t)$ постоянна, вектор-функция x должна удовлетворять условию (3.3), на компоненты вектор-функции $p = \{p_1, \dots, p_n\}$ никаких условий не накладывается, вектор-функция u_* определяется равенством

$$H(p, x, p_{n-1}, u_*) = \min_u H(p, x, p_{n-1}, u). \quad (3.8)$$

Как это следует из предыдущего, для решения вариационной задачи (3.1) непосредственно нельзя применить вариант динамического программирования, изложенный в § 1.

2. Будем предполагать сначала, что множество G открытое; тогда функция u должна удовлетворять условию

$${}_u H = 0. \quad (3.9)$$

Рассмотрим функционал

$$W(x, p, p_{n-1}, F) = \int_0^T \{(x, p) - F p_{n-1} - H\} dt = F(T), \quad (3.10)$$

где функция $F(t)$ определена равенствами

$$F = \dots, \quad F(0) = 0. \quad (3.11)$$

Обозначим через W_{x_i} вариацию W при условии, что варьируется только одна функция x_i . Так как $x_i(0) = x_i(T) = 0$, то

$$W_i \int_0^T x_i, p_i \frac{H}{x_i} x_i dt = \int_0^T p_i \frac{H}{x_i} x_i dt, \quad i = 1, 2, \dots, n.$$

На основании леммы Лагранжа получаем $p_i = H/x_i$. Принимая во внимание условие (3.9) $H/x_i = (H/x_i)_{u \text{ const}}$, мы можем написать $p_i = (H/x_i)_{u \text{ const}}$ и, следовательно, $p = x H$.

Аналогично,

$$W_{p_i} \int_0^T p_i x_i \frac{H}{p_i} dt,$$

или $x \sim_p H$.

Далее,

$$W_{p_{n-1}} \int_0^T p_{n-1} (F) dt,$$

откуда $F = \Phi$, и, наконец,

$$W_F \int_0^T F_{p_{n-1}} dt = F(T).$$

Но $F(0) = 0$, поэтому

$$W_F \int_0^T F_{p_{n-1}} dt = F(T) (p_{n-1}(T) = 1),$$

откуда $p_{n-1} = 0$, $p_{n-1}(T) = 1$.

Итак, условие, которое состоит в том, что функции x и p удовлетворяют краевой задаче (3.3)—(3.9) в предположении, что функция $u(x, p)$ определяется уравнением (3.9), является необходимым условием экстремума функционала (3.10). Эти условия необходимы, но не достаточны. Они определяют также и другие стационарные значения W . Очевидно, что эти состояния не могут быть определены изложенным методом.

Таким образом, исходная вариационная задача для функционала (3.1) заменена другой вариационной задачей для функционала (3.10). Для решения исходной задачи (3.1) нельзя было использовать шкалу управлений и изложенный алгоритм. Вариационная задача для функционала (3.11) уже может быть решена предлагаемым способом, если в качестве «управления» взять вектор U с компонентами $x_1, \dots, x_n, F, p_1, \dots, p_n, p_{n-1}$.

Примечания. 1. Вектор U имеет размерность $2n - 2$. Несмотря на это, при решении задачи нам достаточно использовать шкалу, размерность которой равна $2n$. В самом деле, используя тот факт, что $p_{n-1} = 1$, функционал (3.10) можно преобразовать к следующему виду:

$$W = \int_0^T \{(x, p) [H^*(x, p) - \dot{x}(x, p)]\} dt. \quad (3.12)$$

Это выражение показывает, что задача сводится к построению кривой в пространстве $2n + 1$ измерений (x, p, t) , соединяющей некоторую точку n -мерного многообразия $S_1(t = t_0, x = x_0)$ с точкой многообразия той же размерности $S_2(t = T, x = x_T)$, вдоль которой функционал (3.12) принимает наименьшее значение. Положение концов этой кривой внутри многообразий S_1 и S_2 не фиксируется, а должно быть определено.

2. В наших рассуждениях предположение о том, что функция u определяется уравнением (3.10), было существенным. Если оно не выполняется, то необходимые условия минимума функционала W уже не совпадают с условиями задачи (3.3)—(3.9). Однако иногда удается преодолеть эту трудность введением нового управления. Пусть, например, управление — функция u ограничена условием $|u| \leq 1$. Тогда положим $u = \sin \theta$. Условие максимума по θ будет иметь вид (3.9), и определение нового управления — функции θ может быть осуществлено изложенными методами.

3. Мы рассмотрели только одну из возможных задач — задачу с фиксированными концами. Однако легко убедиться, что целый ряд других задач для функционала (3.1) при условии (3.9) может быть сведен к вариационным задачам для функционала типа (3.10), которые допускают применение данного метода. Например, можно показать, что задача со свободным концом сводится к определению кривой в пространстве x, p, t , соединяющей многообразия $S_1(t = 0, x = x_0)$ и $S_2(t = T, p = 0)$, на которой функционал (3.12) достигает наименьшего значения.

4. В этой работе мы рассмотрели задачи для функционала специального вида (3.1). Однако в рамках излагаемой схемы могут быть рассмотрены и другие типы функционалов, например функционалы вида $I(x, u)$. Может быть рассмотрен и более общий случай, когда система уравнений зависит от функционала $x = f(x, u, F)$.

Единственным ограничением, необходимым для редукции исходной задачи к задаче, допускающей применение методов динамического программирования со шкалой управлений, является возможность так ввести управление, чтобы принцип максимума был справедлив и приводил нас к уравнению $H = 0$. (Разумеется, при этом мы находим только те решения, которые отвечают экстремальным значениям функционала J .)

Заключение

Изложенная схема численного решения вариационных задач имеет ряд достоинств и, по-видимому, может быть использована для разработки стандартных способов численного решения довольно широкого круга задач.

В настоящее время для практического решения вариационных задач используется редукция последних к краевым задачам для обыкновенных дифференциальных уравнений. В свою очередь, для решения этих задач употребляются различные методы организованного поиска: метод градиентного спуска, метод Ньютона и т. д. Эти методы требуют решения задачи Коши. Решение задачи Коши в общем случае неустойчиво, и при больших интервалах времени методы организованного поиска чрезвычайно трудоемки, а иногда и просто нереализуемы на современных ЭВМ, даже в случае хорошего первого приближения.

Предлагаемые методы не требуют решения задачи Коши. На каждом шаге итерационного процесса все краевые условия удовлетворяются. Этот итерационный процесс всегда сходится, независимо от характера первого приближения. (Хорошее первое приближение только сокращает число итераций.)

В данной работе рассматривался некоторый определенный класс задач, которые мы называли задачами, допускающими использование шкалы управлений. Легко указать практически очень важные задачи, которые не могут быть решены описанным способом. Однако в том случае, когда минимизируемый функционал аддитивен для таких задач, также могут быть развиты методы динамического программирования. Однако они потребуют введения нового понятия шкалы состояний. Этому вопросу будет посвящена следующая статья. Автор искренне благодарен Н.Я. Багаевой, Д.Е. Охоцимскому, Б.Н. Пшеничному и Ф.Л. Черноуьско, замечаниями которых он воспользовался при подготовке статьи к печати.

Литература

1. *Охоцимский Д.Е.* К теории движения ракет./Прикл. матем. и механ., 1946, 10, вып. 2, 251–272.
2. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф.* Математическая теория оптимальных процессов. – М., Физматгиз, 1961.
3. *Беллман Р.* Динамическое программирование. – М., Изд-во ин. лит. 1960.
4. *Беллман Р., Гликберг И., Гросс О.* Некоторые вопросы математической теории процессов управления. – М., Изд-во ин. лит., 1962.
5. *Кротов В.Ф.* Методы решения вариационных задач на основе достаточных условий абсолютного минимума. I./Автоматика и телемехан., 1962, 23, № 12, 1571–1581.
6. *Дубовицкий А.Я., Милютин А.А.* Задачи на экстремум при наличии ограничений./Докл. АН СССР, 1963, 149, № 4.
7. *Шатровский Л.И.* Об одном численном методе решения задач оптимального управления./Ж. вычисл. матем и матем. физ., 1962, 2, № 3, 488–491.
8. *Крылов И.А., Черноусько Ф.Л.* О методе последовательных приближений для решения задач оптимального управления./Ж. вычисл. матем. и матем. физ., 1962, 2, № 6, 1132–1139.
9. *Михалевич В.С., Шор Н.З.* О численных методах решения многовариантных плановых и технико-экономических задач./Научно-метод. материалы экономико-матем. семинара ВЦ АН УССР. Вып. 1. Киев, ВЦ АН УССР, 1962.

1963 г.

Методы динамического программирования в теории оптимальных управлений

II. Общий случай аддитивных функционалов

В первой части этой работы [1] были рассмотрены некоторые задачи вариационного исчисления, численное решение которых можно было построить специальной организацией перебора с использованием шкалы управлений. Однако метод, развитый в первой части, позволил рассмотреть только очень узкий класс задач. Например, наши рассуждения переставали быть справедливыми, если речь шла о минимизации функционала

$$\int_0^T F(x, u) dt,$$

где x — вектор, удовлетворяющий дифференциальному уравнению $\dot{x} = f(x, u)$, u — управление.

В данной части работы рассматривается общий случай аддитивных функционалов

$$A(t_1, t_2) = A(t_1) + A(t_2),$$

т. е. функционалов, для которых справедлив принцип Беллмана.

§ 4. Одна вариационная задача теории матриц

1. Все задачи, которые были рассмотрены в первой части работы, и задачи, которые будут рассмотрены в данной части, сводятся, как мы это увидим, к следующей вариационной задаче теории матриц.

Пусть мы имеем трехмерную матрицу

$$M = \{I_{kj}^i\}, \quad i = 1, 2, \dots, N, \quad k = 1, 2, \dots, N_1, \quad j = 1, 2, \dots, N_1,$$

где I_{kj}^i — заданные числа. Требуется разыскать минимум суммы

$$I = \min_{k, j} \sum_{i=1}^N I_{kj}^i. \quad (4.1)$$

Минимум берется по всем k, j при условии что из каждого i -го слоя мы берем один элемент, причем так, чтобы второй нижний индекс i -го слагаемого равнялся первому нижнему индексу $(i + 1)$ -го слагаемого.

2. Для решения этой задачи используется следующий алгоритм. Введем обозначение

$$I_s^l = \min_{k, j} \sum_{i=1}^l I_{kj}^i. \quad (4.2)$$

При вычислении I_s^l вводится дополнительное требование: чтобы элемент I_{ks}^l номера l имел в качестве второго нижнего индекса число s . Так как s меняется от 1 до N_1 , то верхнему индексу l будет отвечать N_1 сумм вида (4.2).

Покажем, как строятся суммы I_s^l . Предположим, что I_r^{l-1} построены. Рассмотрим множество элементов I_{ks}^l , у которых фиксирован второй нижний индекс. Это множество состоит из N_1 чисел. Тогда, очевидно,

$$I_s^l = \min_r \{I_r^{l-1} + I_{rs}^l\}. \quad (4.3)$$

Этим путем для $l = N$ мы находим N_1 сумм I_s^N , после чего

$$I = \min_s I_s^N. \quad (4.4)$$

На каждом шаге (переход от верхнего индекса $l - 1$ к индексу l) мы совершаем N_1 сложений и сравнений. Таким образом, всего мы проделываем $NN_1^2 - N_1$ подобных операций. Стандартная программа решения этой матричной задачи дана в [2].

Примечания. 1. Мы предполагали, что матрица M представляет собой параллелепипед. Может оказаться, что каждому индексу i отвечает своя неквадратная матрица чисел

$$\{I_{kj}^i\}, \quad k = 1, 2, \dots, N_{1i}, \quad j = 1, 2, \dots, N_{2i}.$$

Тогда мы назначаем в качестве числа N_1 наибольшее из чисел N_{1i} и N_{2i} и в пустующие ячейки помещаем число, воспринимаемое машиной в качестве . Если речь идет об отыскании максимума суммы (4.1), то мы в эти ячейки помещаем число .

2. Изложенный алгоритм аналогичен алгоритму перебора, приведенному в [3].

3. Может оказаться, что объем матрицы будет настолько большим, что изложенная схема будет нереализуема на машине. В этом случае мы вынуждены использовать различные итерационные схемы, которые позволяют использовать на каждом шаге незначительную часть матрицы M . Приведем одну из возможных итерационных схем.

Пусть нам задано некоторое «первое приближение»

$$I_0 \quad I_{k_1^0, j_1^0}^1 \quad I_{k_2^0, j_2^0}^2 \quad \dots, \quad k_{i-1}^0 \quad j_i^0.$$

Построим матрицу M_1 , состоящую из $9N$ чисел. Каждому индексу i будет отвечать следующая девятка чисел:

$$I_{k_i^0, j_i^0}^i, I_{k_i^0-1, j_i^0}^i, I_{k_i^0+1, j_i^0}^i; \quad I_{k_i^0, j_i^0-1}^i, I_{k_i^0, j_i^0+1}^i, I_{k_i^0+1, j_i^0-1}^i; \\ I_{k_i^0, j_i^0-1}^i, I_{k_i^0-1, j_i^0+1}^i, I_{k_i^0+1, j_i^0+1}^i.$$

Для данной матрицы решаем задачу определения минимума суммы (4.1), этот минимум обозначим через I_1 .

Используя данное приближение, строим матрицу M_2 так же, как M_1 . Далее строим сумму I_2 и т. д.

По построению $I_0 \quad I_1 \quad I_2 \quad \dots$ Так как матрица M конечна, то на конечном шаге мы получим $I_s \quad I_{s+1}$.

Примечания. 1. Предлагаемая схема последовательных приближений нам может дать только локальный минимум.

2. Изложенная схема является только примером. В зависимости от объема памяти машины и структуры матрицы M могут быть использованы самые разнообразные итерационные способы, основная цель которых — свести поставленную задачу к задаче с меньшим объемом матрицы M .

4. Один из простейших примеров, который сводится к рассмотренной задаче, — это проблема оптимизации функционала

$$I(u) = \int_0^T F(u, u) dt, \tag{4.5}$$

которая изучалась в § 1 данной работы.

Мы аппроксимировали решение ломаной. Для этого на плоскости (u, t) была построена шкала управлений — множество точек $\{u_j^i\}$, лежащих на прямых $t = ih$. Каждая из точек u_j^i соединялась прямой с каждой из точек u_k^{i-1} :

$$u_{j,k}^{i-1} \quad u_j^i \quad \frac{u_k^{i-1} \quad u_j^i}{h} (t = ih).$$

И вдоль этого отрезка прямой вычислялись значения функционала

$$I_{j,k}^{i-1} \quad \int_{ih}^{(i-1)h} F(u_{jk}^{i-1}, \frac{u_k^{i-1} \quad u_j^i}{h}) dt.$$

Числа $I_{j,k}^{i-1}$ образуют матрицу M .

5. В известных условиях эта схема может быть использована при решении краевых задач для обыкновенных дифференциальных уравнений. Предположим, что мы имеем систему

$$x_i \quad f_i(x_1, \dots, x_n) \quad (4.6)$$

и требуется определить ее решение, удовлетворяющее условиям

$$x_i(0) = a_i, \quad i = 1, 2, \dots, k;$$

$$x_i(T) = b_i, \quad i = k + 1, \dots, n. \quad (4.7)$$

Рассмотрим функционал

$$I(x_1, \dots, x_n) = \int_0^T (\sum_{i=1}^n f_i(x_1, \dots, x_n))^2 dt. \quad (4.8)$$

Отыскание решения краевой задачи (4.6)—(4.7) эквивалентно отысканию абсолютного минимума функционала (4.8) на множестве векторов (x_i) , удовлетворяющих условиям (4.7). Таким образом, разыскивая описанным методом всевозможные минимумы функционала (4.8) и отбирая те из них, которые равны нулю, мы найдем все возможные решения исходной краевой задачи. Для получения достаточно точного решения необходимо использовать итерационную схему, при этом процесс счета будет всегда устойчивым и всегда будет иметь место сходимость по функционалу, если решение краевой задачи существует.

Примечание. К матричной задаче, рассмотренной в данном параграфе, легко сводится задача определения геодезических на данной поверхности $F(x, y, z) = 0$. Числа I_{jk}^i — это просто длины отрезков, соединяющих определенным образом упорядоченное множество узлов сетки, натянутой на поверхность. С другой стороны, отыскание геодезической, проходящей через две заданные точки, сводится к решению некоторой краевой задачи для дифференциальных уравнений, которым удовлетворяет геодезическая. Таким образом, тот факт, что с этой системой уравнений ассоциирована определенным образом поверхность, по отношению к которой траектории этой системы обладают определенными минимальными свойствами, позволяет применить непосредственно методику этого параграфа без редукции к задаче типа (4.8).

§ 5. Задачи теории оптимального управления

1. Рассмотрим один частный случай задачи теории оптимального управления — задачу о минимуме функционала

$$I(u) = \int_0^T F(x, u) dt, \quad (5.1)$$

где u — управляющий вектор размерности m , x — вектор размерности n , изменение которого описывается дифференциальным уравнением

$$\dot{x} = f(x, u). \quad (5.2)$$

На управление-вектор u и вектор x будут наложены еще следующие условия:

$$u \in G_u, \quad (5.3)$$

$$x \in G_x, \quad (5.4)$$

$$x(0) \in G_0, \quad (5.5)$$

$$x(T) \in G_T; \quad (5.6)$$

множества G_u и G_x могут быть открытыми или замкнутыми. Условие (5.4) накладывает ограничения на фазовые координаты. G_0 и G_T — многообразия, имеющие размерность меньшую, чем n . Могут рассматриваться также задачи, в которых ограничение (5.6) отсутствует. Такие задачи мы будем называть задачами со свободным концом.

Подобно задаче о минимуме функционала, рассмотренной в § 1, эта задача также может быть сведена к вариационной задаче предыдущего параграфа. Однако для этой цели уже не может быть использована шкала управлений.

2. Введем понятие шкалы состояний. Рассмотрим для этого пространство (x, t) и зададим шаг h . В гиперплоскостях $t = ih$ разместим конечные множества $\{P_j^i\}$ — множества узлов сеточной схемы. Совокупность этих множеств мы будем называть шкалой состояний. Принимается, что точка шкалы состояний $P_j^i \in G_x$. Тем самым траектория, проходящая через точки шкалы состояний, удовлетворяет фазовым ограничениям автоматически.

Для того чтобы свести эту задачу к задаче § 4, нам надо соединить все точки P_j^i с точками P_k^{i-1} и вычислить вдоль этих кривых

$$I_{jk}^{i-1} = \int_{ih}^{(i-1)h} F(x, u) dt. \quad (5.7)$$

Но эти кривые должны удовлетворять дифференциальным уравнениям (5.2).

Пусть нам даны две точки P и Q . Могут представиться два случая:

1) через эти две точки нельзя провести интегральную кривую системы (5.2), удовлетворяющую условию (5.4), каково бы ни было управление $u(t)$, удовлетворяющее условию (5.3). В этом случае точка Q называется недостижимой из точки P , и при построении матрицы M в ячейке, соответствующей значению функционала при переходе из точки P в точку Q , мы помещаем ∞ ;

2) существуют управления, которые переводят систему из состояния P в состояние Q . Это значит, что можно найти такое управление $u(t) \in G_u$, что точки P и Q окажутся на одной интегральной кривой системы (5.2). В этом случае точка Q будет достижимой точкой.

В последнем случае наша задача состоит в отыскании такого управления $u(t)$, которое обеспечивало бы переход системы из точки P в точку Q согласно дифференциальным уравнениям (5.2). Типичным является тот случай, когда решение этой задачи неединственно. Мы должны выбрать то решение, которое доставляет функционалу (5.7) наименьшее значение.

Итак, мы приходим к вспомогательной вариационной задаче: определить управление $u(t)$, которое переводит систему (5.2) из

состояния P_j^i в состояние P_k^{i-1} и сообщает функционалу (5.7) наименьшее значение.

Эта задача более легкая, чем исходная, поскольку в ней отсутствуют фазовые ограничения.

Операцию, которая двум заданным точкам ставит в соответствие управление $u(t)$, переводящее систему из одной точки в другую, и минимальное значение функционала (5.7), которое достигается на этом отрезке траектории, будем называть элементарной операцией. В отличие от операции, которая использовалась в первой части работы, эту операцию будем называть B -операцией. Таким образом, B -операция — это некоторый способ решения вспомогательной вариационной задачи.

Итак, общая схема решения задач оптимального управления такова: строится шкала состояний, учитывающая все фазовые ограничения. Далее составляются стандартная программа для элементарной операции и матрица M . Для этой матрицы решается вариационная задача, поставленная в § 4.

Во всей этой схеме наиболее трудным вопросом является реализация элементарной операции. Его решение зависит от конкретного содержания задачи и трудно для стандартизации. В следующем параграфе будет описан ряд способов построения B -операции.

Примечания. 1. Для практического применения данного метода, т. е. при его программировании, нет необходимости сразу вычислять всю матрицу M . Следует иметь стандартный блок перебора при переходе от i -го слоя к $(i+1)$ -му и вычислять лишь те элементы матрицы M , которые необходимы для реализации этого перебора.

2. Предположим, что для некоторой задачи нам удалось реализовать указанную схему расчета и мы нашли некоторую траекторию Γ_0 управления u_0 и значение функционала вдоль этой траектории I_0 . Мы всегда можем уточнить это решение, уменьшив шаг по t и построив в окрестности Γ_0 новую шкалу состояний с меньшими шагами по переменным x_i (компоненты вектора x). Новую шкалу состояний надо строить так, чтобы точки старой шкалы состояний были среди точек новой. В результате мы получаем новое решение — кривую Γ_1 , управление u_1 и функционал I_1 . Так как кривая Γ_0 находилась среди траекторий, допущенных к сравнению, то $I_1 < I_0$.

Описанный процесс можно повторять неограниченное число раз. В результате мы получим последовательность управлений $\{u_n\}$

и последовательность функционалов $\{I_n\}$, причем $I_0 \leq I_1 \leq I_2 \leq \dots$. В том случае, когда функционал I ограничен снизу, последовательность $\{I_n\}$ сходится. Следовательно, в этом случае последовательность $\{u_n\}$ сходится в смысле сходимости по функционалу.

Итак, развитый метод позволяет строить слабо сходящуюся последовательность управлений.

3. К рассматриваемой задаче может быть сведена задача, в которой ограничения имеют характер интегральных неравенств. Пусть, например, требуется найти фазовую траекторию, вдоль которой для любого $t \in [0, T]$

$$\int_0^t (x, u) dt \leq C, \quad (5.8)$$

где C — заданная постоянная. Введем новую переменную x_{n+1} :

$$x_{n+1} = (x, u).$$

Тогда условие (5.8) будет означать, что $x_{n+1} \leq C$ для любых t из интервала $(0, T)$. Таким образом мы пришли к рассмотренной задаче, но для нового вектора $x^*(t)$, размерность которого на единицу выше.

Задача подобного рода возникает, например, в динамике космического аппарата, движущегося внутри поясов радиации. Наличие радиации, плотность которой определяется некоторой функцией (x, y, z) (x, y, z — декартовы координаты), накладывает определенные требования на траекторию: суммарная радиация не должна превосходить некоторого предела

$$\int_0^T (x, y, z) dt \leq C.$$

Эта задача была рассмотрена в [4].

4. Точно так же могут быть рассмотрены различные варианты изопараметрических задач. Пусть, например, ограничение имеет вид

$$\int_0^T (x, u) dt \leq C. \quad (5.9)$$

Введем новую переменную x_{n+1} , как в предыдущем примечании. Тогда условие (5.9) нам даст граничное условие при $t = T$: $x_{n+1}(T) \leq C$.

§ 6. О построении элементарной операции

Построение элементарной операции является центральным вопросом теории. Здесь очень трудно дать какие-либо общие рекомендации. Поэтому в данном параграфе мы обсудим ряд соображений, позволяющих во многих случаях эффективно реализовать элементарную операцию.

1. Задача построения элементарной операции состоит в разработке алгоритма выбора управления, оптимальным образом переводящего систему из одного состояния в другое, близкое состояние. Эта близость исходного и конечного состояний позволяет развить во многих задачах эффективные способы реализации B -операции.

Рассмотрим сначала один простой пример — задачу о переходе космического аппарата за время t из точки $x_0, y_0, \dot{x}_0, \dot{y}_0$ в близкую точку x, y, \dot{x}, \dot{y} в предположении, что он находится под действием ньютоновского гравитационного поля и снабжен идеально регулируемым двигателем малой тяги. В этом случае функционал (5.1) и система уравнений (5.2) имеют вид

$$I(q) = \int_0^t q^2 dt; \quad (6.1)$$

$$\dot{x} = f_1 q \cos u, \quad \dot{y} = f_2 q \sin u. \quad (6.2)$$

Здесь x и y — декартовы координаты; f_1 и f_2 — компоненты напряженности гравитационного поля:

$$f_1 = \frac{x}{(x^2 + y^2)^{3/2}}, \quad f_2 = \frac{y}{(x^2 + y^2)^{3/2}};$$

q и u — искомые управления.

На величины q и u никаких ограничений не накладывается.

Заменой x, y система (6.2) сводится к системе уравнений первого порядка. Так как фазовые ограничения считаются удовлетворенными выбором точек $P_0(x, y)$ и $P(x, y)$, то вспомогательная задача — это классическая задача и к ней может быть применен принцип максимума.

Далее, точки P_0 и P будем брать настолько близкими, чтобы функции f_1 и f_2 можно было считать постоянными. В этих предположениях составим гамильтониан

$$H = p_1 \dot{x} + p_2 \dot{y} - f_1 p_3 - f_2 p_4 - H^*, \quad H^* = q(p_3 \cos u + p_4 \sin u) q^2. \quad (6.3)$$

Условия для определения функций q и u , согласно Л.С. Понтрягину, имеют вид

$$\begin{aligned} \frac{H^*}{u} & q(p_3 \sin u - p_4 \cos u) = 0, \\ \frac{H^*}{q} & p_3 \cos u - p_4 \sin u - 2q = 0. \end{aligned} \quad (6.4)$$

Откуда

$$\cos u = \frac{p_3}{\sqrt{p_3^2 + p_4^2}}, \quad \sin u = \frac{p_4}{\sqrt{p_3^2 + p_4^2}}, \quad q = \frac{1}{2} \sqrt{p_3^2 + p_4^2}. \quad (6.5)$$

Составим уравнения для импульсов

$$\dot{p}_1 = 0, \quad \dot{p}_2 = 0, \quad \dot{p}_3 = p_1, \quad \dot{p}_4 = p_2. \quad (6.6)$$

Система уравнений (6.6) интегрируется в явном виде:

$$p_1 = \text{const}, \quad p_2 = \text{const}, \quad p_3 = p_{30} + p_1 t, \quad p_4 = p_{40} + p_2 t. \quad (6.7)$$

Вернемся теперь к системе (6.2), используя (6.5) и (6.7):

$$\dot{x} = f_1 \frac{1}{2}(p_{30} + p_1 t), \quad \dot{y} = f_2 \frac{1}{2}(p_{40} + p_2 t).$$

Эта система интегрируется в явном виде. Для определения p_1 , p_2 , p_{30} и p_{40} мы получаем четыре уравнения:

$$x(t) - x(0) = f_1 \frac{1}{2} p_{30} t + \frac{p_1}{2} t^2,$$

$$y(t) - y(0) = f_2 \frac{1}{2} p_{40} t + \frac{p_2}{2} t^2,$$

$$\dot{x}(t) - \dot{x}(0) = \frac{f_1}{2} t + \frac{1}{4} p_{30} t^2 + \frac{p_1}{3} t^3,$$

$$\dot{y}(t) - \dot{y}(0) = \frac{f_2}{2} t + \frac{1}{4} p_{40} t^2 + \frac{p_2}{3} t^3.$$

Таким образом, в рассматриваемой задаче элементарная операция свелась к решению системы линейных алгебраических уравнений.

Примечание. Вместо гипотезы постоянного поля можно использовать гипотезу «линеаризованного» поля:

$$f_1 = \frac{1}{2}x^2, \quad f_2 = \frac{1}{2}y^2,$$

где постоянное число $\frac{1}{2} = 1/(x^2 + y^2)^{1/2}$, где x и y приняты равными средним значениям. В этом случае элементарная операция также сводится к решению системы линейных алгебраических уравнений. Линейная аппроксимация гравитационного поля оказывается очень удобной во многих задачах динамики космических аппаратов. Ее применение позволяет использовать большой шаг по времени, т. е. брать матрицу M сравнительно малого объема.

2. В рассматриваемом примере элементарную операцию удастся записать в явном виде. Этот случай является исключительным. В общем случае мы сталкиваемся с целым рядом трудностей. При реализации элементарной операции следует использовать близость точек P_0 и P . Это осуществляется следующим образом:

а) в силу близости точек P_0 и P переход из состояния P_0 в состояние P реализуется без фазовых ограничений. Поэтому мы можем использовать принцип максимума. Близость этих состояний позволяет также сделать ряд упрощений в системе уравнений. Они должны быть сделаны с таким расчетом, чтобы система уравнений для импульсов могла быть проинтегрирована точно, тогда мы будем иметь представление о структуре управления;

б) если система уравнений для импульсов может быть проинтегрирована, то импульсы будут определены с точностью до постоянных интегрирования. Следовательно, после исключения управления мы приходим к системе уравнений относительно x_i , правые части которых содержат неизвестные постоянные.

Используя краевые условия при $t = 0$, мы получим систему уравнений для определения этих постоянных. Если интервал T мал, то должна быть построена конечноразностная схема, которая позволила бы заменить полученную систему дифференциальных уравнений системой трансцендентных уравнений относительно неизвестных

констант. В качестве примера, который поясняет сказанное, рассмотрим систему уравнений

$$\dot{x}_i = f_i(x_1, \dots, x_n) - p_i(u). \quad (6.8)$$

Управление — функция u — должно быть выбрано так, чтобы при $t =$

$$x_{n-1} = \int_0^t f_{n-1}(u) dt$$

обращалась в минимум. Простейшая схема может быть получена следующим образом.

В системе (6.8) положим $f_i = \bar{f}_i$, где \bar{f}_i — постоянные, равные, например, средним значениям этих функций на интервале изменения. Тогда уравнения для импульсов дадут нам $p_i = \text{const}$, а из принципа максимума будет следовать, что управление выбирается так, что $\dot{x}_i p_i = f_{n-1}(u)$ принимает максимальное значение.

Следовательно, управляющая функция — это некоторая постоянная, зависящая определенным образом от постоянных p_i , т. е. $u = U(p_1, \dots, p_n)$ и, следовательно, $\dot{x}_i(u) = \dot{x}_i(p_1, \dots, p_n)$.

Интегрируя систему (6.8) и используя условия при $t =$, мы придем к системе уравнений

$$\dot{x}_i(p_1, \dots, p_n) = \frac{x_i(t) - x_i(0)}{t} - \bar{f}_i. \quad (6.9)$$

Итак, в данном примере элементарная операция сведена к решению некоторой трансцендентной системы. Формулы (6.9) — это простейшие конечноразностные формулы, написанные с учетом принципа максимума.

3. В общем случае не удастся построить конечноразностную схему столь же простую, как та, которая дается формулой (6.9). В самом деле, если даже одна из функций $\dot{x}_i > 0$, то, положив $f_i = \bar{f}_i$, мы получим, что $x_i = x_{i0} + \bar{f}_i t$; тогда при $t =$ значение x_i будет вполне определено, и выбором величин p_i мы не сможем удовлетворить краевому условию для i -й координаты. Следовательно, в данном случае мы не можем считать управление постоянным. Как бы ни был мал интервал времени, управление на этом интервале должно изменяться во времени. На простейшем примере укажем один из способов преодоления этой трудности. Рассмотрим систему уравнений

$$\dot{x}_1 = f_1(x_1, x_2) + u, \quad \dot{x}_2 = f_2(x_1, x_2) \quad (6.10)$$

и вариационную задачу для функционала

$$I = \int_0^1 F(u) du$$

при граничных условиях

$$x_i(0) = x_{i0}, \quad x_i(1) = x_{i1}. \quad (6.11)$$

Точку $\{x_i\}$ мы считаем близкой к точке $\{x_{i0}\}$ и систему (6.10) линеаризуем, положив $x_i = x_{i0} + \delta x_i$. Тогда уравнения (6.10) и условия (6.11) примут вид

$$\dot{\delta x}_1 = f_{10} + a_{11} \delta x_1 + a_{12} \delta x_2 + u, \quad \dot{\delta x}_2 = f_{20} + a_{21} \delta x_1 + a_{22} \delta x_2, \quad (6.12)$$

$$\delta x_i(0) = 0, \quad \delta x_i(1) = \delta x_{i1}, \quad (6.13)$$

где a_{ij} — некоторые постоянные.

Выпишем выражения гамильтониана

$$H = (f_{10} + a_{11} \delta x_1 + a_{12} \delta x_2) p_1 + (f_{20} + a_{21} \delta x_1 + a_{22} \delta x_2) p_2 + u p_1 - F(u).$$

Условие максимума H нам позволит найти управление

$$u = -p_1. \quad (6.14)$$

Уравнениями для импульсов будут

$$\dot{p}_1 = -a_{11} p_1 - a_{21} p_2, \quad \dot{p}_2 = -a_{12} p_1 - a_{22} p_2. \quad (6.15)$$

Благодаря тому что система (6.10) была заменена линеаризованной, система уравнений для импульсов (6.15) может быть проинтегрирована в явном виде:

$$p_1 = c_1 p_{11} - c_2 p_{12}, \quad p_2 = c_1 p_{21} - c_2 p_{22}, \quad (6.16)$$

где c_i — постоянные, подлежащие определению, p_{ij} — фундаментальные решения системы (6.15).

Таким образом задача свелась к интегрированию системы (6.12), где $u^* = u^*(c_1, c_2, t)$, и определению постоянных c_1 и c_2 из (6.13).

Простейшая конечноразностная схема, которая может быть использована для решения этой задачи, основана на аппроксимации функции u^* постоянными $u^*(c_1, c_2, 0)$ и $u^*(c_1, c_2, T/2)$ на интервалах $(0, T/2)$ и $(T/2, T)$. Соответственно с этим трансцендентная система уравнений для определения c_1 и c_2 будет иметь вид (возможность использовать подобные схемы была отмечена в работе [5]):

$$\begin{aligned} 1 \quad f_{10} &= \frac{f_{10} a_{11}^2}{4} - \frac{f_{20} a_{12}^2}{4} \\ &= \frac{u^*(c_1, c_2, 0) a_{11}^2}{4} - \frac{u^*(c_1, c_2, 0) u^*(c_1, c_2, T/2)}{2}, \quad (6.17) \\ 2 \quad f_{20} &= \frac{f_{10} a_{21}}{4} - \frac{f_{20} a_{22}}{4} = \frac{a_{21} u^*(c_1, c_2, 0)^2}{4}. \end{aligned}$$

4. В рассмотренных примерах использовался принцип максимума Понтрягина. Он дал возможность изучить структуру управления и разыскивать управление среди функций определенного вида. Это обстоятельство, как мы видели, существенно упрощало задачу.

В некоторых случаях можно использовать и другие способы, позволяющие свести построение элементарной операции к решению трансцендентной системы уравнений.

Рассмотрим важный класс дискретных управлений. (К числу дискретных управлений относятся, например, релейные управления.) Простейшей задачей этого класса будет следующая. Определить минимум

$$F(u) = \int_0^T u^2 dt$$

при переходе системы из точки $x = 0$ в точку $x = x_1$ согласно уравнению $\dot{x} = u$. Управление считаем дискретным. Его возможные значения обозначим через u_i ($i = 1, 2, \dots, n$).

Для построения наиболее простой разностной схемы будем считать, что управление остается постоянным в течение отрезков времени t_1 и $t_2 = t - t_1$. Тогда

$$x = x_0 + u_i t_1 + u_j (t - t_1). \quad (6.18)$$

Из уравнения (6.18) вычисляем x . Все те варианты u_i и u_j , для которых $x_1 > x_2$ или $x_1 < 0$, отбрасываем. Далее, выражение функционала мы можем записать в виде

$$I = t_1 F(u_i) + (t - t_1) F(u_j).$$

Задача сводится к перебору всех возможных значений u_i и u_j , что позволит определить $\min I$. Число переборов равно n^2 . В случае релейного управления имеем $n^2 = 4$, при релейном с выключением $n^2 = 9$.

Рассмотрим еще один случай системы второго порядка:

$$\dot{x} = u; \quad x(0) = x(0) = 0, \quad x(t_1) = x_1, \quad x(t_2) = x_2.$$

В этом случае минимальное количество интервалов t_i , на которых управление может быть принято постоянным, равно трем (t_1, t_2 и $t_3 = t - (t_1 + t_2)$).

Система уравнений для определения t_1 и t_2 будет такой:

$$\begin{aligned} x &= x_0 + u_i t_1 + u_j t_2 + u_k (t - t_1 - t_2), \\ x &= x_0 + x_0 + u_i t_1 + u_j t_2 + u_k (t - t_1 - t_2) \frac{t_1}{2} \\ &+ u_j t_2 (t - t_1 - t_2) \frac{t_1}{2} + u_k \frac{t_2^2}{2} \end{aligned} \quad (6.19)$$

и, соответственно,

$$I(u_i, u_j, u_k) = \min_{i,j,k} \{ t_1 F(u_i) + t_2 F(u_j) + (t - t_1 - t_2) F(u_k) \}.$$

В этом построении элементарной операции требуется n^3 переборов. В случае релейного управления — 8 переборов, релейного с выключением — 27 переборов. При решении алгебраической системы (6.19) мы должны отбросить все те комбинации u_i, u_j , и u_k , для

которых величины x_1 и x_2 будут либо отрицательными, либо превосходить x_i^* .

В этом параграфе мы указали только наиболее простые и грубые разностные схемы, которые могут быть использованы. Исследование точности данного метода потребует рассмотрения более сложных разностных схем.

Заметим одновременно, что можно привести примеры задач, в которых нет необходимости использовать конечноразностные схемы и для построения элементарной операции можно взять, например, метод Ньютона.

§ 7. Некоторые дополнения

1. До сих пор мы рассматривали задачи с фиксированным временем. Задачи, в которых время не фиксировано, могут быть рассмотрены подобным методом. Если среди координат системы есть хотя бы одна, о которой известно а priori, что она изменяется монотонно, то задача сводится к рассмотренной заменой независимого переменного (в качестве нового переменного следует принять координату, которая изменяется монотонно). Если среди координат системы нет такой, которая изменяется монотонно, то должна быть изменена структура шкалы состояний. Пусть идет речь о задачах на быстроедействие. Введем скаляр G , характеризующий расстояние до цели управления (см. [5]). Предположим, например, что цель управления — точка с координатами x_i^* . Тогда можно принять

$$\sqrt{(x_i - x_i^*)^2}.$$

Шкалу состояний мы будем размещать на поверхности $G = G_i$. Тогда элементарная операция (которую в этом случае мы будем называть B -операцией) будет состоять в решении следующей вспомогательной вариационной задачи.

Для заданных точек шкалы состояний P^i и P^{i+1} , лежащих на смежных поверхностях $G = G_i$ и $G = G_{i+1}$, определить управление $u \in G_u$, которое за минимум времени переводит систему из состояния P^i в состояние P^{i+1} согласно дифференциальным уравнениям

$$\dot{x}_i = f(x_1, \dots, x_n, u).$$

2. Методы численного решения, которые рассматривались в данной работе, относились исключительно к задачам минимизации аддитивных функционалов. Тем не менее в некоторых случаях задачу минимизации неаддитивных функционалов удается свести к решению серии задач с аддитивным функционалом. Поясним это на примере.

Предположим, что задача состоит в отыскании управления и траектории системы

$$\dot{x} = f(x_1, \dots, x_n, u), \quad (7.1)$$

которая соединяет точки $P_0(x_i = 0)$ и $P(x_i = x_i^*)$ таким образом, что изображающая точка переходит из состояния P_0 в состояние P за время T , а минимум расстояния до некоторой фиксированной точки оказывается максимальным:

$$\max (\min \sqrt{(x_i(t) - x_i^*)^2}).$$

Функционал $\min \sqrt{(x_i(t) - x_i^*)^2}$ не будет аддитивным.

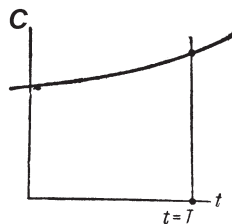
Рассмотрим вспомогательную задачу на быстродействие: найти управление, которое переводит систему из состояния P_0 в состояние P за минимум времени T_1 при условии, что $C_1 \leq C$ где C_1 — некоторая постоянная.

Эта задача с фазовыми ограничениями и аддитивным функционалом. Она может быть решена методами, развиваемыми в данной работе.

Если время T_1 окажется большим, чем T , то мы рассмотрим еще одну подобную задачу на быстродействие при ограничении $C_2 \leq C_1 \leq C$, $C > 0$.

Решив ряд подобных задач, мы можем построить кривую $C(t)$. Точка пересечения кривой $C(t)$ и прямой $t = T$ соответствует искомому режиму (см. фигуру).

В задачах подобного рода можно развить процедуру, аналогичную методу Ньютона определения нулей функции.



Заключение

Изложенный метод неоднократно использовался для решения конкретных задач. Для ряда модификаций были составлены программы

типа стандартных, и все они показали достаточную эффективность. Однако при практическом использовании метода динамического программирования нужно иметь в виду следующее обстоятельство. В настоящее время еще совершенно не разработан вопрос о зависимости предельных значений искомых функций от соотношения шагов по фазовым и временным переменным при неограниченном продолжении итерационного процесса. Нетрудно привести примеры, когда этот предел существенно зависит от данного отношения. При выполнении конкретных расчетов использовался следующий прием. Сначала фиксировался шаг по времени и на каждом из этапов итерационного процесса проводилось дробление шага только по фазовым переменным. В этом случае после получения «предельного значения» для данного разбиения оси времени производилось уменьшение шага по времени и вся процедура повторялась снова. В простых случаях удавалось показать, что этот прием снимает все указанные трудности. Однако в общем случае вопрос о соотношении шагов по координатам и времени, приводящих к нужному пределу, остается открытым.

В заключение я хочу высказать искреннюю благодарность Р.П. Федоренко, который прочел рукопись статьи и сделал ряд замечаний, использованных при редактировании.

Литература

1. *Моисеев Н.Н.* Методы динамического программирования в теории оптимальных управлений. I./Ж. вычисл. матем. и матем. физ., 1964, 4, № 3, 485–494.

2. *Сигал И., Чебаков В.А.* Метод матричного перебора и его применение к одной задаче теории графов./Ж. вычисл. матем. и матем. физ., 1965, 5, № 1.

3. *Михалевич В.С., Шор Н.З.* О численных методах решения многовариантных плановых и технико-экономических задач./Научно-метод. материалы экономико-матем. семинара ВЦ АН УССР. Вып. 1. Киев, ВЦ АН УССР, 1962.

4. *Багаева Н.Я.* Решение двух неклассических задач оптимального управления./Докл. на II Всес. съезде по теор. и прикл. механ. Москва, 1964.

5. *Багаева Н.Я., Моисеев Н.Н.* Об одном способе численного решения задач оптимального управления./Докл. АН СССР, 1963, 155, № 4, 747–750.

1964 г.

Информационная теория иерархических систем

Предисловие

Наряду с классическим определением понятия «информация», которое сделалось источником обширных математических исследований, очень важна, по мнению автора, и прагматическая концепция этого термина. Представление об информации как об абстрактной категории, имеющей меру (энтропию), которая является объектом самостоятельного исследования, заменяется пониманием об информации как об альтернативе неопределенности, которая неизбежно присутствует при любом акте принятия решения. Значит, информация ценна не сама по себе, а в той степени, в какой она помогает достигать целей, ради которых изучается данный управляемый процесс. Такая точка зрения имеет право на существование, поскольку она оказывается эффективной для анализа многих прикладных задач. Более того, она оказывается тем отправным пунктом, который позволяет развить математическую теорию управляемых систем, обладающих иерархической структурой.

Легко заметить, что все более или менее сложные управляемые системы, будь то технические, экономические, социальные или военные, всегда наделены иерархической структурой. Они разделены на подсистемы и отдельные звенья, обладающие самостоятельными правами принятия решений.

Анализ (а тем более синтез) систем, наделенных иерархической структурой, не сводится к классической теории оптимальных систем. И дело здесь даже не в том, что во многих случаях функционирование иерархических систем описывается операторами более сложной природы, нежели дифференциальные. В этих системах изменяется само понятие оптимальности. В самом деле, что значит оптимальная система, если каждое ее звено имеет возможность самостоятельно принимать решения, в частности, максимизировать свой собственный функционал?

Таким образом, иерархическая система — это принципиально многокритериальная система. И, следовательно, первое требование к любой теории, описывающей эволюцию управляемой системы, наделенной иерархической структурой, — найти рациональные математические постановки задач и вложить разумный смысл в понятие оптимальности.

Но прежде чем строить такую теорию, по-видимому, необходимо сначала дать ясный ответ на вопрос, а для чего вообще нужна иерархическая структура, неизбежно представляющая из себя систему дополнительных ограничений и, следовательно, сужающая допустимую область функционирования системы в целом.

Вот для решения этого «нулевого вопроса» и необходима та прагматическая концепция понятия информации, о которой мы уже сказали несколько слов. Только учет разного «уровня информированности» различных элементов системы делает понятным необходимость ее структуризации. Впервые такой информационный подход к анализу иерархических систем был предложен в совместной работе Ю.Б. Гермейера и автора и опубликован в 1971 г.*

В предлагаемом докладе мы развиваем эту концепцию, обращая внимание, главным образом, на содержательный анализ примеров и обсуждение возможностей численного анализа.

§ 1. Предварительный анализ

1. *Зачем нужна иерархия в системах управления?* Введение в данной динамической управляемой системе иерархической структуры имеет своим следствием распределение между различными ее звеньями функций обработки информации и принятия решений по выбору управляющих воздействий. На первый взгляд кажется, что создание иерархической структуры — это просто добавление новых ограничений, т. е. сужение множества допустимых управлений, которое неизбежно влечет снижение показателя эффективности системы в целом. Действительно, может оказаться, что введение иерархической структуры вредно или бессмысленно, но возможно и другое — без разделения функций принятия решений система вообще не сможет функционировать.

* См. Гермейер Ю.Б., Моисеев Н.Н. «О некоторых задачах теории иерархических систем управления» / Сб. «Проблемы прикладной математики и механики». — М., Наука, 1971.

Для того чтобы разобраться в этом вопросе и выработать принципы оценки качества самой структуры, автору кажется необходимым принять во внимание связь между критерием качества системы и требованиями к алгоритмам обработки информации.

Может оказаться, что полностью централизованный сбор и обработка информации либо технически невозможны, либо приводят к значительному запаздыванию в принятии решений, т. е. к принятию решения по устаревшей информации. В обоих случаях это приводит к увеличению неопределенности в процедурах принятия решения, а следовательно, к снижению эффективности систем управления.

Одним из путей преодоления трудностей, вызванных большим объемом информации или сложностью ее переработки, является «распараллеливание» процедур ее сбора и обработки. Однако децентрализация процесса обработки информации неизбежно требует определенного уровня децентрализации также и процедур принятия решений, т. е. создания самостоятельно функционирующих подсистем.

Появление в системе отдельных звеньев, способных за приемлемое время обработать всю относящуюся к ним информацию, учесть все изменения конкретной ситуации и получающих право принятия самостоятельных решений по тем или иным вопросам, и означает появление в системе иерархической структуры.

Для принятия решений в отдельных звеньях системы необходимо уже значительно меньший объем информации, который можно быстрее и лучше обработать и, следовательно, принять решение в условиях значительно меньшей неопределенности.

Таким образом, децентрализация управления приводит к уменьшению неопределенности, связанной со сбором и переработкой информации. Однако децентрализация в свою очередь служит источником новой неопределенности.

В самом деле, как только какая-либо подсистема получает в свое распоряжение право принятия решений, она превращается в самостоятельный организм, т. е. неизбежно приобретает собственные цели, в общем случае не тождественные интересам других подсистем и системы в целом.

Таким образом, основная причина неопределенности в принятии решений, возникающей вместе с образованием иерархии, — неизбежность противоречий между целым и его частями, которые

проявляются в первую очередь в нетождественности интересов всей системы и ее отдельных звеньев.

Вот поэтому можно говорить об оптимальной мере централизации и децентрализации, об оптимальном распределении функций принятия решений между различными уровнями систем.

Такой подход к анализу иерархических систем управления естественно назвать информационным. Он позволяет понять содержание целого ряда задач теории иерархических систем управления, их специфику.

Самый первый вопрос такой — что значит управление в иерархической системе? Далее, какой смысл должен быть вложен в понятие «оптимальность»? Как проектировать, т. е. выбирать рациональную структуру (или наилучшую)? И много других вопросов, которые не имеют аналогий в классических задачах теории управления.

Изложение будет носить индуктивный характер. Мы начнем с неформального анализа простейших примеров, а уже позднее дадим описание тех классов задач, для которых может использоваться развиваемая теория.

Примечание. Информационная концепция построения иерархической структуры систем управления имеет в качестве следствия следующее утверждение: вместе с совершенствованием технологии передачи и обработки информации уровень централизации растет.

2. Информационная концепция иерархической структуры позволяет изучить целый ряд важных особенностей. Однако она не способна учесть социологические и социальные аргументы, которые наряду с характеристиками процедур обработки информации могут оказаться решающими при выборе структуры.

Пример I. Рассмотрим объединение N промышленных предприятий (трест, синдикат), выпускающих однотипную продукцию. Это объединение будем называть Центром. Обозначим через P_i продукцию*, выпускаемую предприятием (производителем) i , и для упрощения рассуждений будем считать величины P_i скалярами. Результат функционирования Центра определяется результатами

* Под буквой P обозначающей продукцию, будем понимать как ее наименование, так и объем.

функционирования отдельных производителей. Оценки второго результата могут быть самыми разными. Так, например, если Центр целиком бюрократическая (управленческая) организация, то оценка его работы, например зарплата служащих, будет просто функцией P_1, \dots, P_N ($J(P_1, \dots, P_N)$). Приведем два возможных типа целевой функции подобного рода.

а) Производится дефицитный продукт, полностью потребляемый рынком; тогда, каково бы ни было плановое задание $P \leq P^*$, где $P = \sum p_i$ целесообразно производить максимально возможный объем продукта P . Следовательно, в этом случае функция J имеет вид:

$$J = J^{(a)} \left(\sum_{i=1}^n P_i \right), \quad (1.1a)$$

где J^a — некоторая монотонная функция; чем больше J , тем выше оценивается (и оплачивается) деятельность Центра.

б) Производится недефицитный продукт, или комплектующий продукт, или, наконец, продукт, на производство которого расходуется дефицитный продукт. В этом случае план должен быть строго выполнен. Его перевыполнение или приведет к омертвлению капитала (в первых двух случаях), или к тому, что некоторый другой продукт, может быть, более важный, не будет произведен в достаточном количестве, потому что сырье перерасходовано на производство продукта P . В этом случае функция J имеет вид

$$J = J^{(b)} \left\| \sum_{i=1}^N P_i \right\|, \quad (1.16)$$

Чем меньше значение монотонно возрастающей функции $J^{(b)}$, тем выше оценивается деятельность Центра.

в) Рассмотрим, наконец, еще один случай, когда Центр является хозрасчетной фирмой, синдикатом. В этом случае функция цели J будет доходом. Доход Центра будет складываться из стоимости

суммарного продукта $c \sum_{i=1}^N P_i$, где c — цена единицы продукта P за

вычетом производственных затрат и инвестиций. Рассмотрим тот частный случай, когда ресурс U , используемый для инвестиций, — экзогенный, т. е. не производящийся внутри системы, и его величина фиксирована

$$\sum_{i=1}^N u_i \leq U. \quad (1.2)$$

Здесь u_i — инвестиционные вложения Центра в его предприятия. Таким образом, ресурс U не участвует в формировании функции, характеризующей доход.

При учете производственных затрат ограничимся только затратами на рабочую силу

$$Q = \sum_{i=1}^N L_i,$$

где L_i — количество рабочих на предприятии номера i , w_i — ставка заработной платы на предприятии того же номера.

В этом случае целевая функция Центра будет иметь вид

$$J^{(a)} = c \sum_{i=1}^N P_i - \sum_{i=1}^N w_i L_i. \quad (1.1b)$$

Центр не имеет права декретировать объемы производства P_i , но он может влиять на них. Предположим, что у Центра есть два способа воздействий на политику производителей:

а) Распределение экзогенного ресурса. Это значит, что Центр имеет право назначать величины инвестиций — u_i .

б) Установление фонда заработной платы. Будем считать, что суммарный фонд заработной платы ограничен

$$Q \leq Q^*. \quad (1.3)$$

Центр имеет право назначать объем фонда для производителей, т. е. устанавливать ограничения вида

$$L_i \leq Q_i^*, \quad i = 1, \dots, N \quad (1.4)$$

причем

$$\sum_{i=1}^N Q_i^* \leq Q^*. \quad (1.3)$$

Рассмотрим теперь действия производителя. Обозначим через x_i фонды Производителя i . предположим еще, что производственный процесс описывается некоторой производственной функцией,

которую для определенности мы будем принимать в форме Кобба—Дугласа:

$$P_i = d_i x_i^{\alpha_i} L_i^{1-\alpha_i}, \quad \alpha_i \in [0, 1] \quad (1.5)$$

Так как по условию производитель i получит от Центра ресурс u_i , то физический объем продукта, который он произведет, будет

$$P_i = d_i (x_{i0} + u_i)^{\alpha_i} L_i^{1-\alpha_i}, \quad (1.6)$$

где x_{i0} — объем фондов в предыдущий плановый период.

Условимся, что доход производителя J_i будет равен стоимости его продукции cP_i за вычетом накладных расходов, которые заключаются только в оплате рабочей силы $\beta_i L_i$, где β_i — ставка заработной платы. Таким образом,

$$J_i = cd_i (x_{i0} + u_i)^{\alpha_i} L_i^{1-\alpha_i} - \beta_i L_i.$$

Задача Центра — так распорядиться ресурсом U и фондами Q_i^* , чтобы максимизировать функцию (1.1). Задача производителя — иметь такое количество рабочих L_i , чтобы максимизировать доход (1.7).

Таким образом, проблема планирования в иерархической системе требует анализа конфликтной (игровой) ситуации: здесь имеется $N + 1$ игроков — Центр и N производителей. Однако эта игра не является антагонистической. Ее участники имеют просто не совпадающие интересы.

Для анализа нам необходима гипотеза информированности. Условимся, что Центр знает производственные функции производителей. Кроме того, Центру необходима еще гипотеза поведения производителя. Условимся считать, что производитель действует разумно. Это значит, что получив ресурс u_i , он так выберет свое управление L_i , чтобы максимизировать доход (1.7). Решение этой задачи при фиксированных x_{i0} и u_i всегда существует, и, следовательно, так как функция $J_i(L_i)$ выпукла вверх, его легко получить в

явном виде из условия $\frac{J_i}{L_i} = 0$:

$$L_i = \hat{c}_i(x_{i0} + u_i), \quad (1.8)$$

где

$$\hat{c}_i = \frac{cd_i(1 - \kappa_i)^{\frac{1}{\kappa_i}}}{i}$$

Решение (1.8) справедливо только для того случая, когда ограничение (1.4) отсутствует. В общем случае решение производителя будет таким:

$$L_i = \begin{cases} \hat{c}_i(x_{0i} - u_i), & \text{если } {}_i\hat{c}_i(x_{0i} - u_i) \geq Q_i^*, \\ \frac{Q_i^*}{i}, & \text{если } {}_i\hat{c}_i(x_{0i} - u_i) < Q_i^*. \end{cases} \quad (1.9)$$

Таким образом, гипотеза о поведении производителя позволяет определить его управление как функцию u_i , т. е. управления Центра. Объем продукта, который производит производитель i , легко подсчитывается по формуле (1.5):

$$P_i = \begin{cases} d_i \hat{c}_i^{1 - \kappa_i}(x_{0i} - u_i), & \text{если } {}_i\hat{c}_i(x_{0i} - u_i) \geq Q_i^*, \\ d_i (x_{0i} - u_i)^{\kappa_i} \frac{Q_i^{*1 - \kappa_i}}{1 - \kappa_i}, & \text{если } {}_i\hat{c}_i(x_{0i} - u_i) < Q_i^*. \end{cases} \quad (1.10)$$

Имея в своем распоряжении выражение (1.10), уже легко вычислить выражение для целевой функции Центра. Так как согласно (1.10) $P_i = p_i(u_i, Q_i^*)$, то и выражение функции $J(p_1, \dots, p_N)$ примет вид:

$$J = J^*(u_1, \dots, u_N, Q_1^*, \dots, Q_N^*). \quad (1.11)$$

Задача Центра состоит теперь в том, чтобы так назначить управляющие воздействия u_i и Q_i^* , чтобы функция (1.11) достигла максимального (или минимального) значения при условиях (1.2) и (1.3). В том случае, когда функция цели имеет вид (1.1в), а ограничение (1.3) отсутствует, доход Центра будет

$$J = \sum_{i=1}^N g_i(x_{0i} - u_i), \quad (1.12)$$

где

$$g_i = \frac{\kappa_i}{1 - \kappa_i} {}_i\hat{c}_i.$$

Теперь Центру уже легко решить задачу планирования. Он должен максимизировать форму (1.12) при условии (1.2). Решение этой задачи тривиально. Обозначим через j то из значений i , которому

$$u_i = \begin{cases} U, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases} \quad (1.13)$$

Задача планирования решена полностью. Разумеется, в общем случае задача о максимизации (1.11) при условиях (1.2) и (1.3) — это сложная задача нелинейного программирования.

3. *Цена децентрализации.* Итак, мы видим, что выбор управления в иерархической системе сводится к некоторой игровой и достаточно сложной задаче. Однако вводя гипотезу поведения, мы сводим игровую задачу к специальной задаче оптимизации. Такая процедура, по существу, аналогична использованию принципа гарантированного результата в теории антагонистических игр, причем в данном случае она даже еще более оправдана.

Теперь мы можем ввести понятие оптимального управления иерархической системой как управления, доставляющего экстремальное значение целевой функции Центра, в предположении, что управление производителей определено на основании гипотезы их поведения. Оптимальное значение целевой функции Центра обозначим через J^* .

Предположим теперь, что в системе управление полностью централизовано. Это значит, что Центр может распоряжаться не только распределением экзогенного ресурса и фондом заработной платы, но и количеством нанимаемых рабочих. Оптимальное значение целевой функции Центра в этом случае обозначим через J^{**} . Очевидно, что

$$J^{**} \geq J^*. \quad (1.14)$$

Введем еще величину

$$\frac{J^{**} - J^*}{J^*} \geq 0. \quad (1.15)$$

Эту величину естественно назвать ценой децентрализации — это цена, которую платит Центр и которая характеризует уменьшение его дохода вследствие организации иерархической структуры. Она всегда неотрицательна, т. е. в рассматриваемом примере децентрализация невыгодна. Ниже мы увидим иную постановку задачи, из которой будет следовать уже необходимость использования иерархической структуры.

Заметим, что легко привести примеры, когда $\theta = 0$ т. е. эффективность функционирования системы не зависит от ее структуры.

Предположим, например, что речь идет о функции (1.1в) и ограничения на фонд заработной платы отсутствуют. Тогда оптимальным управлением Центра будет управление (1.13), независимо от того, введена ли иерархическая структура принятия решений или нет. Этот пример интересен тем, что в нем оптимальное управление производителя, доставляющее максимальное значение функции цели производителя, совпадает с тем управлением, которое выберет для него Центр, если он один будет ответствен за выбор всех управлений. В этом случае интересы (а следовательно, и действия) любого из производителей не противоречат интересам Центра. Но это, разумеется, не значит, что их интересы тождественны. У Центра иные задачи, и он располагает иными возможностями для их решения. После того как Центр выделит производителю то или иное количество ресурса, последний использует его наилучшим для Центра способом.

4. *Управление с помощью штрафов и поощрений.* Пример II. В примере I были описаны два способа управления. Первый — это управление с помощью экзогенного ресурса. Одна из особенностей этого управления состоит в том, что производитель получает его в свое распоряжение в строго фиксированном количестве, не зависящем от способов его использования.

Второй способ — это управление с помощью задания ограничений (фонд заработной платы). В одном отношении оба эти способа аналогичны. Они сообщались заранее производителю и никак не зависели от характера решений, которые он примет.

Рассмотрим теперь пример, в котором Центр имеет возможность непосредственно влиять на целевую функцию, меняя ее значение в зависимости от того управления, которое принимает производитель.

Сохраним условие примера I, будем считать, что деятельность производителя описывается производственной функцией Кобба—Дугласа (1.5), причем положим $\kappa_i = \frac{1}{2} (i = 1, \dots, N)$.

Целевую функцию производителя мы запишем в виде:

$$J_i(L_i) = cd_i x_i^{1/2} L_i^{1/2} - \tilde{\gamma}_i(P_i). \quad (1.16)$$

Здесь $\tilde{\gamma}_i(P_i)$ — дополнительное вознаграждение (или штраф), которое выплачивается Центром производителю в зависимости от результатов его деятельности; $\tilde{\gamma}_i(P_i) \in G_{\tilde{\gamma}_i}$, где $G_{\tilde{\gamma}_i}$ — некоторые множества.

Величины x_i будем считать фиксированными и писать (1.16) в виде

$$J_i(L_i) = c_i L_i^{1/2} - \tilde{\gamma}_i(P_i). \quad (1.17)$$

Гипотезу поведения производителя оставим старой. Это значит, что он выбирает свое управление из условия максимума функции (1.17). Следовательно, в интересах Центра форма штрафа или поощрения должна сообщаться производителю заранее, т. е. производитель должен быть проинформирован о структуре функции $\tilde{\gamma}_i(P_i)$.

Условие максимума $J_i(L_i)$ позволяет определить значение L_i^* , доставляющее максимум функции $J_i(L_i)$. L_i^* будет функционалом, зависящим от вида функции $\tilde{\gamma}_i$:

$$L_i^* = L_i^*(\tilde{\gamma}_i).$$

Точно так же функционалом от $\tilde{\gamma}_i$ будет и оптимальный объем продукта P_i^* :

$$P_i^* = P_i^*(\tilde{\gamma}_i).$$

Следовательно, задача определения оптимального управления Центра — это задача такого определения штрафов $\tilde{\gamma}_i(P_i)$, которые доставляют экстремальное значение функции цели Центра $J(P_1, \dots, P_N)$, которую мы теперь перепишем в виде:

$$J = J(P_1, \dots, P_N). \quad (1.18)$$

Задача определения экстремумов функционала (1.18) является сложной и нестандартной задачей оптимизации, поскольку сами функционалы J_i определяются из решения оптимизационной задачи. Для того очень простого и частного случая иерархической системы, который мы рассматриваем, решение задачи эффективного определения оптимального штрафа требует изобретения специальных методов.

Если функция цели Центра явно не зависит от штрафов и если на величину этих штрафов не накладывается никаких ограничений, то решение этой задачи не единственно и тривиально.

В самом деле, пусть $P_1^*, P_2^*, \dots, P_N^*$ — объемы продукции производителей, которые доставляют максимальное значение функции цели Центра. Тогда мы можем систему штрафов принять, например, в виде:

$$\tilde{J}_i(P_i) = (P_i - P_i^*)^2 - cd_i x_i^{1/2} L_i^{1/2} - \alpha_i L_i, \quad (1.19)$$

где α_i — произвольное число. В этом случае

$$J_i = (P_i - P_i^*)^2.$$

Следовательно, производитель так будет распоряжаться своими ресурсами, чтобы объем его продукции был равен P_i^* . Структура штрафа (1.19) делает интересы производителя не противоположными интересам Центра.

Легко видеть, что функций вида (1.19), которые делают интересы производителей совпадающими с интересами Центра, может быть построено бесчисленное множество.

Однако неограниченный штраф или поощрение вряд ли могут представлять какой-либо практический интерес. В задачах, «достаточно реально» поставленных, величина штрафа либо ограничена, т. е. удовлетворяет условию вида

$$(\tilde{J}_1, \dots, \tilde{J}_N) \in G, \quad (1.20)$$

где G — некоторое множество, либо значение функции J зависит от функций \tilde{J}_i :

$$J = J(P_1, \dots, P_N; \tilde{J}_1, \dots, \tilde{J}_N). \quad (1.21)$$

В том случае, когда на функции штрафа наложены ограничения вида (1.20) или целевая функция Центра явно зависит от функций штрафа, задача определения максимального штрафа становится трудной задачей оптимизации.

Заметим, что основная сложность здесь та же, что и в задачах синтеза. Необходимо отыскивать функции $\tilde{~}_i(P_i)$. Поэтому так же, как и в задачах синтеза, функции $\tilde{~}_i$ следует считать принадлежащими к некоторому достаточно простому классу, элементы которого определяются небольшим числом параметров.

Рассмотрим более подробно эти два последних случая. Предположим, что Центр представляет собой некоторое производственное объединение, выпускающее однотипную продукции, т. е. некоторый аналог односекторной экономики. Тогда оценка его деятельности будет связана с некоторой функцией, зависящей только от объема суммарного продукта, созданного производителями $F(P_i)$. Кроме того, предположим, что существует некоторый «фонд поощрения» R , причем

$$R = \sum_{i=1}^N \tilde{~}_i(P_i). \quad (1.22)$$

Рассмотрим последовательно два случая.

А) Фонд поощрения R является экзогенным, т. е. он предоставляется в распоряжение Центра каким-либо высшим органом, тогда целевая функция Центра будет

$$J(P_1, \dots, P_N) = F \sum_{i=1}^N P_i. \quad (1.23)$$

Б) Фонд поощрения выделяется из «дохода Центра», и тогда

$$J = J(P_1, \dots, P_N; \tilde{~}_1, \dots, \tilde{~}_N) = F \sum_{i=1}^N P_i - R(\tilde{~}_1, \dots, \tilde{~}_N). \quad (1.24)$$

Функцию F мы условимся брать в следующем виде

$$F = \sum_{i=1}^N P_i - c_1 \sum_{i=1}^N P_i - c_2 \sum_{i=1}^N P_i^2. \quad (1.25)$$

Здесь c_1 — некоторая оптовая цена, P — некоторый заданный (допустимый, требуемый) объем валового продукта. Если производственный продукт

равен $\sum_{i=1}^N P_i - P$, то возникают дополнительные расходы.

В зависимости от технического или экономического содержания проблемы смысл этого слагаемого может быть самый разный.

Доход Центра в форме (1.25) охватывает довольно широкий класс задач, встречающихся на практике. Прежде всего это про-

блема комплектных поставок. Если совокупный продукт $\sum_{i=1}^N P_i \leq P$,

то, следовательно, более высокий уровень иерархии, который комплектует изделия, не сможет их поставить в нужном количестве.

Если $\sum_{i=1}^N P_i > P$, то избыток произведенного продукта (например,

деталей) не сможет быть использован и будет заморожен. В обоих

случаях величина $c_2 \sum_{i=1}^N P_i - P$ может рассматриваться как штраф

за невыполнение комплектных поставок. Величина c_2 может считаться величиной постоянной.

Подобная задача возникает и в тех случаях, когда продукт является недефицитным и известна емкость рынка или объем контрактных поставок. В этом случае P и будет тем объемом продукта, который должен быть произведен. Если произведено продукта

$\sum_{i=1}^N P_i > P$, то возникают дополнительные затраты, например, на его

хранение или вследствие замораживания капитала. В этом случае

$$c_2 = \begin{cases} 0, & \text{если } \sum_{i=1}^N P_i \leq P, \\ \bar{c}_2 \sum_{i=1}^N P_i - P, & \text{если } \sum_{i=1}^N P_i > P, \end{cases}$$

где \bar{c}_2 — постоянная. Здесь $c_2 \sum_{i=1}^N P_i - P$ — либо некоторая постоянная, либо вогнутая функция аргумента.

Рассмотрим сначала случай А. Фонд поощрения — величина экзогенная, и, следовательно, суммарный штраф (или поощрение) должен удовлетворять условию (1.22).

Обозначим через P_i^* объем продукта, который, с точки зрения Центра, должен произвести производитель i , и функцию штрафа примем в виде:

$$\tilde{~}_i(P_i) = k_i(P_i - P_i^*)^2 - e_i P_i. \quad (1.26)$$

Тогда фонд производителя будет, согласно (1.16):

$$J_i(L_i) = (c - e_i) d_i x_i^{1/2} L_i^{1/2} - k_i(P_i - P_i^*)^2 - e_i L_i, \quad (1.27)$$

где

$$c_i = (c - e_i) d_i x_i^{1/2}, \quad k_i = d_i x_i^{1/2}.$$

Управление производителя определяется из условия: $\frac{J_i}{L_i} = 0$ или, в явном виде:

$$L_i = \frac{1}{4} \frac{c_i - 2k_i P_i^*}{k_i} \frac{1}{P_i^*}. \quad (1.28)$$

И следовательно, целевая функция Центра будет

$$J = c_1 \prod_{i=1}^N \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(k_i d_i^2 x_i)} - c_2 \prod_{i=1}^N \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(k_i d_i^2 x_i)} P_i^{*2}. \quad (1.29)$$

Объем «плановых заданий» для производителей P^* , коэффициент поощрений e_i и штрафов k_i находятся в распоряжении Центра. Центр может ими воспользоваться для оптимизации функции (1.29) при ограничении (1.22), которое в данном случае имеет вид:

$$\prod_{i=1}^N k_i \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(k_i d_i^2 x_i)} P_i^{*2} = R. \quad (1.22)$$

Таким образом, задача планирования Центра сведена к обычной задаче отыскания экстремума (1.29) при условиях типа равенства (1.22) и ограничениях вида

$$P_i^* \geq 0, \quad k_i \geq 0, \quad e_i \geq 0, \quad i = 1, \dots, N. \quad (1.30)$$

Если коэффициент c_2 имеет вид (1.25), то это достаточно сложная, но вполне обозримая задача нелинейного программирования.

Рассмотрим теперь случай Б, в котором фонд поощрения R не является экзогенным, а формируется из дохода Центра. Этот факт, очевидно, не отражается на политике производителя, который независимо от этого фонда выбирает свое управление, следуя формуле (1.28), как только ему сообщена структура штрафа (1.26). Следовательно, в этом случае задача сводится к выбору величин P_i^* , e_i и k_i , которые максимизируют функцию

$$\begin{aligned} J(P_1^* \dots, P_N^*, e_1, \dots, e_N, k_1, \dots, k_N) \\ = \sum_{i=1}^N (c_1 - e_i) \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(c - k_i d_i^2 x_i)} \\ - c_2 \sum_{i=1}^N \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(c - k_i d_i^2 x_i)} P_i^* \\ - \sum_{i=1}^N k_i \frac{d_i^2 x_i (c - e_i - 2k_i P_i^*)}{2(c - k_i d_i^2 x_i)} P_i^* \end{aligned} \quad (1.31)$$

при ограничениях (1.30).

Задача (1.30)—(3.31) является достаточно простой задачей квадратичного программирования.

Примечание. Функцию штрафа (1.26) можно представить в виде

$$\tilde{\gamma}_i(P_i) = a_i L_i - b_i L_i^{1/2} - l_i, \quad (1.26)$$

где

$$a_i = k_i d_i x_i,$$

$$b_i = e_i d_i x_i^{1/2} - 2k_i d_i x_i^{1/2} P_i^*,$$

$$l_i = k_i P_i^{*2}.$$

Легко видеть, что выбор стратегии производителя L_i не будет зависеть от параметра l_i ; таким образом, из трех параметров, которыми определяется штраф (1.26), только два являются существенными.

Если функция цели Центра зависит только от величин P_i , то параметры l_i можно отнести к ограничениям.

§ 2. Общая статическая задача

1. *Постановка задачи.* Рассмотрим теперь общий случай статической задачи, сохраняя для удобства терминологию, которую мы использовали при рассмотрении примеров. Мы снова будем говорить о Центре и производителях, хотя то, о чем будет говориться ниже, относится не только к экономическим, но и к техническим системам.

Итак, будет изучаться система, состоящая из N производителей, результат функционирования которых оценивается функцией

$$J_i = J_i(u_i, i, i(i), i), \quad (2.1)$$

где u_i — экзогенный ресурс, находящийся в распоряжении Центра. Он стеснен ограничением

$$u_i \leq G_u, \quad (2.2)$$

например,

$$\sum_{i=1}^N u_i \leq U,$$

i — вектор управлений производителя. Он может быть стеснен ограничениями вида

$$i \leq G_i, \quad (2.3)$$

причем структура множества G_i в общем случае может зависеть от решений Центра, т. е. находиться в его распоряжении. $i(i)$ — это также управления Центра. Значения функций $i(i)$ определяются действиями производителя, его управлением i . На выбор этих функций наложены ограничения, которые мы будем иметь в виде:

$$i \in G_i. \quad (2.4)$$

Наконец, i — это некоторые внешние характеристики процесса. Они могут быть известны производителю i точно или с какой-то степенью точности. Будем их считать случайными величинами, принадлежащими некоторому классу.* Этот факт будем записывать так:

$$i \in g_i. \quad (2.5)$$

Интересы производителя — это максимизация функции J_i .

Опишем теперь Центр. Его целевая функция зависит от тех же параметров

$$J = J(u_1, \dots, u_N, \mu_1, \dots, \mu_N, \mu_1(i_1), \dots, \mu_N(i_N), \mu_1, \dots, \mu_N, G_1, \dots, G_N)$$

Однако уровень информации о случайных параметрах i у него иной:

$$i \in G_i. \quad (2.6)$$

Центр хуже информирован об особенностях процесса. Это значит, что множества G_i более широкие, чем g_i . Последнее обстоятельство условимся записывать в виде:

$$g_i \subset G_i. \quad (2.7)$$

Именно это обстоятельство и определяет создание иерархической системы.

* В частности, мы можем считать ее величиной неопределенной, например, если i — скаляр, то $i \in [a, b]$ и равномерно распределена на этом отрезке, или гауссовские моменты ее принадлежат данному отрезку.

Для того чтобы теперь получить математическую задачу, мы должны условиться о «правилах игры». Условимся считать, что Центр всегда делает первый ход: он сообщает производителю

- а) значение экзогенного ресурса u_i ,
- б) структуру функции штрафа, т. е. функцию $\varphi_i(\cdot)$,
- в) характер ограничений G_i .

Теперь производитель должен принять решение, т. е. распорядиться вектором x_i в условиях неопределенности. Условимся считать, что Центр всегда знает принцип принятия решения производителем. Эту гипотезу мы условимся обозначать через Γ_i и писать ее в следующем виде:

$$\hat{J}_i = \max_{x_i \in G_i} J_i(u_i, \varphi_i(x_i), x_i).$$

В качестве естественных примеров гипотез Γ_i можно привести следующие:

- а) если u_i — величина, известная производителю, то гипотеза Γ_i — оптимизация J_i по x_i ;
- б) если φ_i — случайный гауссовский вектор и его математическое ожидание $\bar{\varphi}_i$ известно производителю, то гипотеза Γ_i — оптимизация по x_i при $\varphi_i = \bar{\varphi}_i$ т. е.,

$$\hat{J}_i = \max_{x_i \in G_i} J_i(u_i, \varphi_i(x_i));$$

- в) если $\varphi_i \in g_i$, где g_i — ограниченное замкнутое множество и φ_i равномерно распределено в g_i , то гипотеза Γ_i — это максимизация по x_i гарантированного результата по φ_i , т. е.

$$\hat{J}_i = \max_{x_i \in G_i} \min_{\varphi_i \in g_i} J_i(u_i, \varphi_i(x_i), x_i).$$

После того как гипотеза Γ_i , сформулирована, проблемы управления, которые должны решаться Центром, превращаются в некоторые специальные задачи оптимизация. В самом деле, гипотезы Γ_i позволяют определить управления производителей — величины x_i — как функции управлений Центра и параметра φ_i :

$$x_i = x_i(u_i, \varphi_i(x_i), G_i, \varphi_i). \tag{2.8}$$

Таким образом, функционал J зависит только от величия, которыми может распоряжаться Центр, и случайных факторов i :

$$J = J^*(u_1, \dots, u_N, \Gamma_1, \dots, \Gamma_N, G_1, \dots, G_N). \quad (2.9)$$

Поскольку выражение (2.8) определяет некоторый оператор, то задача оптимизации функционала (2.9), конечно, не относится к числу стандартных. Наконец, для того чтобы сделать задачу определенной, Центр должен еще принять решение о принципах выбора собственных стратегий, поскольку функционал содержит неконтролируемые параметры i . Здесь также очень трудно дать какие-либо общие рекомендации. В одних случаях Центр может использовать гипотезы гарантированного результата, в других случаях — вводить ту или другую функцию риска и т. д. Центр имеет права оперирующей стороны, и формулировка принципа обращения с неконтролируемыми параметрами i находится полностью в компетенции Центра и лежит вне задачи, как и гипотеза поведения Γ_i , но она должна быть четко сформулирована Центром, без чего дальнейший анализ становится бессмысленным. Гипотезу о поведении Центра будем обозначать через Γ . Таким образом, для того чтобы провести анализ задачи, необходима система гипотез $\{\Gamma, \Gamma_1, \Gamma_2, \dots, \Gamma_N\}$.

Прежде чем переходить к анализу задач, сделаем некоторые замечания.

а) Представляют практический интерес те ситуации, когда Центр не полностью информирован об интересах, т. е. о целевых функциях производителей. Этот случай полностью вкладывается в описанную схему: параметры i мы можем трактовать как величины, которые описывают степень неопределенности наших знаний о целевых функциях производителей.

б) Еще раз подчеркнем, что существование самостоятельных интересов у производителей — это объективный факт. Эти интересы возникают немедленно, как только производитель получает права самостоятельного организма, т. е. права самостоятельно принимать решения. Центр может воздействовать на эти интересы, изменять их в нужном направлении, расширять или сужать возможности производителей в достижении этих целей.

в) Одной из особенностей иерархических систем управления является возможность управлять с помощью ограничений, возможность сужать или расширять права предприятий, входящих в

объединение. В конкретных случаях это может быть разрешение или запрет некоторых технологий, ограничение на трудозатраты и т. д. Необходимость подобных ограничений как раз и является следствием нетождественности интересов Центра и производителей, следствием невозможности сделать эти интересы вполне совпадающими.

г) Описанная модель является весьма общей и охватывает целый ряд специальных особенностей, имеющих практический интерес. Так, например, в ее рамках легко учитывается стоимость переработки информации, качество информации и т. д.

2. *Задача об экзогенном ресурсе.* В силу большой общности модели трудно говорить о каких-либо общих достаточно эффективных численных методах ее анализа. Поэтому мы рассмотрим ряд последовательно усложняющихся примеров. Начнем сначала с рассмотрения задачи об экзогенном ресурсе.

Предположим, что интересы Центра описываются функцией

$$J = J(u_1, \dots, u_N), \quad (2.10)$$

а интересы производителей — функциями

$$J_i = J_i(u_i, v_i), \quad (2.11)$$

причем

$$(u_1, \dots, u_N) \in G_u, \quad (2.12)$$

а на управление v_i не накладывается никаких ограничений, т. е.

$$v_i \in R^{n_i}. \quad (2.13)$$

В качестве гипотезы поведения Γ_i примем гипотезу о максимизации функции (2.11). Тогда условие, которому необходимо должно удовлетворять управление v_i , будет

$$\frac{\partial J_i}{\partial v_i} = 0. \quad (2.14)$$

Условие (2.14) можно представить в виде

$$J_i(u_i) \leq 0, \quad i = 1, \dots, N. \quad (2.15)$$

Таким образом, задача Центра сводится к отысканию векторов u_1, \dots, u_N , доставляющих максимум функции (2.10) при условиях (2.12), (2.15), а это стандартная задача нелинейного программирования.

К аналогичному результату мы придем и в том случае, если величины u_i подчинены ограничениям, допускающим использование функций штрафа. В этом случае вместо функции $J_i(u_i, u_i)$ мы будем рассматривать некоторую функцию $\tilde{J}_i(u_i, u_i, \tilde{u}_i)$, где \tilde{u}_i — коэффициент при функции штрафа.

Большое практическое значение имеют различные варианты теории возмущения, позволяющие уточнять приближенный, диспетчерский метод решения. Рассмотрим одну из возможных схем теории возмущений для того случая, когда на управление u_i не наложено никаких ограничений или они сняты введением функции штрафа. Предположим, что Центр распорядился тем или иным образом своим ресурсом, т. е. назначил некоторые величины $u_i \approx \tilde{u}_i$. Точно так же и производители выбрали свои решения $u_i \approx \tilde{u}_i$. Каким образом «улучшить» это решение?

Положим

$$u_i = \tilde{u}_i + \delta u_i, \quad (2.16)$$

Подставляя (2.16) в (2.11), заменим это выражение следующим:

$$J_i(\tilde{u}_i + \delta u_i, \tilde{u}_i + \delta u_i) = J_i(\tilde{u}_i, \tilde{u}_i) + J_{i, u_i}(\tilde{u}_i, \tilde{u}_i) \delta u_i + \frac{1}{2} J_{i, u_i u_i}(\tilde{u}_i, \tilde{u}_i) \delta u_i^2 + o(\delta^2),$$

где $o(\delta^2)$ — совокупность членов, не содержащих управление u_i , и членов, порядок которых более высокий, чем $\|\delta u_i\|^2$.

На основании гипотезы Γ_i цель производителя состоит в таком выборе u_i , который доставляет минимум функции

$$J_i(u_i, u_i) = J_i(\tilde{u}_i, \tilde{u}_i) + \frac{1}{2} J_{i, u_i u_i}(\tilde{u}_i, \tilde{u}_i) \delta u_i^2 + o(\delta^2).$$

Отбрасывая последнее слагаемое и считая, что на u_i не наложено никаких ограничений, находим a_i в явном виде:

$$a_i = A_i u_i, \quad (2.17)$$

где вектор равен $a_i = J_{i, i} J_{i, i}$, а матрица $A_i = J_{i, i} J_{i, i} u_i$.

Заметим, что если производитель в ответ на ход Центра, т. е. на его информацию о $u_i = \tilde{u}_i$, выбрал свою стратегию согласно гипотезе Γ_j , то и $a_i = 0$ и $A_i = 0$, т. е. $a_i = 0$. Отсюда следует не только вывод о том, что в рассматриваемом случае данный вариант теории возмущений непригоден, но и более важный результат о малой эффективности управления с помощью экзогенного ресурса.

Если Центр хочет незначительно изменить стратегии производителей, например, в том случае, если незначительно изменилась его цель, т. е. его целевая функция, то он должен весьма существенно изменить характер распределения ресурса, поскольку из анализа выражения (2.17) следует, что

$$a_i = o(u_i^2).$$

3. *Задача о поощрениях.* Будем считать, что интересы Центра описываются функцией (2.10), а интересы производителя — функционалом:

$$J_i = J_i(u_i, \phi_i(u_i)), \quad (2.18)$$

где $\phi_i(u_i)$ — штраф или поощрение, зависящие от решений, которые принимает производитель. Таким образом, мы рассматриваем задачу, в которой целевая функция Центра не зависит от функций штрафа. Для того чтобы задача имела смысл, как мы видели, на размер штрафа или поощрения должно быть наложено некоторое ограничение. Для определенности будем считать штраф скалярным, а суммарный фонд поощрений заданным:

$$\sum_{i=1}^N \phi_i(u_i) = Q. \quad (2.19)$$

Примечание. Если $Q = 0$, то этот случай мы можем интерпретировать как перераспределение Центром дохода производителей в зависимости от того, насколько стратегия производителя способствует достижению целей Центра.

В предыдущем параграфе мы уже обратили внимание на те трудности, которые мы встречаем при попытках построить процедуры эффективного определения оптимального штрафа, т. е. функций $J_i(x_i)$. Поэтому, следуя идеям теории синтеза, будем отыскивать функцию $J_i(x_i)$, принадлежащую к некоторому классу функций, определяемому небольшим числом параметров. Так же, как и в предыдущем параграфе, будем искать $J_i(x_i)$ в форме квадратичной функции

$$J_i(x_i) = a_i x_i^2 + b_i x_i + c_i \quad (2.20)$$

Представление функции поощрения в форме (2.20) охватывает большое количество практически важных задач.

После замены (2.20) функционал (2.18) превращается в функцию, зависящую от x_i, a_i, b_i, c_i :

$$J_i = J_i(x_i, a_i, b_i, c_i) \quad (2.21)$$

Если на управление x_i не наложено никаких ограничений или если мы снимаем ограничение с помощью функции штрафа, то необходимое условие оптимальности будет следующим:

$$\frac{dJ_i}{dx_i} = \frac{J_i}{x_i} - \frac{J_i}{a_i} (b_i + 2c_i/x_i) = 0 \quad (x_i, a_i, b_i, c_i) \quad (2.22)$$

В этом случае задача отыскания оптимального штрафа сводится к отысканию величин x_i, a_i, b_i, c_i ($i = 1, \dots, N$), доставляющих максимум функции

$$J = J(x_1, x_2, \dots, x_N)$$

при ограничениях

$$\prod_{i=1}^N (a_i x_i^2 + b_i x_i + c_i) = Q,$$

$$f_i(x_i, a_i, b_i, c_i) = 0,$$

т. е. к стандартной задаче нелинейного программирования.

Таким образом, с вычислительной точки зрения обе задачи аналогичны: введением функций штрафа, снимающих ограничения на управление производителя, обе рассматриваемые задачи сводятся к некоторой задаче нелинейного программирования, содержащей ограничения типа равенств и неравенств. Однако с содержательной точки зрения их природа значительно отличается. Эта разница легко устанавливается при анализе с помощью теории возмущений.

Пусть $\hat{a}_i, \hat{b}_i, \hat{c}_i, \hat{x}_i$ — некоторые значения неизвестных, которые мы хотим уточнить, опираясь на наши предположения. Получим

$$\begin{aligned} a_i &= \hat{a}_i + \delta a_i, \\ b_i &= \hat{b}_i + \delta b_i, \\ c_i &= \hat{c}_i + \delta c_i, \\ x_i &= \hat{x}_i + \delta x_i. \end{aligned}$$

Тогда

$$\begin{aligned} J_i(\hat{x}_i) &= \frac{dJ_i}{dx_i}(\hat{x}_i) \delta x_i + \frac{dJ_i}{da_i}(\hat{x}_i) \delta a_i + \frac{dJ_i}{db_i}(\hat{x}_i) \delta b_i + \frac{dJ_i}{dc_i}(\hat{x}_i) \delta c_i + \\ &+ o(\delta x_i^2, \delta a_i^2, \delta b_i^2, \delta c_i^2). \end{aligned}$$

Согласно гипотезе Γ_i , величины δx_i выбираются из условия

$$\frac{dJ_i}{dx_i}(\hat{x}_i) = 0,$$

откуда для определения δx_i мы получаем следящее уравнение теории возмущений:

$$\begin{aligned} \frac{dJ_i}{dx_i}(\hat{x}_i) \delta x_i + 2 \frac{dJ_i}{da_i}(\hat{x}_i) \delta a_i + \frac{dJ_i}{db_i}(\hat{x}_i) \delta b_i + \frac{dJ_i}{dc_i}(\hat{x}_i) \delta c_i &= 0. \end{aligned} \quad (2.23)$$

Уравнение (2.23) показывает, что поправки на управление производителя являются линейными функциями поправок на управление Центра, причем этот вывод остается в силе даже и тогда, когда величина $\hat{\alpha}_i$ была выбрана оптимальным образом, т. е. так, что

$$\frac{J_i}{\alpha_i} = 0.$$

Отсюда следует важный вывод о том, что в окрестности оптимального управления управление с помощью функций поощрений значительно эффективнее управления с помощью распределения экзогенного ресурса.

Этот вывод получен на основе рассмотрения в «малом». Следовательно, он теряет свои силы для общего случая. Действительно, в окрестности оптимального решения Центру легче воздействовать на производителя, легче приспособить его функционирование к достижению своих целей. Однако если речь идет о глобальной перестройке стратегии Центра (требующей, например, крупных инвестиций для создания новых фондов), то может оказаться, что управление о помощью экзогенного ресурса будет эффективнее.

Здесь уже трудно давать какие-либо общие рекомендации.

4. *Замечание о неединственности.* Все предыдущие рассуждения имели смысл в том и только в том случае, когда уравнение

$$\frac{J_i}{\alpha_i} = 0 \tag{2.24}$$

определяет единственное решение или значения $J_i(\hat{\alpha}_i^s)$, где $\hat{\alpha}_i^s$ — корни уравнения (2.24) — не равны между собой. Но может оказаться так, что существует, например, две стратегии производителя $\hat{\alpha}_i^1$ и $\hat{\alpha}_i^2$, для которых

$$J_i(\hat{\alpha}_i^1) = J_i(\hat{\alpha}_i^2). \tag{2.25}$$

Это значит, что с точки зрения производителя обе стратегии $\hat{\alpha}_i^1$ и $\hat{\alpha}_i^2$ — эквивалентны. Однако, с точки зрения Центра, эти стратегии могут быть совсем разными. Простейший пример — различные наборы ассортимента изделий, имеющие одинаковую суммарную стоимость. С точки зрения производителя, эти наборы абсолютно одинаковы. Но Центр может быть заинтересован в комплектной поставке изделий. В этом случае, разумеется, Центр должен иметь

из этих наборов лишь один вполне определенный. Рассматриваемый случай не является патологией, он довольно типичен и в той или иной форме часто встречается в жизни.

Если имеет место равенство (2.25), то задача управления Центра значительно усложняется. Так, например, с помощью экзогенного ресурса повлиять на выбор производителем того из \hat{i} , который нужен Центру, просто невозможно. Здесь для управления принципиально необходим специальный штраф или поощрение.

5. *Замечание о дополнительных связях.* До сих пор мы считали, что функционирование производителя i определяется только его стратегией i и действиями Центра. Для анализа данной ситуации, для превращения ее в задачу математики нам было необходимо иметь набор гипотез о поведении производителей, мы их обозначили через Γ_i и гипотезу о поведении Центра. Но может оказаться, что целевая функция производителя i зависит еще и от стратегий некоторых других производителей:

$$J_i = J_i(i, j_1, \dots, j_s).$$

Этот случай также достаточно типичен и заслуживает подробного рассмотрения. Величины j_s могут трактоваться, например, как взаимные поставки.

Ситуация, которая теперь возникает, уже существенно сложнее рассмотренной ранее. Прежде всего необходима еще система Γ_{ij} — гипотез, которые должен формулировать производитель i о стратегии производителя j . Но наиболее трудными здесь будут вопросы управления и формирования гипотезы о поведении Центра. Один из частых и трудных вопросов, которые относятся к этому кругу, — что считать оптимальным планом для производителей и как сформулировать математические задачи его определения?

Решение этих вопросов требует, очевидно, неформального анализа.

6. *О роли информированности.* На протяжении всего изложения подчеркивался характер информированности. Центр всегда делал первый ход. Он заранее сообщал производителям те сведения, без которых производитель просто не мог принять определенного решения. Но заметим, что в условиях, когда Центр сформулировал гипотезы Γ_i , никакой другой информации о поведении производителя ему не нужно — никакая дополнительная информация не изменит характера его стратегии.

§ 3. Пример двухуровневой динамической задачи

1. *Особенности динамических задач.* Управление, в собственном смысле слова, появляется вместе с динамикой — когда возникает проблема целенаправленного влияния на процессы, которые разворачиваются во времени. Однако это вовсе не значит, что статические задачи не интересны для проблем управления. Они дают оценку для некоторых интегральных характеристик процесса. Но еще важнее их методическое значение. Они позволяют не только достаточно наглядно изучить математические особенности задач, возникающих в теории иерархических систем, но и оценить основные идеи, которые могут быть использованы для построения численных методов.

Однако содержание динамических задач неизмеримо богаче, чем статических. Здесь возникает целый ряд новых проблем, не имеющих аналогов в статике. В самом деле, если в статических задачах управление — это некоторый единичный акт принятия решения, то в динамических задачах рассматривается некоторая цепочка (конечная или бесконечная) последовательных актов принятия решений. Следовательно, у Центра возникают совершенно новые, значительно более эффективные возможности влиять на выбор стратегий производителя. Наконец, повторяемость актов управления позволяет в какой-то степени компенсировать недостаток информации у Центра использованием идей и методов адаптивных систем. В данном докладе, разумеется, мы ограничиваемся изложением лишь самых начальных фактов этой теории.

2. *Децентрализация в централизованной экономике.* Подчеркнем еще раз, что примеры, которые мы рассматриваем, относятся к проблеме децентрализации принятия решений в централизованной экономике. Наличие общих глобальных целей как раз и является основным признаком такой системы. Разворачивая некоторую иерархическую систему подчиненности, мы заменяем строгую регламентацию некоторыми экономическими механизмами, способными более точно реагировать на внешние изменения в экономической конъюнктуре, техническом прогрессе, изменения социальных факторов и т. п. Централизация, которая проявляется в возможностях Центра формировать или влиять на формирование целевых функций производителя, никак не исключает существование и возможности учета флюктуации рыночного спроса.

Пример, который рассматривается в этом параграфе, носит так же, как и примеры § 1, методический характер. Однако он уже гораздо ближе к реальности и включает в себя целый ряд особенностей, которые свойственны функционированию крупных производственных объединений.

3. *Основные уравнения.* Сохраним обозначения § 2: $x_i(t)$ — это фонды производителя i , которые он имеет в момент времени t ; через u_i и v_i мы обозначим поток инвестиций, соответственно, Центра и производителя, т. е. интенсивность вложения капитала (в единицу времени). Тогда эволюцию фондов мы будем писать в следующем виде:

$$\dot{x}_i = u_i - v_i - k_i x_i, \quad (3.1)$$

где k — коэффициент амортизации. Таким образом, лаги в данной модели не рассматриваются.

Обозначим через P_i продукт, выпускаемый производителем i в единицу времени, через \hat{F}_i — соответствующую производственную функцию:

$$P_i = \hat{F}_i(x_i, l_i),$$

где

$$l_i = l_i L_i.$$

L_i — по-прежнему обозначает количество рабочей силы, l_i — среднюю ставку заработной платы.

Если l_i — величина фиксированная, то мы будем иметь обычную производственную функцию

$$\bar{F}(x, l) = F(x, L) dx^* L^{1-\alpha}, \quad (3.2)$$

где ставка заработной платы включена в константу d , и тогда L будет, как и в § 1, некоторой управляющей функцией. Но представляет интерес и другой случай, когда количество рабочей силы нормально. Оно определяется однозначно объемом фондов $L = L(x)$. Тогда ставка заработной платы превращается в одно из управляющих воздействий и производственная функция будет иметь вид

$$\hat{F}(x, \dots) = F^*(x, \dots). \quad (3.3)$$

Представление производственной функции в форме (3.3) учитывает зависимость производительности труда от ставки заработной платы. Эта зависимость может определяться явлениями весьма разной природы. Прежде всего здесь могут учитываться факторы социологические — влияние фонда поощрения. Но могут быть учтены и совсем другие обстоятельства — может быть нанят на работу более квалифицированный персонал.

Обозначим через c_i цену продукта, тогда «рыночная стоимость» продукта выпускаемого в единицу времени будет $c_i P_i$ — для простоты будем считать продукт однородным.

Из этой суммы должны быть выделены инвестиции v_i , средства на оплату рабочей силы $\dots L_i$, выплаты Центру $\dots (P_i)$. Таким образом, следующее балансовое соотношение будет таким:

$$c_i P_i - v_i - \dots - \dots (P_i) = \dots. \quad (3.4)$$

Функцию $\dots (t)$ условимся называть «доходом производителя», это та сумма, которая остается в распоряжении производителя и может быть им израсходована на поощрения рабочих, на социальные нужды, строительство и т. д.

Примечания. 1. Соотношения (3.1)—(3.4) полностью описывают условия, в которых функционирует производитель весь цикл его функционирования; они отражают основную формулу политической экономии: деньги (капитал, количество которого x_i описывается соотношением (3.1)) — товар (объем которого P_i описывается соотношением (3.2) и (3.3)) — деньги (соотношение (3.4)).

2. Модель (3.1)—(3.4) незамкнутая — она содержит управления v_i , \dots , L_i , которыми производитель может распоряжаться.

3. В реальности все эти соотношения, разумеется, значительно более сложные. В них, например, отсутствуют амортизационные отчисления, налог на фонды, оборотный капитал и т. д. Некоторые из этих обстоятельств будут учтены в том примере, который будет рассмотрен ниже.

4. *Центр*. Так же, как и в примерах, рассматриваемых в § 1, мы будем считать, что функционирование Центра полностью определяется результатами функционирования производителей. Характеристики функционирования Центра будут, таким образом, некоторыми функционалами, зависящими от управлений производителей. Ограничимся рассмотрением только интегральных функционалов, являющихся обобщением целевых функций, рассмотренных в § 1.

а) Производится дефицитный продукт, полностью потребляемый рынком. В этом случае деятельность Центра будет оцениваться тем выше, чем больше значение функционала

$$J = \int_0^T \sum_{i=1}^N P_i dt. \quad (3.5-1)$$

Выражение (3.5-1) отвечает тому случаю, когда продукт, производимый объединением производителей, один и тот же. В более общем случае многономенклатурного производства аргументом будет некоторая взвешенная функция.

б) Центром производится комплектная поставка оборудования. Суммарное задание четко определено. Оно равно P^* . Тогда Центр обязан так руководить производителями, чтобы отклонение от плана поставок было минимальным. В качестве функционала может выступать, например, такой:

$$J = \max_{t \in (0, T)} \| P^* - P_i \|. \quad (3.5-2)$$

в) Центр является хозрасчетной организацией. В этом случае доход Центра будет складываться из суммы доходов производителей. Здесь возможны различные варианты взаимоотношений Центра и производителя. Мы на них не будем останавливаться, так как они требуют уже детального анализа экономического механизма.

Разумеется, перечисленные типы функционалов не ограничивают всего многообразия целевых функций Центра.

Для достижения своих целей Центр имеет в своем распоряжении определенные возможности. Мы будем рассматривать три типа управляющих воздействий.

а) Экзогенный ресурс, который расходуется на централизованные капиталовложения u_i . Они подчинены определенным ограничениям

$$(u_1, \dots, u_N) \in G_u, \quad (3.6)$$

например:

$$\int_{t=0}^T u_i(t) U(t) dt \leq R, \quad (3.7-1)$$

или

$$\int_{t=0}^T u_i(t) dt \leq R. \quad (3.7-2)$$

Экономический смысл ограничений (3.7) очевиден.

б) Ограничения на активность производителей. В распоряжении производителя имеется много видов управления, например, фонд заработной платы Q_i . На эту величину Центром могут быть наложены ограничения вида

$$Q_i \leq R_i, \quad t \in (0, T)$$

или

$$\int_0^T Q_i(t) dt \leq Q_i^*.$$

Могут быть сформулированы ограничения и на объем инвестиций и т. д.

в) Штрафы и поощрения — функции $\Gamma_i(P_i)$, зависящие от функционирования производителей.

5. *Гипотезы поведения.* Для того чтобы данную ситуацию вложить в рамки математической задачи, необходимо постулировать гипотезы Γ и Γ_i — правила игры. Так же, как и в предыдущем параграфе, мы будем считать, что Центр сообщает всю ту информацию, которая необходима производителю для принятия решения: ему известны функции $u_i(t)$, $\Gamma_i(P_i)$ и структура ограничения. В этом случае производитель может так распорядиться своим ресурсом, чтобы максимизировать свой доход. Одна из возможных гипотез Γ_i будет такой

$$\max_0^T (c_i P_i - v_i \Gamma_i(P_i)) dt. \quad (3.8)$$

Однако сроки планирования Центра и производителя могут быть разными. Производитель может быть «менее дальновиден», нежели Центр. Он, например, может не знать дальних целей Центра. Поэтому один из вариантов гипотезы Γ_i может быть и таким:

$$\max_0^{T_i} (c_i P_i - v_i - L_i(P_i)) dt, \quad (3.9)$$

где $T_i \leq T$. Обратный знак равенства на практике, наверное, не встречается, поскольку на новый плановый период Центр может назначить новые правила игры, например, изменить структуру штрафов и поощрений.

Задача производителя, таким образом, состоит в отыскании таких v_i , L_i и P_i , которые доставляют максимум функционалу (3.8). Гипотеза поведения производителя сводит его задачу планирования к некоторой специальной задаче оптимального управления. В результате ее решения стратегии производителей будут некоторыми операторами, действующими из пространства стратегий Центра в пространство стратегий производителей.

Центр должен так распорядиться своими управлениями, чтобы доставить максимальные значения функционалу вида (3.5). Эта задача уже не является стандартной задачей оптимального управления.

В отличие от статических задач, даже в самом простом случае мы приходим к новым сложным типам оптимизационных задач, аналитическое и качественное использование которых очень трудно. В следующем параграфе мы сделаем попытку наметить некоторые возможные схемы численного анализа.

По существу, задачи планирования, о которых идет речь, — это новые типы дифференциальных игр.

6. *Цена децентрализации, роль случайных факторов.* Говоря о статических задачах, мы старались обратить внимание читателя на то, что иерархическая структура управления необходима лишь в том случае, когда нижние звенья могут использовать информацию, которая по тем или другим обстоятельствам недоступна Центру. Если считать задачу полностью детерминированной и лишенной всех неопределенностей, то цена децентрализации, как и в статических задачах, будет положительной.

Для того чтобы приблизить рассматриваемую модель к реальности, надо учесть разнообразные случайные факторы. Перечислим некоторые из них.

а) В действительности увеличение фондов не определяется только количеством вкладываемого капитала. Колебание стоимости оборудования на рынке, стоимости зарплаты, разнообразные технические новшества или, наоборот, отступления от принятой технологии требуют писать уравнение (3.1) в виде

$$\dot{x}_i = kx_i - f_i(u_i, v_i, i^{(1)}), \quad (3.10)$$

где $i^{(1)}$ — некоторый фактор, неизвестный Центру.

б) Особенности производственного процесса у производителей никогда точно не могут быть известны Центру. Поэтому производственную функцию следует писать в виде

$$P_i = \hat{F}_i(x_i, i, i^{(2)}), \quad (3.11)$$

где $i^{(2)}$ — также некоторый фактор, неизвестный Центру.

в) Наконец, функция цели производителя также может содержать целый ряд неопределенностей. Прежде всего цена на продукт может зависеть от местной рыночной конъюнктуры, не вполне известной Центру. Это значит, что цена имеет вид

$$c_i = \hat{c}_i - i^{(3)}$$

и, следовательно, функция цели производителя будет такой:

$$J_i = \int_0^T (\hat{c}_i - i^{(3)}) P_i - v_i - i \, dt. \quad (3.12)$$

Наконец, сама функция цели может быть известна недостаточно точно.

Примечание. Я хотел бы заметить, что предположения о существовании «рыночного хаоса» никак не противоречат принципу централизации управления. Экономический организм имеет некоторые глобальные цели, достижение которых преследуется Центром. Рыночный спрос, его флюктуация — это неизбежный элемент функционирования общества, невозможность регламентации всех особенностей потребления. Учет этой реальности просто поднимает планирование на более высокий уровень, делает его более научно обоснованным.

§ 4. Общая динамическая задача

1. *Функционирование производителя.* Рассматривается система, состоящая из N производителей, фазовое состояние которых обозначается через x_i . Векторы x_i удовлетворяют следующей системе дифференциальных уравнений:

$$\dot{x}_i = f_i(u_i, p_i, p_i(x_i), p_i), \quad i = 1, \dots, N, \quad (4.1)$$

где $u_i(t)$ — экзогенный ресурс, находящийся в распоряжении Центра. Ограничения, которым должны удовлетворять эти векторы, будем писать в виде

$$(u_1, \dots, u_N) \in G_u; \quad (4.2)$$

p_i — вектор управления производителя —

$$p_i \in G_i; \quad (4.3)$$

$p_i(x_i)$ — функции поощрения, находящиеся в распоряжении Центра —

$$p_1, \dots, p_N \in G; \quad (4.4)$$

$p_i(t)$ — некоторые неизвестные или случайные функции времени (или параметров). Уровни информированности об этих величинах у Центра и производителя разные. На основании сведений, находящихся в распоряжении производителя,

$$p_i = g_i. \quad (4.5)$$

В распоряжении Центра находится другая информация:

$$p_i \in G_i, \quad (4.6)$$

причем

$$g_i \in G_i. \quad (4.7)$$

Практически наиболее интересен тот случай, когда информация производителя просто более точная, т. е.

$$g_i \in G_i. \quad (4.8)$$

Условимся считать раз и навсегда, что производитель информирован о действиях Центра. Это значит, что он может принимать решение, т. е. распоряжаться назначением управляющих функций и параметров, зная объем экзогенного ресурса, который ему выделен, структуру функций поощрения $\gamma_i(\cdot)$ и структуру множества G_i . В этих условиях, согласно гипотезе Γ_i производитель выбирает управление, стремясь доставить максимальные значения функционалу

$$J_i = \int_0^{T_i} F_i(x_i, u_i, \gamma_i(\cdot), \cdot, t) dt. \quad (4.9)$$

Примечание. Еще раз подчеркнем, что гипотеза Γ_i — это гипотеза, которую формулирует Центр, которая необходима Центру для принятия решений. Может оказаться, что производитель не будет ей следовать. Анализ этого факта представляет специальный интерес и далеко выходит за рамки данного доклада.

Задача максимизации функционала (4.9) при условиях (4.1), (4.2), (4.3), (4.4) и (4.5) является задачей оптимального управления. Однако она, как правило, выходит за рамки задачи Понтрягина. В самом деле, F_i как функция γ_i может быть некоторым функционалом. Операция F_i над γ_i может быть, например, операцией математического ожидания или операцией минимизации, если гипотеза Γ_i — это получение гарантированного результата. В этом случае

$$F_i = \min_{\gamma_i} \hat{F}(x_i, u_i, \gamma_i(\cdot), \cdot, t). \quad (4.10)$$

Кроме того, ограничения (4.3), если говорить об экономических задачах, являются, как правило, фазовыми, т. е. структура множеств G_i зависит от фазовых переменных. Например, в задаче предыдущего параграфа инвестиции v удовлетворяют, очевидно, ограничению

$$0 \leq v \leq c_i P_i - \dot{x}_i - \dot{y}_i - \dot{z}_i(P_i).$$

2. *Поведение Центра.* В распоряжении Центра имеются следующие возможности управлять производителем:

- а) распределять экзогенный ресурс u_i ;
- б) назначать функции поощрения $v_i(x_i)$;
- в) определять структуру ограничений G_i .

Последнее из управлений является спецификой иерархических систем и требует некоторых пояснений. В § 1 мы, например, видели, что ограничения G включают условие ограниченности фонда заработной платы. В терминах настоящего параграфа это условие будет выглядеть так:

$$v_i^{(j)} - v_i^{-j},$$

где $v_i^{(j)}$ означает компоненту вектора v_i , а v_i^{-j} — некоторое число, которое находится в распоряжении Центра. Ограничения могут быть и другого типа, например, соблюдение технологических пропорций и т. д. Ограничимся рассмотрением ограничений вида

$$V_i(x_i) - V_i(t), \tag{4.11}$$

где $V_i(x_i)$ — некоторая функция компонент вектора x_i , а $V_i(t)$ — функция, которая находится в распоряжении Центра.

Целевую функцию Центра мы будем писать в виде

$$J = \int_0^T F(x_1, \dots, x_N, u_1, \dots, u_N, v_1, \dots, v_N, t) dt \tag{4.12}$$

Поскольку производитель будет информирован о действиях Центра, то величины x_i и v_i будут некоторыми операторами от $u_i(t)$, v_i и x_i , т. е.

$$J = \int_0^T F^*(u_1, \dots, u_N, v_1, \dots, v_N, t) dt. \tag{4.13}$$

Однако выражение (4.13) содержит еще случайные (или неопределенные) факторы v_i , и для того чтобы описать действия Центра, должно быть регламентировано правило обращения с этими параметрами (гипотеза Г). В результате мы придем к функционалу, который будем писать условно в следующем виде:

$$J^* = \min_{i \in \Gamma} \int_0^T F^*(u_1, \dots, u_N, v_1, \dots, v_N, t) dt. \tag{4.14}$$

Возможные примеры гипотез Γ были приведены в предыдущем параграфе.

3. *Задача распределения экзогенного ресурса в условиях полной информированности.* Простейшей задачей рассматриваемого плана является, вероятно, именно эта задача. В ней $x_i \geq 0$, $u_i \geq 0$. В этом случае задача производителя сводится к определению вектора $x_i(t)$, доставляющего максимум функционалу

$$J_i = \int_0^{T_i} F_i(x_i, u_i, t) dt \quad (4.15)$$

при ограничениях

$$\dot{x}_i = f_i(x_i, u_i, t), \quad (4.16)$$

$$x_i \in G_i. \quad (4.17)$$

Предположим, что G_i не зависит от фазовых переменных, а тогда для ее решения может быть использован принцип максимума. Значит, переменные x_i и u_i удовлетворяют следующей краевой задаче:

$$\dot{x}_i = f_i(x_i, u_i, t), \quad (4.18)$$

$$\dot{F}_i = -\frac{\partial f_i}{\partial x_i} F_i - \frac{\partial F_i}{\partial x_i}, \quad (4.19)$$

$$H(x_i, u_i, F_i) = \max, \quad (4.20)$$

$$x_i(t_0) = x_{i0}, \quad x_i(T_i) = 0. \quad (4.21)$$

Последнее из условий (4.21) — следствие того, что при $t = T_i$ на величины x_i не накладывается никаких ограничений. Расчет оптимальной траектории в этой задаче достаточно сложен, даже если функции u_i фиксированы. Но для того чтобы решить задачу Центра, нам необходимо иметь зависимость фазовой траектории и управления производителя от структуры функций $u_i(t)$. Поэтому построение рациональных численных схем в такой задаче представляется не только важной, но и весьма сложной проблемой. Задача значительно упрощается, если множество G_i совпадает со всем пространством. В этом случае необходимое условие максимума J_i имеет вид:

$$\frac{H}{i} = f_i, \quad \frac{f_i}{i} = \frac{F_i}{i} \quad (x_i, u_i, t) = 0, \quad i = 1, \dots, N, \quad (4.22)$$

и задача Центра превращается в задачу отыскания экстремума функционала при дополнительных ограничениях типа равенств. Эта задача значительно более обозрима. Конечно, ее размерность весьма вы-

сока и равна $\sum_{i=1}^N 2n_i$, где n_i — размерность вектора x_i , но нетрудно по-

строить итерационные схемы, каждый шаг которых требует решения N задач, размерность каждой из которых равна $2n_i$.

Рассуждения, которые здесь были приведены, имеют своей целью наметить некоторые рекомендации для численного анализа рассматриваемых механизмов.

Итак, прежде всего следует стремиться снять ограничения. Если задача существенно нелинейна, то для этой цели следует использовать функцию штрафа. Если управление u_i входит линейно, как это было в рассмотренном примере, то для снятия ограничений более целесообразно вводить новые управления, например, если u_i скаляр, а ограничения имеют вид

$$-u_i(t, x) \leq u_i(t, x),$$

то вместо u_i можно ввести функцию v_i :

$$u_i = \frac{1}{2} (u_i + u_i) \sin v_i.$$

Далее, особая роль в теории иерархических систем принадлежит различным вариантам теории возмущений, причем, вероятно, наиболее эффективной теорией такого рода будет такая, в которой произведена линеаризация по фазовым переменным и удержаны члены второго порядка малости по отношению к поправкам на управление. Другими словами, целесообразно использовать процедуры теории возмущений, являющиеся естественным обобщением тех, которые были изложены в § 2.

Заметим, что анализ, произведенный с помощью подобной теории возмущений, даст результат, аналогичной тому, которой мы

имели в § 2. Если \bar{u}_i является оптимальным решением для производителя при $u_i = \tilde{u}_i$, то

$$u_i = o(\|u_i\|^2),$$

т. е. управление с помощью экзогенного ресурса не является эффективным в окрестности оптимального решения производителя.

§ 5. Пример трехступенчатой иерархической системы

1. *Предварительные замечания.* Модели экономических механизмов, которые мы рассматривали в этом докладе, были предельно упрощены. Вероятно, только в самых простых случаях можно свести анализ к модели, имеющей два уровня. В самом деле, пусть Центр — это некоторые производственные объединения, а производители — хозяйственные предприятия. Мы рассмотрели функционирование этого множества экономических организмов как некоторой независимой совокупности, изолированной от всего остального мира. Но ведь существует еще государство, которое имеет юридическое и фактическое право вмешиваться в деятельность объединений, ставить им определенные задачи, влиять тем или иным образом на стратегию производителей. Частично мы учли эти возможности, рассматривая экзогенный ресурс — капитал, выделяемый государством в распоряжение центров. Мы назвали его экзогенным ресурсом. По существу он имитировал централизованные капиталовложения.

Но действительность значительно сложнее. Прежде всего этот экзогенный ресурс является капиталом, формируемым в результате действия тех же самых производителей. Его формирование — это своеобразный и сложный механизм, включающий в себя налоги на фонды, налог на оборот и другие виды отчислений. Далее, распределение экзогенного ресурса по объединениям — это еще один механизм общегосударственного планирования. Наконец, еще существует банк, деятельность которого пронизывает все звенья иерархической системы. Поэтому даже самая простая модель общегосударственного механизма должна иметь три уровня иерархии. Однако переход к трехуровневой системе качественно усложняет математическое содержание задачи. Для того чтобы представить себе содержание задачи анализа подобного механизма, рассмотрим условную модель трехступенчатой иерархической системы, которую назовем государство — центры — производители.

2. *Модель «Программа».* Предположим, что в данной экономической системе, которую мы условно назовем государством и которая предполагается замкнутой*, вырабатывается M разных продуктов P_1, P_2, \dots, P_M . Для выполнения Программы необходимо, чтобы эти продукты вырабатывались в течение планового периода в определенных количествах. Другими словами, в результате разработки Программы определены функции $P_j^*(t)$. Следовательно, реализация Программы в каждый данный момент характеризуется системой показателей

$$r_j(t) = \frac{P_j(t)}{P_j^*(t)}.$$

Поскольку $P_j^*(t)$ — некоторый абсолютно необходимый норматив, то оценка реализации программы выпуска продукта номера j в течение планового периода можно характеризовать числом

$$\bar{r}_j = \min_{t \in (0, T)} r_j(t).$$

Тогда очевидно, что государство тем лучше выполнит Программу, чем больше будет функционал

$$J = \min_j \bar{r}_j = \min_j \min_t \frac{P_j(t)}{P_j^*(t)}. \quad (5.1)$$

3. *Центры.* Предположим, что весь экономический организм государства состоит из N объединений, которые мы будем называть центрами. Каждое объединение i состоит из N_i производителей. Центры выполняют только управленческие (бюрократические) функции. Они не производят продукт непосредственно. Задача центров — распределить капиталовложения, которые выдаются государством. Этот ресурс U_i выдается государством безвозмездно — он моделирует централизованные капиталовложения. Кроме того, государство выделяет в распоряжение центра фонд заработной платы $Q_i(t)$, который центр должен распределить среди производителей, входящих в данное объединение. Кроме того, центру сообщается его программа $P_j^{i*}(t)$ ($i = 1, \dots, N, j = 1, \dots, M$) производства.

* Включение в эту модель экспорта-импорта и конъюнктуры международного рынка не вносит с математической точки зрения качественных усложнений. Все особенности трехуровневой иерархии могут быть прослежены и на замкнутой модели.

Примечание. Если центр i производит только однотипную продукцию, например, P_i , то все $P_j^i = 0$ для $j \neq i$. Как и всякое бюрократическое управленческое учреждение он заинтересован в наилучшем выполнении Программы, которое обеспечивает его стабильность. Функционал, характеризующий благополучие центра, будем обозначать через J_i . Одна из возможностей описать действие центра — это максимизация функционала

$$J_i = \min_j \min_t \frac{P_j^i(t)}{P_j^{i*}(t)}. \quad (5.2)$$

Существуют также полностью хозрасчетные объединения, которые имеют целевые функции, отличные от (5.2). Их обсуждение на содержательном уровне нас увело бы очень далеко от основной задачи доклада — представить себе особенности математических гипотез, лежащих в основе функционирования подобного трехступенчатого механизма. Для того чтобы охватить более широкий класс задач, условимся считать, что цели центра описываются некоторым функционалом J_i , зависящим только от результатов производственной деятельности производителей, входящих в объединение

$$J_i = J_i(P_1^i, P_2^i, \dots, P_M^i). \quad (5.3)$$

4. *Банк.* В распоряжении государства находится Банк, который является носителем его интересов. Этим условным термином объединяем как собственно банковую службу, так и фискальную налоговую систему.

Через $z(t)$ мы будем обозначать общий объем денег, находящийся в Банке, т. е. в распоряжении государства.

В качестве поступлений в Банк мы будем рассматривать следующие составляющие:

а) Налог на фонды. Фонды производителя k , входящего в объединение i (производитель (i, k)), будем обозначать через x_{ik} . Обозначим через α_x некоторый числовой коэффициент и будем считать, что производитель отчисляет из своей прибыли величину $\alpha_x x_{ik}$:

$$z_x = \sum_{i=1}^N \sum_{k=1}^{N_i} \alpha_x x_{ik}. \quad (5.4)$$

б) Налог с оборота. Предположим, что производитель (i, k) создал в единицу времени продукт $\{P_1^{ik}, P_2^{ik}, \dots, P_M^{ik}\}$. Обозначим через

c_j рыночную (или оптовую государственную) цену продукта P_j . Тогда государству отчисляется некоторая сумма c_j^* , т. е. производитель получает за продажу продукта

$$\sum_{j=1}^M (c_j - c_j^*) P_j^{ik}, \quad (5.5)$$

а государству отчисляется от всех производителей

$$z_p = \sum_{i=1}^N \sum_{k=1}^M c_j^* P_j^{ik} - \sum_{j=1}^M c_j^* P_j, \quad (5.6)$$

где P_j — суммарный j -й продукт.

в) Процент от кредита центрам. Банк может кредитовать центры. Если y_i — количество денег, которые центр i взял в займы у Банка, то он выплачивает Банку в единицу времени $y_i y_i$. Таким образом, Банк в форме процента на кредит получает в течение единицы времени

$$z_y = \sum_{i=1}^N y_i y_i. \quad (5.7)$$

г) Процент от кредита производителям. В этом докладе речь идет о централизованной экономике, носителем целей которой мы условились называть государство. Нетождественность интересов всех уровней иерархии может привести к тому, что государству выгодно иметь возможность тем или иным способом непосредственно, минуя центры, влиять на активность производителей. Обозначим через y_{ik} количество денег, полученных производителем (i, k) в кредит от Банка, через y_{ik} — процент, тогда Банк в течение единицы времени получает

$$z_{yy} = \sum_{i=1}^N \sum_{k=1}^M y_{ik} y_{ik}. \quad (5.8)$$

Коэффициенты y_i и y_{ik} мы будем называть учетной ставкой Банка — это некоторые параметры, находящиеся в распоряжении государства.

Общий доход государства расходуется следующим образом:

а) Непроизводственные расходы $R(t)$, которые являются заранее фиксированными функциями времени.

б) Централизованные капиталовложения R_u . Они складывается из суммы U_i средств, безвозмездно выделенных государством центрам

$$R_u = \sum_{i=1}^N U_i. \quad (5.9)$$

в) Фонда заработной платы Q

$$Q = \sum_{i=1}^N Q_i. \quad (5.10)$$

Данный фонд ограничен объемом продуктов, которые можно приобрести на эти деньги, т. е.

$$Q = \sum_j c_j P_j, \quad (5.11)$$

где суммирование распространено по всем тем индексам j , которые отвечают потребительским продуктам («рыночный» продукт).

г) Кредит центрам R_y . Он складывается из тех средств, которые в единицу времени объединения берут займы у Банка

$$R_y = \sum_{i=1}^N v_i. \quad (5.12)$$

д) Кредит производителям R_{yy}

$$R_{yy} = \sum_{i=1}^N \sum_{k=1}^{N_i} v_{ik}. \quad (5.13)$$

Кредиты берутся центрами и производителями — это их управления, но максимальный объем кредитов лимитируется Банком — это управление государства:

$$0 \leq v_i \leq v_i(t, y_i) \quad (5.14)$$

$$0 \leq v_{ik} \leq v_{ik}(t, y_{ik}, x_{ik}, P_j^{ik}). \quad (5.15)$$

Таким образом, Банк имеет право отказать в кредитах, причем в общем случае эти управления зависят от объема капиталов, уже взятых в займы, и от результатов функционирования.

Кроме ограничений (5.14) и (5.15) на величину текущих займов, должны быть наложены ограничения на общую сумму кредита, находящегося в распоряжении кредитруемого:

$$0 \leq y_i \leq y_i(t), \quad (5.16)$$

$$0 \quad y_{ik} \quad y_{ik}(t). \quad (5.17)$$

Теперь изменение общего капитала, находящегося в распоряжении государства — Банка, описывается следующим скалярным уравнением:

$$\dot{z} \quad z_x \quad z_p \quad z_y \quad z_{yy} \quad R \quad R_u \quad Q \quad R_y \quad R_{yy}, \quad (5.18)$$

где величины, входящие в правую часть уравнения, описываются формулами (5.4)—(5.10), (5.12), (5.13).

Таким образом, задача государства состоит в том, чтобы так распорядиться управляющими воздействиями $U_i(t)$, $Q_i(t)$, c_x^* , y_i , y_{ik} , v_i , v_{ik} , y_i , y_{ik} , чтобы максимизировать функционал (5.1).

5. *Функционирование центра.* Гипотеза информированности предполагает, что государство (Банк) сообщает свои решения центрам и производителям. Таким образом, центр располагает некоторым фондом капитальных затрат $U_i(t)$, который в пределах, допустимых ограничениями (5.14) и (4.16), он может взять из Банка. Этот фонд расходуется на инвестиции производителей и выплату процентов по кредиту

$$U_i \quad v_i \quad \sum_{k=1}^{N_i} u_{ik} \quad y_i y_i \quad t. \quad (5.19)$$

Строго говоря, здесь должен стоять знак неравенства, так как объединение должно иметь некоторый резерв:

$$\sum_{k=1}^{N_i} u_{ik} \quad y_i y_i \quad U_i \quad v_i. \quad (5.19-1)$$

Величина y_i в выражениях (5.19) и (5.19-1) определяется уравнением

$$\dot{y}_i \quad v_i. \quad (5.20)$$

Распределение ресурса и выбор размеров кредитов у Банка — не единственные управляющие воздействия центра. Центр распределяет фонд заработной платы

$$Q_i \quad \sum_{k=1}^{N_i} Q_{ik}^* \quad q_i,$$

где q_i — строго регламентированное число. Оно либо фиксировано (фонд заработной платы объединения), либо является однозначной функцией функционала J_i .

Наконец, центр назначает функция поощрения производителя

$$j_{ik}(P_j^{ik}) \quad (j = 1, \dots, M; i = 1, \dots, N; k = 1, \dots, N_i).$$

Функции поощрения (или штрафа) должны быть подчинены системе ограничений. Достаточно естественным является, например, такое

$$\sum_{j=1}^M \sum_{k=1}^{N_i} j_{ik}(P_j^{ik}) - i = 0. \quad (5.21)$$

Здесь i — некоторая регламентированная часть благ, присваиваемая центром. Она может быть поставлена в зависимость от действий производителей, например, так:

$$i = \sum_{j=1}^M \sum_{k=1}^{N_i} j_i P_j^{ik}. \quad (5.22)$$

Выражения (5.22) являются некоторым «внутренним» аналогом налога с оборота. Равенства (5.21) означают права центра в известной степени перераспределить доходы своих производителей в зависимости от того, насколько их активность отвечает целям центра. Деятельность центра состоит в таком выборе своих управляющих воздействий, который максимизирует функционал (5.2).

6. *Производитель.* Нижнее звено этой иерархической системы занимает производитель. Каждый производитель (i, k) обладает определенным капиталом (фондами) x_{ik} , необходимым для производства продукта P_j^{ik} . Эволюция капитала x_{ik} описывается уравнением

$$\dot{x}_{ik} = \kappa_{ik} x_{ik} - u_{ik}, \quad (5.23)$$

где u_{ik} — нецентрализованные инвестиции. Производство мы будем описывать производственными функциями F_j^{ik} вида

$$P_j^{ik} = F_j^{ik}(x_{ik}^j, L_{ik}^j). \quad (5.24)$$

Здесь x_{ik}^j — доля капитала x_{ik} , неиспользуемая для производства товара номера j :

$$\sum_{j=1}^M x_{ik}^j x_{ik}. \quad (5.25)$$

Стоимость продукции, реализованной производителем, будет

$$\sum_{j=1}^M c_j P_j^{ik}.$$

Общее количество денег (поток денег), которое получает производитель в единицу времени, складывается из стоимости реализованной продукции и кредита, получаемого из Банка, v_{ij} . Это количество денег расходуется следующим образом:

а) Выплачивается государству налог с оборота $\sum_{j=1}^M c_j^* P_j^{ik}$.

б) Выплачивается налог на фонды $\sum_{j=1}^M x_{ik}^j$.

в) Выплачивается заработная плата $\sum_{j=1}^N Q_{ik}^j L_{ik}^j$,

причем

$$Q_{ik} = Q_{ik}^*.$$

г) Производятся внутренние инвестиции v_{ik} .

д) Выплачивается центру штраф (или получается поощрение) $\sum_{j=1}^M (P_j^{ik})$.

Разность

$$\sum_{j=1}^M (c_j - c_j^*) P_j^{ik} - v_{ik} - \sum_{j=1}^M x_{ik}^j - \sum_{j=1}^M Q_{ik}^j L_{ik}^j - \sum_{j=1}^M (P_j^{ik}) \quad (5.26)$$

представляет собой чистый доход производителя, который он может расходовать по своему усмотрению. Условимся называть его социальным фондом предприятия.

Цель производителя (i, k) максимизировать интеграл

$$J_{ik} = \int_0^T v_{ik}(t) dt. \quad (5.27)$$

Для того чтобы производитель мог принять решение, мы должны принять определенную гипотезу о его информированности. Мы будем считать, что до принятия решения производителю сообщены и структура функций поощрения, и размер централизованных инвестиций.

7. *Заключение.* Для того чтобы данная трехступенчатая иерархическая система могла функционировать, должны быть сформулированы определенные «правила игры». Прежде всего должен быть определен порядок ходов. Первый ход делает государство — оно сообщает центрам и производителям значения своих управляющих параметров и функций. После этого задача сведена к анализу двухступенчатой иерархии. В этой системе следующий ход делают центры — они сообщают производителям значения своих управляющих воздействий. Теперь производитель должен сделать свой ход — принять свои решения, т. е. выбрать значения своих управляющих функций и параметров. Он их не сообщает центрам. Центры принимают решения, исходя из гипотез Γ_{ik} о поведении производителя номера (i, k) . Гипотеза Γ_{ik} состоит в том, что производитель максимизирует функционал (5.27). Его решения — управляющие воздействия производителей — будут функционалами от управляющих воздействий центров и государства. Теперь очередь за центрами. Они принимают решения и не сообщают их государству. Последнее принимает решение, основываясь на гипотезах Γ_{ik} и гипотезах Γ_i о поведении центров. Гипотезы Γ_i состоят в максимизации целевых функций центров. Этого достаточно, чтобы принять решение на верхнем уровне — уровне государства, поскольку управляющие функции центров будут некоторыми функционалами от управляющих воздействий государства. Будем называть план распределения ресурсов государства и стратегию использования бюджета оптимальными, если они доставляют максимальное значение функционалу (5.1) при гипотезах Γ_i и Γ_{ik} . Аналогично определяются оптимальные планы центров и производителей.

Многое из того, что говорилось в предыдущем параграфе о численной реализации оптимальных планов в двухступенчатой задаче, переносится и на случай трехступенчатой иерархической системы.

§ 6. Комментарии

1. *О мере информации.* В этом докладе все время использовались слова: информация, информированность, ценность информации и т. д. Понятие ценности информации, по мнению автора, неотделимо от понятия цели, целевой функции. В предисловии мы уже подчеркнули этот факт. При достижении одной цели информация о

некотором фонде или процессе может оказать решающее влияние на выбор стратегии и, следовательно, на значение целевой функции. Та же информация может быть абсолютно бесполезной в случае, если цель изменится. Таким образом, имеет смысл говорить о мере информации или о ее качестве лишь по отношению к данной целевой функции.

Предположим, что целевая функция зависит от некоторого процесса (t) , и мы знаем, например, что

$$a(t) = b(t) \quad t \in [0, T]. \quad (6.1)$$

Информация (6.1) позволяет рассмотреть две задачи:

- а) найти $\max J(\cdot) = J$ при условии (6.1);
- б) найти $\min J(\cdot) = J$ при том же условии.

Разность

$$J - J \quad (6.2)$$

и означает ценность информации. Таким образом, в качестве меры информации мы можем принять меру неопределенности при вычислении целевой функции.

Оценка (6.2) соответствует принципу гарантированного результата. Если принята другая гипотеза поведения, например, введена некоторая функция риска, то вместо оценки (6.2) надо будет принять другую аналогично построенную оценку. Другими словами, ценность информации будет зависеть также и от гипотезы поведения. Однако поскольку гипотеза поведения сводит игровую задачу к оптимизационной, то об этом можно было и не говорить, понимая под функционалом J функционал, к которому мы пришли после того, как использовали гипотезу поведения.

2. *Агрегирование.* Одна из центральных проблем исследования операций — это проблема агрегирования. В разной степени ее решение необходимо в экономическом планировании, военном деле, социологии и т. д. В настоящее время опубликовано много различных работ, посвященных трактовке этого понятия. И все (или почти все) работы, посвященные этой теме, имеют одно общее. Проблема агрегирования рассматривается как некоторый специальный

метод решения оптимизационной задачи. Классическим примером подобного подхода является метод декомпозиции Данцига—Вульфа в линейном программировании.

Мне кажется, что решение проблемы агрегирования должна лежать совсем в другой плоскости. Необходимо четко постулировать следующий факт: в процессе агрегирования происходит не только сжатие, но и (неизбежно) потеря информации. По агрегированной информации нельзя восстановить ее дезагрегированный источник, так же как имея решение уравнений Навье—Стокса, нельзя восстановить судьбы молекул, слагающих найденное течение жидкости.

Но если процесс агрегирования неизбежно приводит к потере информации, то она, согласно сказанному, приводит к неопределенности в значении целевой функции. По-видимому, можно говорить о допустимом уровне агрегирования, если неопределенность, которая порождается разностью $J - J^*$, находится в допустимых пределах.

Есть еще несколько дефектов этой проблемы. Во-первых, разные уровни иерархической системы имеют информацию разного качества об исходном дезагрегированном массиве и, во-вторых, необходимо еще учитывать стоимость обработки массивов разной степени агрегирования. Все это в целом, по-видимому, может служить основой для построения такой теории агрегирования, в которой степень агрегирования определяется соответствующим значением целевой функции, и может быть поставлена математическая задача оптимального агрегирования.

3. *Анализ и синтез иерархических систем.* В представленном докладе мы занимались проблемой анализа: изучали особенности данной, уже спроектированной системы, с заданными уровнями иерархии, с четко регламентированными правилами функционирования.

Однако, как и для любых систем управления, в теории иерархических систем следует выделить проблемы синтеза. Здесь два типа задач, которые можно назвать так: а) задачи текущего планирования (управления); б) задачи структурного синтеза.

Они значительно отличаются друг от друга, и о них следует говорить независимо.

а) Под задачей синтеза в теории управления понимают построение такой управляющей функции, значения которой однозначно определяются фазовыми координатами управляемой системы. Мы эту задачу назвали задачей текущего управления. По существу один

из вариантов ее постановки приведен в докладе. Функция поощрения как раз и есть синтезирующее управление. Однако приведенная постановка задачи не является исчерпывающей. Можно привести много примеров из реальной жизни, в которых четкая регламентация функции поощрения отсутствует. В зависимости от некоторых внешних причин происходит изменение стратегии центра. Другими словами, меняются правила игры. Это обстоятельство порождает целый класс новых задач, которые естественно рассматривать под углом зрения центра. Можно ли говорить о некоторых оптимальных правилах «изменений правил» игры, имея в виду неизбежную реакцию нижних звеньев, возможности блефа и т. д. Интересно проследить на примерах, к чему приводит отсутствие всяких правил, т. е. анализ некоторой инфраигры — игры без правил.

б) В задачах текущего управления мы синтезировали функцию управления при заданной структуре. Однако естественно поставить вопрос о выборе структуры, о возможности существования оптимальной иерархической структуры, т. е. такой вопрос, о котором мы говорили в предисловии. Множество возможных структур — это всегда некоторое дискретное множество. Следовательно, задача выбора оптимальной структуры — это задача оптимизации на некотором дискретном множестве. Она, конечно, качественно отличается от задачи а). В целом проблема проектирования оптимальной иерархической системы сводится к совместному решению проблем дискретной и непрерывной информации. В общем случае эти проблемы очень трудно разорвать, поскольку с изменением структуры меняется система целевых функций и систем внутренних связей, а следовательно, и гипотез поведения. Но последние, как мы видели, являются основой редукции проблемы к оптимизационной, т. е. основой для выбора оптимальных программ (задача оптимального планирования) и синтеза управляющих воздействий (функций поощрения).

Построение некоторой общей схемы проектирования представляется, конечно, очень важной и трудной задачей большого прикладного и теоретического значения. Но на сегодняшнем этапе, я думаю, еще более важно дать законченный анализ некоторых конкретных систем, который позволит выделить то общее, что объединяет иерархические системы разной природы.

4. *Целевые функции звеньев.* Это самостоятельная проблема и она имеет много аспектов. В некоторых случаях, например, если

речь идет об экономических системах, объединяющих хозрасчетные звенья, и если мы имеем к тому же возможность влиять на их формирование с помощью штрафов и поощрений, эта проблема решается относительно просто.

Гораздо сложнее проблема целевых функций бюрократических, управленческих организаций типа центров. Для выяснения этой проблемы необходим глубокий социологический анализ. Пока что о ней можно говорить только предположительно.

По-видимому, такого типа организации несут много черт, характерных для организма. Непосредственно объем дохода, даже у хозрасчетных объединений или корпораций, еще не представляет собой основную цель функционирования управленческого аппарата. По-видимому, в этих ситуациях надо следовать Эшби и считать, что управленческий аппарат стремится главным образом сохранить свою устойчивость. Но если область гомеостазиса достаточно широка, то в действиях центров будет достаточно много произвола. Следовательно, если верна наша гипотеза, то для управления управленческим аппаратом необходима некоторая регламентация правил неформального характера, например, роспуска аппарата или его руководства в случае, если центр (объединение, корпорация) не выполнит определенных показателей. Данная точка зрения открывает реальные возможности для формализованного описания поведения центра.

Если речь идет о технических системах, в которых отдельные звенья не имеют никаких целевых функций, но способны по тем или иным алгоритмам выбирать управляющее воздействие, то проблемы ставятся в совершенно иной плоскости. Здесь мы должны сами назначить целевые функции подсистем (или, что то же самое, алгоритмы обработки информации). Проблема назначения целевых функций будет некоторой специальной задачей аппроксимации — аппроксимации целевой функции системы целевыми функциями подсистем.

5. *Заключение.* Соображения, которые изложены в докладе, были попыткой найти радикальный подход к исследованию иерархических систем. Сейчас еще трудно говорить, насколько общи и действительны эти подходы. Необходим прежде всего анализ конкретного материала — анализ функционирования конкретных систем. Попытка уже сегодня строить подробно разработанные вычислительные схемы, обладающие высокой степенью универсальности, — вряд ли

наиболее рациональный путь развития теории. Примеры, которые мы рассмотрели, показывают, что аналитическая природа возникающих задач очень сложна и вряд ли сколь-нибудь универсальные схемы будут эффективны. Мне кажется, что дальнейшее развитие теории должно идти по двум каналам. Во-первых, необходимо развивать качественные методы, позволяющие на относительно простых примерах исследовать разнообразные особенности игрового характера, которые возникают в теории иерархических систем. Во-вторых, необходимо развить технику машинного моделирования, т. е. построения систем, имитирующих функционирование реальных систем, и в том числе процедур принятия решений. Последнее и будет, вероятно, главным направлением практического использования развиваемой теории.

1973 г.

Иерархические структуры и теория игр

В течение нескольких последних лет сотрудниками Вычислительного центра АН СССР была опубликована серия работ, которая позволяет говорить о формировании новой дисциплины — теории иерархических систем управления. В ее основе лежит представление о влиянии процессов обработки информации на качество принимаемых решений. Может оказаться, что централизованный сбор и обработка информации в известных условиях либо вообще невозможны, либо приводят к такому уровню неопределенности, при котором нельзя говорить об обоснованности принимаемых решений. В подобных условиях централизованная система управления оказывается неэффективной. Единственной альтернативой в этом случае становится «распараллеливание» процедур обработки информации и принятия решений, т. е. организация в системе управления иерархической структуры, а значит и определенной степени децентрализации. Анализ возникшей таким способом системы управления приводит к математическим задачам игровой природы.

Изменяя тем или иным образом юридическую основу взаимоотношений между различными звеньями иерархически организованной системы управления, мы изменяем тем самым взаимную информированность сторон, структуру процедур принятия решений и формализованное описание целевых функций подсистем, т. е. изменяем характер игровых задач, возникающих при исследовании иерархических систем управления. Благодаря этому особое значение для развиваемой теории приобретают исследования влияния неформальных факторов на содержание и математические особенности задач теории игр, успешно проводимые последние годы Ю.Б. Гермейером и его учениками.

Предлагаемая концепция иерархически организованных систем управления существенно отличается от тех концепций, которые в своей основе опираются на идеи блочного программирования.

Предметом исследования является не отыскание точного решения тех или иных оптимизационных задач, а исследование процедур функционирования систем и выяснение вопроса о рациональной степени централизации и децентрализации.

Настоящая статья является введением в теорию иерархически организованных систем управления. Первый раздел носит вспомогательный характер. В нем изложена система взглядов, выработавшихся в ВЦ АН СССР относительно места и значения отдельных постановок игровых задач. Следующие разделы посвящены изложению последовательно усложняющейся цепочки задач и призваны иллюстрировать развиваемые идеи и одновременно показать читателю основные трудности, возникающие при анализе.

С рукописью статьи ознакомились И.А. Ватель, Ю.Б. Гермейер, Ф.И. Ерешко, В.С. Михалевич и другие ученые, сделавшие полезные замечания. Особенно тщательно статью прочел Г.С. Поспелов и дал ряд советов, существенно улучшивших изложение. Всем упомянутым лицам автор выражает свою искреннюю благодарность.

1. Замечания о гипотезах в теории игр

1. Прежде чем обсуждать возможные пути построения математической теории управления системами, обладающими иерархической структурой, хотелось бы сделать несколько замечаний о предположениях, которые приводят к задачам теории игр.

Пусть речь идет о принятии решения. Это будет означать, что мы выбираем некоторый элемент $x \in X$, где X — некоторое множество. Далее будем считать, что X — множество пространства Банаха B_x . Выбор $x \in X$ должен быть мотивирован. Это значит, что существует какая-либо цель, достижение которой преследует выбор x . Если эта цель состоит в максимизации некоторой скалярной функции $F(x)$, то проблема принятия решения сводится к отысканию точки, реализующей максимум функции $F(x)$, определенной на $X \in B_x$.

Может оказаться, однако, что выбор любого $x \in X$, т. е. принятие любого решения, не позволяет однозначно определить результат. Так бывает всякий раз, когда результат зависит не только от наших действий. Эту ситуацию будем записывать в форме функции уже двух переменных $F(x, y)$, где $y \in Y \in B_y$.

Параметр y характеризует неопределенность, имеющуюся в операции (задаче). Это могут быть действия реального противника,

отсутствие необходимой информации (и в частности, случайные помехи) или неопределенность наших желаний. На последнее особенно следует обратить внимание, поскольку к этому случаю сводится, в частности, проблема принятия решения, когда существует несколько целей.

Каким образом мотивировать теперь выбор $x \in X$ и превратить проблему принятия решений в условиях неопределенности в строгую математическую задачу?

Для этого необходимо сделать некоторые дополнительные предположения о поведении других лиц, участвующих в операции и преследующих свои цели, или, скажем, о «поведении» природных неопределенностей. Эти предположения назовем гипотезой поведения. В зависимости от характера этих гипотез будем иметь те или другие математические задачи.

Подчеркнем еще раз следующее тривиальное утверждение. Не делая никаких предположений о процедуре выбора элемента y и области его определения, нельзя сформулировать никаких мотивов относительно выбора решения x .

Частным случаем гипотез поведения является предположение о том, что противник выбирает y из соображений

$$\min_{y \in Y} F(x, y) = \max_{y \in Y} (F(x, y)). \quad (1.1)$$

Эта гипотеза поведения приводит к теории антагонистических игр и их анализу в духе фон Неймана [1]. Одной гипотезы поведения еще недостаточно, чтобы полностью описать действия «противника». Необходима некоторая гипотеза информированности участников операции об интересах и выборах стратегий партнеров, а также о той последовательности, в которой осуществляются выборы каждым игроком своих стратегий. Эту гипотезу будем называть гипотезой об информированности и о порядке ходов. Аналогичные представления используются и в классических схемах [1] за исключением информированности об интересах участков операции (партнеров в игре). Например, можно предположить, что мы делаем «первый ход» и «противник» будет его знать, т. е. будет знать наш выбор x . Далее он произведет свой выбор y согласно принятой гипотезе поведения (сделает второй, заключительный, ход в игре).

Из этих гипотез информированности и поведения (1.1) следует знаменитый принцип гарантированного результата. Он сводится к

тому, что противник строит свое решение в виде функции $y = \hat{y}(x)$, которая определяется из условия (1.1). Это означает, что $x \in X$ — функция $\hat{y}(x)$ доставляет минимум $F(x, y)$:

$$\min_{y \in Y} F(x, y) = F(x, \hat{y}(x)) = F^*(x). \quad (1.2)$$

В этом случае выбор решения $x = \hat{x}$ сводится к определению элемента x , доставляющего максимум функции $F^*(x)$. Таким образом,

$$F^*(\hat{x}) = \max_{x \in X} \min_{y \in Y} F(x, y). \quad (1.3)$$

Выбор решения согласно процедуре (1.3) называется принципом гарантированного результата. В этом случае выбор решения свелся снова к проблеме оптимизации. Однако аналитическая природа функций вида $F^*(x)$ обычно бывает очень сложна. Даже если функция $F(x, y)$ многократно дифференцируема, то относительно $F^*(x)$ можно утверждать только дифференцируемость по направлению.

В последнее время к методам отыскания экстремумов функций вида $F^*(x)$ привлечено внимание математиков, и в этой области получено много новых важных результатов.

Меняя гипотезу об информированности и (или) гипотезу поведения, будем получать различные типы формулировок математических задач. Так, например, имеет смысл предположить, что «противник» выбирает y , зная наш отклик на этот выбор, т. е. $x = x^*(y)$. В этом случае противник, если он следует гипотезе поведения (1.1), выбирает такой элемент $y = y^*$, который доставляет минимум функции $F(x^*(y), y)$:

$$\min_{y \in Y} F(x^*(y), y) = F(x^*(y^*), y^*). \quad (1.4)$$

Тогда остается только одно — так выбрать функцию $x^*(y)$, чтобы доставить максимум функционалу (1.4). Это решение обозначим через $x^{**}(y)$:

$$F(x^{**}(y), y) = \max_{x^*(y)} F(x^*(y), y). \quad (1.5)$$

Задача (1.5) уже совсем другой природы, чем (1.3), это вариационная задача. Она сводится к отысканию не элемента \hat{x} , а некоторой функции $x^{**}(y)$, т. е. нашего отклика на ход противника.

Пользуясь терминологией теории управления, получим некоторую задачу синтеза.

2. Изменим теперь класс гипотез поведения. Вместо предположения, что «противник» выбирает свою стратегию $y \in Y$ из условия $\max_x \{ F(x, y) \}$, предположим, что он выбирает y из условия $\max_{y \in Y} (x, y)$, где в общем случае

$$F \quad . \quad (1.6)$$

Такое изменение гипотезы поведения приводит к новому классу игр. Теория, которая здесь возникает, шире теории антагонистических игр. К тому же она дает язык, удобный для описания многочисленных и важных приложений.

Эта теория неантагонистических игр, или игр с непротивоположными интересами.

Отметим, что рассмотренный нами случай (1.6) для игр двух лиц естественным образом вкладывается в наиболее общий класс игр n лиц. А такие игры при $n > 2$ по существу являются играми с непротивоположными, но, вообще говоря, и не совпадающими интересами.

Традиционно в играх многих лиц рассматриваются принципы принятия рациональных решений на основе либо ситуаций равновесия [2], либо коалиционных действий. В то же время наличие в игре права первого хода у одного из игроков при учете его надлежащей информированности существенно расширяет возможности этого игрока. В самом деле, формируя свою стратегию в виде функции от ожидаемой информации и сообщая ее партнерам, игрок, обладающий приоритетом в действиях, может ограничить множество окончательных исходов в игре. Это позволяет ему гарантировать себе результат, как правило, превышающий тот, который он получает в ситуации равновесия, а зачастую и тот, который гарантирован ему в коалиции. Как увидим ниже, характернейший атрибут иерархической системы — приоритет в действиях верхних звеньев по отношению к нижним — позволяет существенно продвинуться в изучении процессов принятия решений в таких системах.

Для того чтобы подчеркнуть отличие рассматриваемой концепции от традиционных схем, будем эти игры называть играми Гермейера [3], который первым начал систематическое изучение влияния роли

неформальных факторов на формирование стратегий и результат игры. Подробная библиография работ по этому направлению дана в [3].

Поскольку интересы противника, который распоряжается элементом y , не противоположны нашим, то и сам термин «противник» мало соответствует существу дела. Поэтому в дальнейшем вместо слов «мы» и «противник» будем говорить игрок x и игрок y . Игрок x контролирует стратегию x , игрок y — стратегию y .

Предположим, что игрок x делает первый ход. Он выбирает x и сам сообщает его игроку y . (Заметим, что в играх Гермейера одному игроку может оказаться выгодным передавать информацию другому игроку.) Тогда игрок y определяет свою стратегию $y = \hat{y}$:

$$(x, \hat{y}) = \max_{y \in Y} F(x, y),$$

\hat{y} — некоторая функция $\hat{y}(x)$.

Теперь игрок x строит свою гарантирующую стратегию из условия достижения наилучшего гарантированного результата:

$$\max_{x \in X} F(x, \hat{y}(x)) = F(\hat{x}, \hat{y}(\hat{x})).$$

Здесь и далее используется предположение об однозначности $\hat{y}(x)$. Точно так же игрок y может сообщать игроку x свой выбор в форме функции $x(y)$. Заметим, что в данном случае эту функцию можно трактовать как функцию штрафа или поощрения.

Игрок x выбирает тогда элемент y^* из условия

$$\max_{y \in Y} F(x(y), y) = F(x(y^*), y^*).$$

Заметим, что элемент y^* будет теперь уже некоторым оператором

$$y^* = y[x(y)].$$

Принятие решения для игрока x сведется тогда к выбору такой функции $x^*(y)$, которая доставляет максимум функционалу

$$F(x(y^*), y^*).$$

3. Рассмотрим еще одну особенность данного класса игр. Пусть речь идет об антагонистических играх

$$(x, y) \in F(x, y), y \in Y,$$

и игрок Γ_1 имеет возможность влиять на структуру ограничений игрока Γ_2 — изменять множество Y . Например, он может заменить Y на $Y_1 \subset Y$. Игроку Γ_1 эта операция выгодна всегда, так как

$$x \in \arg \min_{y \in Y_1} F(x, y) \subset \arg \min_{y \in Y} F(x, y)$$

и, следовательно,

$$\max_x \min_{y \in Y_1} F(x, y) \geq \max_x \min_{y \in Y} F(x, y).$$

В играх с непротивоположными интересами этот факт уже может не иметь места. Если пользоваться языком теории управления, то может оказаться полезным с точки зрения игрока Γ_1 расширить множество допустимых стратегий своего «противника», т. е. игрока Γ_2 .

2. Пример двухступенчатой иерархической системы

1. В последние годы иерархические структуры принятия решений в экономике начали подвергаться всестороннему анализу ([4], [5]). При этом используются различные способы формализации и модельного описания.

С нашей точки зрения, теория неантагонистических игр является наиболее естественным языком для описания иерархических систем управления в экономике. Предлагаемый пример должен продемонстрировать этот факт. Однако глубокая связь, которая существует между теорией игр Гермейера и теорией иерархических систем, не исчерпывается вопросами общности языка. Многие факты, установленные из анализа игровых ситуаций, имеют, как оказывается, глубокий экономический смысл.

2. Рассмотрим промышленное объединение (фирму), состоящее из производственных предприятий, которые выпускают продукты P_1, \dots, P_N . Здесь P_i является вектором, имеющим для каждого предприятия свою размерность. Аналогично вводимые ниже величины являются векторами соответствующих размерностей. Обозначим через x_i основные фонды предприятия номера i . Изменение фондов описывается уравнением

$$\dot{x}_i = k_i x_i - u_i(t) + v_i(t), \quad (2.1)$$

где k_i — матрица коэффициентов амортизации; u_i — инвестиции фирмы (экзогенный ресурс); v_i — внутренние капиталовложения; $u_i v_i$ — это потоки капиталовложений — вложения в течение единицы времени; $v_i(t)$ — управляющая функция предприятия; $u_i(t)$ — управляющая функция фирмы.

Процесс производства будем описывать производственной вектор-функцией:

$$P_i = P_i(x_i, L_i, w_i), \quad i = 1, \dots, N. \quad (2.2)$$

Здесь L_i — количество рабочей силы, w_i — средняя по предприятию ставка заработной платы. Отметим, что запись функции (2.2) является нетрадиционной, а именно: в нее введена зависимость от параметра w_i с целью изучения в дальнейшем возможностей воздействия на предприятия через этот параметр. L_i и w_i также находятся в распоряжении предприятия и подчиняются ограничениям вида

$$\begin{aligned} L_i^* - L_i &\geq 0, \quad w_i^* - w_i \geq 0, \\ L_i &\leq Q_i, \quad i = 1, \dots, N. \end{aligned} \quad (2.3)$$

Смысл первых двух ограничений очевиден. Последнее же означает, что фонд заработной платы ограничен.

Создав вектор продуктов P_i , предприятие реализует его: сдает на склад фирмы или передает в торговую сеть. Обозначим через C_i вектор цен на продукцию i -го предприятия. Тогда за единицу времени предприятие получает сумму (C_i, P_i) .

Из этой суммы нужно заплатить зарплату рабочим $w_i L_i$, сделать вложение в фонд фирмы $w_i(P_i)$, произвести внутренние капиталовложения v_i и компенсировать текущие затраты R_i . Обозначим через f_i остаток этих расходов:

$$f_i(t) = (C_i, P_i) - \{w_i L_i + w_i(P_i) + v_i(t) + R_i(x_i)\}. \quad (2.4)$$

Величину $f_i(t)$ назовем социальным фондом предприятия. Если внешние инвестиции $u_i(t)$ заданы, если известна функция $w_i(P_i)$,

которую естественно назвать функцией штрафа (или поощрения), если известен фонд заработной платы $Q_i(t)$, то задача предприятия номера i — так распорядиться добавочными инвестициями $v_i(t)$, ставкой заработной платы $w_i(t)$ и количеством рабочей силы $L_i(t)$, чтобы максимизировать свой социальный фонд. Этот функционал обозначим через

$$I_i = I_i(v_i, w_i, L_i). \quad (2.5)$$

Так поставленная задача является задачей теории оптимального управления.

Примечание. До сих пор не были уточнены многие ограничения. Для того чтобы сделать задачу корректной, надо, например, потребовать, чтобы

$$t \in [0, T], \quad v_i(t) \geq 0.$$

Функционал I_i может иметь самую произвольную природу. Например, естественно принять, что

$$I_i = \min_{t \in [0, T]} \{ v_i(t) \}. \quad (2.6)$$

Максимизация функционала (2.6) означает максимизацию минимального по времени размера социального фонда. Не менее естественно рассматривать интегральные функционалы вида

$$I_i = \int_0^T v_i(t) dt, \quad (2.7)$$

где T — плановый период.

3. Рассмотрим теперь функционирование правления фирмы, которой подчиняются предприятия. По своему смыслу эта организация сама никаких ценностей не производит. Эффективность ее деятельности оценивается в зависимости от того продукта, который производят предприятия.

Критерий (доход фирмы) будем писать в форме

$$I = I(P_1, \dots, P_N, w_1(P_1), \dots, w_N(P_N)). \quad (2.8)$$

Примечание. В выражении (2.8) подчеркивается, что доход фирмы зависит от структуры функций поощрения (или штрафа) $w_i(P_i)$. Если целевую функцию фирмы (2.8) принять независимой от $w_i(P_i)$, а фонд поощрений — неограниченным, то тривиальным будет следующее утверждение.

Всегда могут быть назначены поощрения (штрафы) $w_i(P_i)$, ..., $w_N(P_N)$ такие, что предприятия будут выбирать свои управления наилучшим способом с точки зрения интересов фирмы.

В самом деле, пусть v_i^* , z_i^* и $L_i^*(i = 1, \dots, N)$ — функции, которые приводят к выпускам $P_i^*(i = 1, \dots, N)$, доставляющим максимум (2.8). Тогда можно подобрать такие $w_i(P_i)$ ($i = 1, \dots, N$), что функционалы $I_i(i = 1, \dots, N)$ будут достигать максимума на v_i^* , z_i^* и $L_i^*(i = 1, \dots, N)$.

В качестве функционала I могут выступать самые разные величины: точность соблюдения государственного плана или максимизация числа комплектов оборудования, или, наконец, чистый доход.

4. Итак, мы пришли к ситуации, о которой говорилось в первом разделе.

Существуют два игрока — фирма и предприятия, имеющие свои собственные цели. Здесь нет симметрии: фирма имеет право первого хода. Задача фирмы так распределить экзогенный ресурс

$$\sum_{i=1}^N u_i = U, \quad (2.9)$$

фонд заработной платы

$$\sum_{i=1}^N Q_i = Q \quad (2.10)$$

и назначить такие поощрения $w_i(P_i)$ предприятиям за выпуск продукции, чтобы максимизировать доход I .

Мы видим, что фирма располагает тремя способами управления — она может распределять экзогенный ресурс U , вводить функции штрафов или поощрений $w_i(P_i)$ и ограничивать активность предприятий.

3. Об одном классе дифференциальных игр

1. Рассмотренный пример является частным случаем следующей динамической системы:

$$\dot{x}_i = f_i(x_i, u_i, v_i, t), \quad i = 1, \dots, N. \quad (3.1)$$

Здесь $x_i(t)$ — вектор-функции времени со значениями из E_{n_i} ; $u_i(t)$ и $v_i(t)$ — вектор-функции со значениями из E_{m_i} и E_{r_i} соответственно.

На функции $x_i(t)$ наложены ограничения

$$t \in T \Rightarrow x_i(t) \in G_i. \quad (3.2)$$

На управляющие функции $u_i(t)$ наложены ограничения вида

$$t \in T \Rightarrow \{u_1, u_2, \dots, u_N\} \in G_u. \quad (3.3)$$

Относительно ограничений на $v_i(t)$ будет сказано ниже.

С динамической системой (3.1) связывается система функционалов \dot{I}_i :

$$\dot{I}_i = \dot{I}_i(x_i, v_i, w_i(x_i, t)), \quad i = 1, 2, \dots, N, \quad (3.4)$$

где $w_i(x_i, t)$ — управления типа синтеза со значениями из E_{k_i} .

На функции v_i , w_i и фазовые переменные x_i накладываются ограничения типа

$$t \in T \Rightarrow \{v_i(x_i, w_i)\} \in G_v. \quad (3.5)$$

Кроме того, вводится понятие генерального функционала \dot{I} :

$$\dot{I} = \dot{I}(x_1, \dots, x_N, u_1, \dots, u_N, w_1, \dots, w_N, G_u). \quad (3.6)$$

Описание игры. В начальный момент $t=0$ состояние системы, т. е. векторы $x_i(0)$, фиксировано. Игрок i делает первый ход. Он сообщает игрокам j функции $w_j(x_j, t)$ и $u_j(t)$ и множество G_u .

Гипотеза поведения игроков i : эти игроки выбирают функции $v_i(t)$ так, чтобы максимизировать функционалы \dot{I}_i . Игрок i имеет своей целью так выбрать функции u_i , w_i , G_u , чтобы максимизировать генеральный функционал \dot{I} .

Время T , в течение которого происходит игра, может быть фиксированным или свободным. То же самое можно сказать и о значениях $x_i(T)$.

2. Описанный класс задач теории дифференциальных игр весьма широк. В зависимости от структуры функционалов можем получить тот или другой тип игры.

Если

$$N \geq 1, \quad \dot{I} = \|x\|, \quad I_1 = \|x\|,$$

а T не фиксировано, то получаем один из вариантов игр преследования.

Какие-либо общие заключения относительно столь широко поставленной задачи, в том числе и методы вычислительного характера, вряд ли могут быть предложены. Для того чтобы продвинуться дальше, следует сузить рассматриваемую проблему.

Условимся в дальнейшем, что функционалы \dot{I}_i и \dot{I} интегральные:

$$\dot{I} = \int_0^T F(x_1, \dots, x_N, u_1, \dots, u_N, w_1, \dots, w_N, G_u) dt, \tag{3.7}$$

$$\dot{I}_i = \int_0^T F_i(x_i, v_i, w_i) dt.$$

Следующее упрощение состоит в снятии ограничений. Это может быть сделано разными способами.

Одним из примеров ограничения (3.5) является условие $v_i(t) \geq 0$, где $v_i(t)$ определяется равенством (2.4). В этом случае можем вместо функции $v_i(t)$ ввести новую управляющую функцию $y_i(t)$ с помощью равенства

$$v_i(t) = y_i(t) \{ (c, P_i) - L_i - w_i(x_i, t) - R_i(x_i) \}.$$

В этом случае снятие ограничений не вносит какого-либо изменения в природу задачи. В других случаях будем вводить функции штрафов.

В результате заменим исходную задачу другой. Ее решение будет аппроксимировать решение исходной задачи, но искомые параметры и функции уже не будут стеснены ограничениями.

Повторим ее новую формулировку. Игрок i сообщает игрокам j функции $u_i(t)$ и $w_i(x_i, t)$ таким образом, чтобы максимизировать функционал

$$I_i^* = \int_0^T F^*(x_1, \dots, x_N, u_1, \dots, u_N, w_1, \dots, w_N) dt. \quad (3.8)$$

Игроки j выбирают свои стратегии $v_j(t)$ так, чтобы максимизировать функционалы

$$I_j^* = \int_0^T F_j^*(x_j, v_j, w_j) dt. \quad (3.9)$$

В выражениях (3.8), (3.9) «звездочка» означает, что функционалы записаны в форме, учитывающей функции штрафа, с помощью которых мы сняли ограничения.

3. Для решения задачи в такой упрощенной постановке могут быть предложены методы, использующие необходимые условия экстремума. Рассмотрим сначала задачу для игрока i . Введем функцию Гамильтона H_i :

$$H_i = f_i(x_i, u_i, v_i, t) + F^*(x_i, v_i, w_i).$$

Поскольку игрокам j игрок i сообщает заранее функции $u_i(t)$ и $w_i(x_i, t)$ и поскольку в задаче (3.9) на функцию $u_i(t)$ никаких ограничений не наложено, то необходимое условие максимума функции Гамильтона можно записать в виде

$$\frac{\partial H_i}{\partial v_i} = f_{i,v_i} + \frac{\partial F^*}{\partial v_i} = 0.$$

Это условие будем писать в форме

$$g_i(x_i, u_i, v_i, t) = 0. \quad (3.10)$$

Условие (3.10) — векторное. Оно имеет размерность вектор-функции $v_i(t)$.

Множители Лагранжа λ_i удовлетворяют следующему векторному уравнению:

$$\dot{x}_i = \frac{H_i}{x_i} - \frac{f_i}{x_i} - \frac{F_i^*}{x_i} - \frac{F_i^*}{w_i} \frac{w_i}{x_i}. \quad (3.11)$$

Поскольку никаких ограничений на правый конец траектории не наложено, приходим к следующему результату.

Для того чтобы функция $v_i(t)$ была оптимальной для игрока i , необходимо, чтобы она удовлетворяла условию (3.10), где \dot{x}_i , x_i удовлетворяют уравнениям (3.1) и (3.11) и граничным условиям

$$x_i(0) = x_{i0}, \quad \dot{x}_i(T) = 0, \quad (3.12)$$

где x_{i0} — известные векторы.

4. Написать подобного рода необходимые условия для \dot{x}_i мы не можем, поскольку, помимо «стандартной» управляющей функции $u_i(t)$, в распоряжении игрока находится выбор функции $w_i(x_i, t)$. Таким образом, игрок решает задачу типа синтеза, для которой не существует достаточно простых необходимых условий.

В теории синтеза систем управления синтезирующую функцию разыскивают обычно в некотором классе функций.

В теории иерархических систем управления естественно разыскивать функцию поощрения $w_i(x_i, t)$ в виде

$$w_i(x_i, t) = a_0^i + a_1^i x_i + a_2^i x_i^2. \quad (3.13)$$

В самом деле, коэффициентам $a_s^i(t)$ легко дать экономическую интерпретацию. Действительно, представим (3.13) в виде

$$w_i(x_i, t) = b_1^i x_i + b_2^i (x_i - x_i^*)^2, \quad (3.14)$$

где функции $b_1^i(t)$, $b_2^i(t)$ и $x_i^*(t)$ легко определяются через $a_0^i(t)$. Функция $x_i^*(t)$ — это плановое задание предприятию, $b_1^i x_i$ — поощрение ($b_1^i > 0$) за выпуск продукции, $b_2^i (x_i - x_i^*)^2$ — наказание за невыполнение плана ($b_2^i > 0$).

Если принять аппроксимацию (3.13), то задача игрока будет следующей задачей оптимального управления: определить функции $x_i(t)$ и $\dot{x}_i(t)$, $i = 1, \dots, N$, удовлетворяющие уравнениям (3.1) и (3.11), и функции $v_i(t)$, $u_i(t)$, $a_0^i(t)$, $a_1^i(t)$ и $a_2^i(t)$, доставляющие максимум функционалу (3.8) и удовлетворяющие ограничениям (3.10) и (3.12).

5. В заключение этого раздела сделаем несколько замечаний о вычислительном аспекте сформулированной задачи (задача).

а) Задача является краевой задачей весьма высокой размерности. Размерность фазового вектора (x, \dot{x}) равна $2 \sum_{i=1}^N n_i$, размерность управляющего вектора (u, v, a_0, a_1, a_2) равна

$$\sum_{i=1}^N m_i + \sum_{i=1}^N r_i + 3 \sum_{i=1}^N k_i.$$

Высокая размерность задачи делает ее с вычислительной точки зрения весьма трудной. Облегчающим обстоятельством является то, что на значения $x_i(T)$ не наложено никаких ограничений. Последнее обстоятельство позволяет использовать эффективные итерационные методы [6].

б) Вычислительные трудности заставляют использовать рациональные варианты теории возмущений. В частности, нетрудно построить метод теории возмущений, который сводит задачу к решению N задач размерности $2n_i$ каждая из двух задач Коши для системы линейных дифференциальных уравнений порядка $2n_i$.

Этот факт может служить источником для построения рациональных схем имитационных процедур.

в) Теория возмущений позволяет установить ряд важных результатов качественного характера.

4. Задачи с неполной информацией

1. Как ни сложны задачи, которые обсуждались в предыдущих разделах, они тем не менее еще не содержат одной важной особенности иерархических систем. В рамках рассмотренных моделей еще нельзя объяснить необходимость введения иерархической структуры в сложных управляемых комплексах (технических, экономических, социальных и т. д.). В самом деле, введение организационной структуры — это некоторые дополнительные ограничения. Тот факт, что функция $v_i(t)$ выбирается из условия максимума функционала $\dot{I}_i(v_i)$, может только снизить значение максимума функционала $\dot{I}(u, v, w)$. Если вернуться к экономической интерпретации, то описанный факт означает, что фирме выгодней лишить предприятия права выбирать функции $v_i(t)$, т. е. полностью централизовать систему.

Обозначим через \hat{I}_{op} оптимальное значение функционала Центра (фирмы) в условиях децентрализации, в условиях полной централизации — через \hat{I}_{op} . Для тех примеров, которые рассмотрены в разделах 2 и 3, всегда

$$\hat{I}_{op} - I_{op} \geq 0. \quad (4.1)$$

Величину $\hat{I}_{op} - I_{op}$ можно принять в качестве «цены» децентрализации. Если $\hat{I}_{op} - I_{op} > 0$, то децентрализация невыгодна.

В рассмотренных случаях всегда $\hat{I}_{op} - I_{op} \geq 0$. И тем не менее в реальных системах структура управления носит всегда иерархический характер. Следовательно, модели 2 и 3 не содержат аргументов, объясняющих необходимость иерархии в процедурах принятия решений.

В чем может быть смысл этих аргументов? Автору кажется, что главным мотивом является следующий факт. Предприятие может располагать информацией, которая либо принципиально недоступна Центру (руководству фирмы), либо приходит с запаздыванием, либо ее централизованная обработка стоит очень дорого. Именно поэтому фирме может оказаться выгодным передать часть своих прав предприятию, т. е. разрешить ему стремиться к достижению своих собственных целей.

2. Прежде чем переходить к формализации этого факта, рассмотрим один пример.

Предположим, что эволюция управляемой системы описывается уравнением

$$\dot{x} = f(x, u, \theta, t), \quad (4.2)$$

где x — фазовая переменная, u — управление, θ — некоторый параметр или функция, которая заранее не известна. Априорная информация о функции f такова:

$$f(t) \in G, \quad (4.3)$$

где G — некоторое множество.

Управление $u(t) \in G_u$ выбирается из условия

$$\hat{I}^* = \max_u I(u). \quad (4.4)$$

Найдем гарантированный результат в условиях информированности (4.3). Предположим, что выбрано некоторое управление $u = \hat{u}(t)$. Тогда значение функционала определяется не только управлением \hat{u} , но и функцией $(t): I = I(\hat{u}, \cdot)$. В этих условиях гарантированный результат $\hat{I}(\hat{u})$ будет

$$I(\hat{u}) = \min_G I(\hat{u}, \cdot).$$

И, следовательно, выбирая управление $\hat{u}(t)$ из условия (4.4), можно гарантировать следующий результат, который обозначим \hat{I}^* :

$$\hat{I}^* = \max_u \min_G I(u, \cdot). \quad (4.5)$$

Предположим теперь, что кроме априорной информации (4.3) мы знаем заранее, что в каждый момент времени нам будет известно значение параметра \cdot .

Как использовать эту дополнительную информацию?

В этом случае, очевидно, управление следует искать в виде $u = u(\cdot, t)$. Тогда, если будем следовать условию (4.4), то найдем значение функционала I^* , которое будет зависеть от \cdot :

$$I^*(\cdot) = \max_u I(u, \cdot).$$

Следовательно, результат (доход), который можно гарантировать в этом случае, будет

$$I^* = \min_G \max_u I(u, \cdot).$$

Но согласно теореме о минимаксе

$$\hat{I}^* = I^*. \quad (4.6)$$

Этот результат совершенно тривиален. Он означает, что если правильно использовать информацию, которой будем располагать, то результат всегда окажется лучше (или во всяком случае не хуже).

Этот пример и дает ключ к построению математической теории иерархических систем и количественной оценке качества иерархической структуры.

Примечания. 1. Все, что сказано в этом пункте, может быть рандомизировано и переписано с помощью языка теории вероятностей.

2. Рассуждения в данном примере могут быть легко распространены на более общие случаи информированности. Например, вместо гипотезы о том, что в каждый момент времени будет известно значение параметра θ , интересно и важно рассмотреть гипотезу о том, что в каждый момент будет известно, что

$$G^t \subset G. \tag{4.7}$$

3. После этих замечаний вернемся к обсуждению модели иерархической системы, которая является обобщением модели предыдущего раздела.

Динамический процесс описывается следующей системой уравнений:

$$\dot{x}_i = f_i(x_i, u_i, v_i, \theta_i(t), t), \quad i = 1, \dots, N. \tag{4.8}$$

Все обозначения, принятые в уравнении (4.8), совпадают с обозначениями предыдущего раздела. Для того чтобы придать экономический смысл нашим рассуждениям, уравнение (4.8) будем называть производственным процессом предприятия i .

Новым в уравнении (4.8) по сравнению с уравнением (3.1) является присутствие функции $\theta_i(t)$. Априори известно, что

$$\theta_i(t) \in G_i. \tag{4.9}$$

Будем считать, что эта и только эта информация доступна Центру (фирме). В процессе производства эта величина (состояние оборудования, качество материала, производительность труда и т. д.) становится известной предприятию.

Сохраним гипотезу о поведении предприятия. Будем считать, что оно стремится максимизировать функционал (3.4). Сохраним также гипотезу о его информированности. Центр делает первый ход, сообщая предприятию величину экзогенного ресурса $u_i(t)$ и функцию поощрения $w_i(x_i)$. Получив эту информацию, предприятие может реализовать акт составления плана. Но (и это очень

существенно) этот план должен иметь характер синтеза. Управляющие воздействия v_i , которые находятся в распоряжении предприятия, разыскиваются в форме

$$v_i = v_i(t, x_i, u_i, w_i). \quad (4.10)$$

Следовательно, и траектория x_i будет функцией от этих параметров:

$$x_i = x_i(t, x_i, u_i, w_i). \quad (4.11)$$

Значит, функционал (3.6), характеризующий интересы Центра, также можно представить в аналогичной форме:

$$I = I(x_1(t, x_1, u_1, w_1), \dots, x_N(t, x_N, u_N, w_N), u_1, \dots, u_N, w_1, \dots, w_N, G_u), \quad (4.12)$$

Максимизируя выражение (4.12), получаем управляющие воздействия Центра также в форме синтеза:

$$\begin{aligned} u_i &= u_i(t, x_i, \dots, x_N), \\ w_i &= w_i(t, x_i, \dots, x_N), \\ G_u &= G_u(t, x_i, \dots, x_N). \end{aligned} \quad (4.13)$$

Поскольку x_i — априори неизвестные факторы, то проблема априорной оценки качества функционирования данной иерархической системы требует еще дополнительного обсуждения. Заметим, что априорная оценка — это основной результат развиваемой теории. Умение ее вычислить позволяет сравнивать различные варианты иерархической организации, создаваемые для решения одной и той же задачи.

Для построения такой оценки автору кажется естественным использовать принцип гарантированного результата. Если в одних и тех же условиях гарантированный результат системы A больше гарантированного результата системы B , то следует предпочесть систему A .

Из процесса построения решения (4.13) следует, что гарантированным результатом будет число, вычисленное по следующему правилу:

$$I^* = \min_i \max_{G_i} \max_{u_i, w_i, G_u} I(x_1(t_1, u_1, w_1), \dots, x_N(t_N, u_N, w_N)), \quad (4.14)$$

$$u_1, \dots, u_N, w_1, \dots, w_N, G_u).$$

4. Рассмотрим теперь тот случай, когда управление динамической системой (4.8) осуществляется из единого Центра. Это означает, что все функции — не только u_i, w_i, G_u , но и v_i — находятся в распоряжении Центра и выбираются из условия максимума функционала (3.6).

Заметим прежде всего, что структура ограничений G_u перестает быть управляющим воздействием. В интересах Центра (фирмы) предельно ослабить любые ограничения на выбор управляющих воздействий. Далее, о параметрах G_i известна априорная информация (4.9) и никакой другой информации Центр иметь не будет. Следовательно, гарантированный результат функционирования системы в условиях полной централизации будет следующим:

$$\hat{I} = \max_{u, v, w, G_i} \min_{G_i} I(x_1, \dots, x_N, u_1, \dots, u_N, w_1, \dots, w_N). \quad (4.15)$$

Сравнивая \hat{I} и I^* , можно оценить качество введенной структуры.

Если $\hat{I} - I^* > 0$, то введение иерархии в системе управления ухудшает качество работы системы. Если $\hat{I} - I^* < 0$, то введение иерархической структуры делает экономический механизм более эффективным.

5. Заключительные замечания

1. Автором излагается некоторая общая схема анализа и оценки эффективности иерархической структуры в управляемой системе — именно схема. Сами модели были предельно упрощены. Это упрощение относится прежде всего к числу ступеней иерархической лестницы. Самая простая модель государственного управления требует как минимум три уровня принятия решений. Кроме того, не рассматривались горизонтальные связи между предприятиями, которые могут также явиться источником конфликта и требуют формулировки соответствующих гипотез. Но изложенная схема позволяет найти рациональные пути исследования и в более общих случаях.

Исследование иерархических структур сегодня представляется нам одним из важнейших направлений прикладной математики,

теории управления и экономики. Развитие производительных сил, усложнение технологий и все возрастающая роль планового, не рыночного, начала в распределении приводят к тенденции усиления роли централизованного управления. Однако, как было видно, имеет смысл говорить о рациональной (а в ряде случаев и оптимальной) мере централизации и децентрализации.

Этот оптимальный уровень децентрализации в принятии решений зависит от многих причин, в частности от технологии обработки данных и передачи информации. По мере совершенствования этой технологии уровень централизации будет возрастать.

В социалистических странах с их плановой централизованной экономикой экономические механизмы, обладающие иерархической организацией, имеют очень важное значение. Воздействуя надлежащим образом на целевые функции нижних уровней, верхние уровни могут обеспечить выполнение необходимых плановых показателей, которые возникают при реализации программного метода управления.

2. Одним из самых трудных вопросов обсуждаемой теории является вопрос о целевых функциях подсистем. Конечно, в простейших схемах их можно отождествлять с доходом и с помощью функций штрафов и поощрений оказывать влияние на их формирование. Однако было бы заблуждением думать, что все сводится к факторам чисто экономическим. В конечном счете каждый из организмов, следуя Эшби, стремится быть внутри некоторой области гомеостаза, которая определяется не только экономическими факторами, но и престижем, ощущением самостоятельности (трудовой комфорт), юридическими нормами и т. д.

3. Мы видели, что анализ даже простейших задач теории иерархических систем приводит к очень громоздким вычислительным задачам. Автор не думает, что можно (во всяком случае в ближайшие годы) ждать появления достаточно общих и эффективных вычислительных методов. Кроме того, надо считаться с тем очевидным фактом, что в реальных системах верхние уровни далеко не всегда будут знать достаточно хорошо целевые функции своих подсистем (звеньев).

Вот почему автор считает, что успех практического применения изложенных идей будет зависеть прежде всего от того, как мы научимся строить системы, имитирующие функционирование реальных иерархических структур, и организовывать процедуры диалога человек — машина.

Литература

1. фон Нейман Дж., Моргенштерн О. Теория игр и экономическое поведение. – М., Наука, 1970.
2. Нэш Дж. Бескоалиционные игры./В сб. «Матричные игры». – М., Физматгиз, 1961.
3. Гермейер Ю.Б. Игры с непротивоположными интересами (Теория принятия решений при искомом единстве). – М., Изд-во МГУ, 1972.
4. Аганбегян А.Г. Система моделей народнохозяйственного планирования. – М., Мысль, 1972.
5. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем. – М., Мир, 1973.
6. Моисеев Н.Н. Численные методы теории оптимальных систем. – М., Наука, 1970.

1973 г.

Имитационные модели

Чарльз Перси Сноу в статье «Две культуры», опубликованной в ежегоднике «Наука и человечество. 1970», пишет о той пропасти, которая образовалась между естественными науками, с одной стороны, и науками гуманитарными — с другой. Эта пропасть продолжает расширяться, и обе группы интеллигенции, разделенные ею, теряют постепенно не только общий язык: у них возникает разное понимание одних и тех же вещей, разные интересы и как следствие — исчезает взаимное общение. Автор считает, что создавшаяся ситуация опасна для человеческой цивилизации, ведет к ее распаду на две ветви, и видит главную причину этого в недостатках системы образования в наиболее развитых странах.

Ч.П. Сноу сравнивает системы образования, принятые в США, Англии и СССР, отмечает преимущество советской системы подготовки специалистов и характер взаимопонимания ставит в прямую зависимость от структуры образования.

Ч.П. Сноу — блестящий писатель и первоклассный физик. Он знает обе стороны пропасти. Он свой среди естественников и литераторов. Ему дано тонко чувствовать устремления и взгляды обеих групп интеллигенции. Его высказывания вполне компетентны и убедительны.

Правда, они несколько категоричны и излишне абсолютизированы. Общность марксистской идеологии, например, объединяет людей независимо от их профессии и, разумеется, служит надежным цементом советского общества. Тем не менее и у нас происходит известное разобщение отдельных групп интеллигенции по профессиональному признаку, но мне кажется, что недостатки образования — не только не единственная, но и не главная причина того явления, которому посвятил свою статью Ч.П. Сноу.

Я думаю, что этот процесс разобщения в известной степени закономерный и объективный. Он, конечно, зависит от содержания и системы воспитания, и образования, но мне кажется, что он

определяется прежде всего уровнем развития производительных сил. Он происходит, хоть и в разной степени, во всех развитых странах. Но специализация и, как следствие ее, процесс разобщения — это лишь одна из сторон сложной диалектической связанности, дифференциации и синтеза человеческих знаний, неизбежная, но не трагическая. Об этом и будет идти речь.

Огромный рост общего объема знаний, с которым оперирует человечество, требует все большей и большей локализации интересов специалистов. Универсалы постепенно исчезают.

Происходит непрерывное деление потока общечеловеческих знаний на отдельные протоки и ручейки, которые в свою очередь превращаются в мощные реки. У различных групп интеллигенции возникают различные интересы, малопонятные другим. Не только естественники начинают плохо понимать профессиональный язык гуманитариев, но и сами естественники начинают все хуже и хуже понимать друг друга. Уже сегодня даже математики разных специальностей не всегда могут объяснить друг другу смысл и значение своих исследований. Об этом еще 50 лет тому назад с тревогой говорил крупнейший немецкий математик Давид Гильберт, который опасался, что, растекаясь, поток человеческих знаний потеряет силы и иссякнет.

Обособление интересов, может быть, заметнее между гуманитариями и естественниками, но это лишь одно из проявлений общего процесса роста и дифференциации человеческих знаний.

Опасения Гильберта и Сноу можно понять, но мне кажется, что они несколько преувеличены. В самом деле, наряду с процессом разделения происходит и обратный процесс непрерывного синтеза идей и знаний, накопленных в разных научных дисциплинах. Именно на стыке разных научных «цивилизаций» появляются наиболее значительные новые исследования. Химическая физика, физическая химия, молекулярная биология — все это примеры синтеза различных научных школ и концепций, примеры, демонстрирующие диалектическое единство процессов дифференциации и синтеза человеческих знаний.

Подобный процесс начался и в общественных науках. Математика, например, начинает использоваться в экономике, социологии, лингвистике и т. д. Я думаю, что в ближайшем будущем появится еще целый ряд новых научных дисциплин, объединяющих гуманитарное и естественнонаучное мышление, требующих совместных

усилий специалистов в области общественных наук, математиков, физиков и инженеров. Решающую роль в этом процессе призваны сыграть электронные вычислительные машины.

Прежде чем переходить к обсуждению этого тезиса, сделаем небольшой экскурс в историю электронной вычислительной техники. Это необходимо нам, чтобы объяснить некоторые особенности сегодняшнего периода ее развития.

Первые электронные вычислительные машины (ЭВМ), которые начали использоваться в практической деятельности человека, были введены в строй на грани 1940–1950-х годов. В СССР это была машина «Стрела» и знаменитая БЭСМ-1, положившая начало серии БЭСМ. Эти ЭВМ принято называть машинами первого поколения. Громоздкие и ненадежные ламповые устройства тем не менее обладали весьма хорошими арифметическими возможностями, позволяя производить арифметические действия в несколько тысяч раз быстрее человека. Этого оказалось достаточным, чтобы проводить расчеты инженерных и физических задач такого объема, который раньше казался фантастическим.

Роль этих ЭВМ трудно переоценить. Вряд ли без них был бы возможен в 1950-е годы тот прогресс в физике, в создании ракет и в других областях науки и техники, который мы наблюдали.

Таким образом, ЭВМ первого поколения сыграли революционизирующую роль в технологии инженерных расчетов и исследований в физике. Однако эти ЭВМ оказались почти бесполезными при решении экономических задач. Особенность последних состоит в том, что их решение требует переработки больших массивов информации. ЭВМ первого поколения обладали очень небольшой памятью, порядка нескольких тысяч слов. Этого было явно недостаточно. Возможности ЭВМ первого поколения оказались изученными (и исчерпанными) за 4–5 лет.

В конце 1950-х годов начали создаваться ЭВМ второго поколения. С точки зрения инженера-эксплуатационника их основное отличие — замена ламп транзисторами и практически полная надежность. Однако с точки зрения пользователя главное состояло не в этом.

Значительное расширение памяти открыло возможность широкого использования ЭВМ в экономике. Следующие 10 лет прошли под знаком ускоряющегося внедрения этой техники в экономику и управление производством. Оказалось, что использование ЭВМ

дает огромный экономический эффект. Центр тяжести использования ЭВМ за эти 10 лет переместился из области научных исследований в сферу производственной деятельности человека (около 80% машин в разных странах заняты в экономике и управлении).

ЭВМ второго поколения обладали еще одной важной особенностью. Значительно возросло их быстродействие. Так, например, БЭСМ-6, классическая машина второго поколения, может производить миллион операций в секунду. Таким образом, мощность БЭСМ-6 примерно во столько же раз превосходит мощность первых машин первого поколения, насколько ЭВМ первого поколения превосходили человека (2–3 тысячи операций БЭСМ-1 и 2–3 операции в секунду, которые способен производить человек).

Последнее обстоятельство поставило перед математиками целый ряд сложных проблем. Оказалось, что не так-то просто полностью использовать их огромную мощность. Для того чтобы «поставить задачу на машину», математик должен переписать процедуру ее решения на языке, который понимает машина, как говорят, «в кодах машины». Хороший программист может делать в день около 5 команд. Поэтому на отладку больших задач, которые решались на ЭВМ первого поколения (например, аэродинамический расчет, связанный со входом тела в атмосферу) и программа которых имела тысячи команд, тратились годы кропотливого и изнурительного труда математика.

И вот математику с его привычной техникой программирования предоставляют возможность работать на машине в 1000 раз более быстродействующей. Значит, для того чтобы ее загрузить, надо в 1000 раз больше математиков-программистов. Но это еще не все. Чтобы исчерпать возможности ЭВМ, т. е. оправдать целесообразность создания новых машин, надо не только увеличить количество задач, доступных ЭВМ первого поколения, надо еще научиться решать качественно более трудные задачи. ЭВМ второго поколения может решать задачи, содержащие сотни тысяч команд. Но при старой технологии составление каждой такой программы потребует сотен лет.

Комментарии излишни! Перед математиками 1960-х годов стала дилемма: либо признать неспособность использовать полностью все возможности ЭВМ второго поколения, либо... Либо создать качественно иную систему общения человек — машина. Разумеется, история пошла по второму пути. Были созданы алгоритмические

языки, позволившие качественно по-иному транслировать вычислительные процедуры с человеческого языка на машинный. Грубо говоря, математики научились перекладывать тяжесть программирования со своих плеч на машинные.

Для того чтобы полностью изучить все возможности ЭВМ второго поколения и научиться их использовать, человеку потребовалось около 12 лет.

Как ни велико было значение ЭВМ второго поколения, однако настоящая революция в методах научного анализа, как мне кажется, будет связана с освоением машин третьего поколения, которые начали создаваться в конце 1960-х годов.

Чем существенным эти ЭВМ отличаются от машин второго поколения? Если спросить инженера — специалиста по ЭВМ, то он, вероятно, ответит так: ЭВМ третьего поколения — это вычислительные системы, основанные на интегральных схемах. Действительно, переход на интегральные схемы, которые позволили заменить целые блоки машин одним «элементом», — это важный этап технологической эволюции ЭВМ. Он позволяет перейти, в частности, к поточной сборке ЭВМ, многократно увеличивает надежность и ведет к удешевлению ЭВМ в перспективе на 2–3 порядка. Но с точки зрения пользователя важны другие характеристики. Прежде всего традиционные — быстродействие и память.

Быстродействие. Здесь не произошло того качественного скачка, который был между машинами первого и второго поколений. Самые мощные системы третьего поколения имеют быстродействие порядка десяти миллионов операций в секунду, т. е. они мощнее, чем БЭСМ-6, не более чем в 10 раз. Поэтому не рост быстродействия, как он ни важен, определяет сегодня значение этих ЭВМ.

Память. Создание дисковых носителей с очень высокой скоростью поиска и считывания нужной информации, позволяющей к тому же практически неограниченно наращивать объем памяти, — это значительно более важная для нас характеристика.

Уже сегодня объем памяти каждой из действующих систем третьего поколения позволяет иметь в своем распоряжении информацию, которой располагает публичная библиотека средних размеров. Причем поиск необходимой информации, нужной цитаты, цифры и т. д. занимает неизмеримо меньше времени, чем требуется самой квалифицированной, налаженной библиотеке.

Этот факт будет иметь разнообразные последствия. Возьмем лишь одно из них — постановку архивного дела. Переписывая документы на магнитные носители или сразу ведя запись на специальных устройствах типа перфорированных или магнитных лент, мы будем иметь архивы, где «мгновенно» можно отыскать любой документ, получить любую справку. Качественно изменится характер труда многих специалистов, связанных с поиском и анализом документов, прежде всего историков, изучающих архивы. Они будут освобождены от тяжелой технической работы. Их основная задача в будущем — ставить вопросы поисковой системе. Труд станет значительно более творческим. Я думаю, что уже в текущем десятилетии мы окажемся свидетелями повсеместного перехода к новым методам хранения и извлечения необходимой информации.

Но перечисленным не исчерпываются все отличия ЭВМ третьего поколения от своих предшественниц. Эти машины, как уже заметил читатель, принято называть системами. Это действительно системы, в которых центральный процессор — основное арифметическое устройство, которое, собственно, и предназначено для того чтобы производить вычисления, взаимодействует с целым рядом вспомогательных или, как говорят, периферийных устройств. Эти устройства, может быть, и составляют основное, что внесено в вычислительную технику вместе с появлением ЭВМ третьего поколения.

Эти периферийные устройства вместе с увеличившейся производительностью и памятью предельно упрощают диалог человек — машина, позволяют создавать человеко-машинные системы, объединяющие возможности ЭВМ быстро производить формально-логические операции с неформальным мышлением и талантом человека.

Что это за устройства? Это прежде всего различного рода системы отображения. Если все результаты расчетов, всю информацию, которую создает ЭВМ, представить в форме цифр, то получится таблица, анализировать которую трудно. Нужно уметь представлять информацию в агрегированном виде, позволяющем наглядно увидеть тенденции изучаемого процесса. Простейшая система отображения — это представление информации в графической форме.

Но система отображения — это прямая связь машина — человек, а современные периферийные устройства позволяют эффективно реализовать и обратную связь человек — машина. Один из примеров таких устройств — световое перо.

Предположим, что исследователь, который изучает некоторый процесс, хочет выяснить, как влияют те или другие зависимости на его течение. И пусть эти зависимости задаются в графическом виде. Тогда световое перо позволяет ввести в ЭВМ новую информацию, т. е. новые зависимости, прямо «от руки». Исследователю достаточно с помощью этого пера нарисовать новый график. Система отображения позволит ему сразу увидеть, как изменилось течение процесса вследствие изменения исходной информации, т. е. структуры исходных зависимостей.

Перечисленные технические новшества и окажут революционизирующее влияние на эволюцию технологии научных исследований.

* * *

Появление новых дорогостоящих электронных вычислительных устройств ставит трудные проблемы перед их создателями и учеными разных специальностей. Как научиться их правильно использовать? Как с максимальной полнотой использовать их возможности, с наибольшей эффективностью заставить служить эти системы прогрессу человеческого общества?

Вопросы эти важны с разных точек зрения и прежде всего — с экономической. Вкладывая миллиарды в создание электронной вычислительной техники, естественно ожидать и соответствующей отдачи.

Я не собираюсь здесь анализировать экономическую сторону вопроса, а хочу попытаться показать, что решать эти проблемы предстоит именно ученым разных специальностей, в том числе и гуманитариям. А для этого сначала необходимо объяснить понятие «модель процесса» и эволюцию этого понятия, когда речь касается процессов, происходящих в человеческом обществе.

Понятие модели возникло в физике. Уже более 300 лет, начиная со времен Ньютона, для описания изучаемых процессов физики используют язык математики. Законы Ньютона были одной из первых важных моделей процессов, изучаемых физикой, — моделью механического движения. По мере расширения и углубления наших знаний совершенствовались также и математические модели*, которые все с большей точностью описывали изучаемые явления. Появились модели электромагнитного поля, процессов квантовой механики и т. д.

* В дальнейшем со словом «модель» мы будем связывать только понятие математической или формализованной модели — приближенного описания процесса с использованием языка математики.

Эти модели основывались на огромном экспериментальном материале, который и определял их достоверность. В свою очередь, эти модели приобретали огромную эвристическую силу и становились средством, с помощью которого устанавливались новые факты. Открытие Нептуна и Плутона были лишь эпизодами истории развития метода моделей, когда новые физические открытия были преподнесены человечеству на кончике пера.

Но наиболее яркой демонстрацией тех возможностей, которые дает использование математических моделей, было овладение атомной энергией. Вспомним, что никакого экспериментального материала о характере течения цепной реакции в распоряжении физиков до этого еще не было!

Область применения метода формализованных моделей все время расширялась. Уже во многих разделах химии, биологии и других естественных наук широко используют методы математического описания. В последние десятилетия этот метод перешагнул границу общественных наук. Его «первой жертвой» сделалась экономика. Справедливости ради необходимо сказать, что математические модели в экономике, той из общественных наук, которая в большей степени, чем другие, использует количественные характеристики, появились уже давно. Однако действительно широкое использование моделей в экономике началось только в послевоенный период, когда успехи электронной вычислительной техники сделали доступным анализ достаточно сложных моделей.

Оказалось, что, как правило, описание общественных процессов (модель общественного процесса) незамкнуто. В модели присутствуют свободные параметры или функции, которые находятся в распоряжении человека и могут быть выбраны (назначены) по его желанию. Это отражает тот факт, что процессы, протекающие в обществе, — управляемые. Распоряжаясь надлежащим образом ресурсом, человек может в значительных пределах влиять на течение экономического, военного, политического — одним словом, исторического процесса. Поэтому модели общественных процессов оказались на много порядков сложнее тех, которые до сих пор изучались математическими методами. Электронная вычислительная техника и явилась базой для построения методов их анализа. В послевоенное время широкое распространение получили идеи оптимизации — идеи отыскания такого решения, которое было бы наилучшим в смысле того или

иного критерия. Оказалось, что целый ряд экономических проблем может быть решен с помощью подобных моделей. Эти модели мы условимся называть оптимизационными.

Простейший пример оптимизационных моделей — модель транспортных перевозок. Задача состоит в том, чтобы построить такой план перевозок какого-либо продукта из баз снабжения (складов) потребителям (например, магазинам), чтобы запросы потребителя были выполнены, а затраты на перевозку были минимальны. Эта задача может быть четко сформулирована как математическая задача и решена методами математики.

Несмотря на то, что задача о перевозках формулируется как математическая задача, в ней имеется один элемент, который качественно отличает ее от обычных задач естествознания, — критерий. Мы хотим подешевле осуществить перевозки. Слово «хотим» — это некоторая «модель поведения действующего лица». В этой простой задаче все очень просто. Но в сложных задачах понять, чего хочет действующее лицо, а без этого невозможно сформулировать задачу, — бывает часто очень трудно. Значит, уже в самом начале использования математических методов при анализе общественных процессов мы сталкиваемся с необходимостью объединения неформального анализа — выработки структуры наших желаний (модель поведения) с формальными математическими методами — процедурой решения.

К сожалению, анализ реальных ситуаций далеко не всегда можно свести к таким четко поставленным оптимизационным задачам, как задача о построении плана наиболее дешевых перевозок. В задаче могут быть разного рода неопределенности. Планируя, например, перевозки зерновых, мы сталкиваемся с неопределенностью природы: мы не знаем заранее достаточно точно погодных условий, а они будут влиять на сроки созревания и, следовательно, выполнение плана. Планы международной торговли существенно зависят от не известных полностью действий наших конкурентов. Наконец, очень часто бывает трудно и невозможно сформулировать критерий, охарактеризовать результат операции одним показателем из-за неопределенности желаний.

Объективная реальность состоит в том, что обычно нам приходится принимать решения в условиях неопределенности. Для решения подобных задач уже не годятся оптимизационные модели. Для их исследования разработаны специальные подходы и специальные

модели, которые принято называть конфликтными или игровыми моделями.

Оптимизационные и игровые модели сегодня уже прочно вошли в арсенал средств изучения процессов в экономической и военной сферах. Они становятся основой современного планирования и прогнозирования.

Несмотря на то что подобные модели уже начали широко использоваться и никто из практиков не ставит под сомнение их эвристическую ценность, они сами ставят перед специалистами в области общественных наук целый ряд гносеологических проблем, аналогичных тем, которые возникли в конце XIX в. в физике и послужили источником агностицизма. Проблема анализа моделей общественных процессов с позиций теории познания, вероятно, в ближайшие годы привлечет к себе внимание философов, как это случилось в свое время с проблемами физики.

* * *

Модели, о которых шла речь, пригодны для анализа лишь относительно наиболее простых явлений. Сегодня мы, математики, которых часто обвиняют в том, что они стремятся все и вся математизировать, отлично понимаем, что лишь небольшая часть проблем, стоящих перед человечеством, поддается математической формализации и описанию на языке математики. И это не следствие слабости математики. Эта ситуация отражает тот факт, что человек приобретает уверенность в истинности какого-либо обстоятельства не только с помощью чисто логических процедур. Ассоциация, интуиция, эмоциональное восприятие и т. д. являются столь же законными способами познания, как и чисто логические процедуры. Последние имеют лишь одно преимущество — они могут быть легко реализованы на ЭВМ. Машина, к сожалению, не может мыслить неформально.

Поэтому дальнейшее совершенствование методов научного анализа в общественных науках должно основываться не на стремлении к стопроцентной математизации этих дисциплин, а к естественному объединению возможностей ЭВМ, которая делает логические операции в миллионы раз быстрее и лучше человека, с человеческим творчеством, со способностью человека мыслить неформальными категориями.

Этот альянс уже происходит. Оценить его перспективы, найти правильные пути его реализации — это сегодня одна из важнейших задач большой науки. Реализация этого альянса и приведет, с моей точки зрения, к появлению новых научных направлений, которые потребуют синтеза гуманитарной и естественнонаучной «цивилизаций».

Техническую базу развития соответствующих методов исследования я вижу в создании человеко-машинных систем. В их основе лежит возможность обработки и переработки мнений экспертов — специалистов в разных областях науки, возможность достаточно быстрой проверки следствий тех решений, которые они принимают, возможность организации процесса обучения (тренинга) и т. д.

По существу, все эти элементы уже присутствуют в современных методах организации сложных экспертиз. За последние 10–15 лет создана весьма эффективная технология получения экспертных оценок и прогнозов. В основе этой технологии лежит идея расчленения большого, трудного вопроса на ряд мелких, доступных для решения. Из ответов на эти мелкие вопросы тем или другим способом научились «собирать» ответ на исходный трудный вопрос. Сегодня мы можем оценивать степень достоверности такого анализа. Нам понятны и границы применимости системы методов, развитых за последние годы. Подробное изложение этих подходов уведет нас в сторону от основной задачи этой статьи. Отметим только, что даже без использования электронной вычислительной техники иногда удается получать ответы очень высокой достоверности. Знакомство с этой технологией может, вероятно, оказать большое влияние на совершенствование инструментария общественных наук.

Однако методы организации экспертных процедур и получение с их помощью оценок и прогнозов — это лишь первый шаг на пути объединения формальных и неформальных методов научного анализа. Метод экспертиз, в общем-то, имеет довольно ограниченную сферу применимости. Конечно, результаты экспертных прогнозов существенно зависят и от таланта и компетенции экспертов, и от организации процедур их обучения и обработки результатов. Тем не менее для экспертиз современного типа всегда нужен прецедент. Представим себе, что в 1940 году собрали бы всех виднейших физиков и с использованием всех современных

достижений метода экспертиз сделали бы прогноз научно-технического прогресса на 20 лет. Можно быть уверенным, что картина мира, которая была бы нарисована, не содержала бы ни атомных электростанций, ни спутников, ну и, конечно, в ней не нашло бы места самое замечательное достижение человеческого гения — электронная вычислительная техника.

Значительно более эффективным средством анализа и прогноза являются так называемые имитационные модели. Этот новый тип моделей появился в самые последние годы и связан с новым этапом развития вычислительной техники — с появлением машин третьего поколения. Использование имитационных моделей с помощью ЭВМ второго поколения весьма затруднительно. Что же такое имитационные модели?

Выше мы говорили о моделях оптимизационных и игровых. Оба эти класса моделей дают возможность четкой формулировки определенных математических задач. Для этого, в свою очередь, необходима формулировка на математическом языке целого ряда условий, определяющих протекание процесса, и прежде всего — понятия цели управления. Уже упоминались трудности, с которыми приходится сталкиваться, когда мы не можем четко сформулировать критерий. В этом случае возникает неопределенность в оценке возможных вариантов развития изучаемого процесса. Но часто оказывается, что не только критерий, но и понятие цели управления мы не можем сформулировать на языке математики. Кроме того, цель может изменяться в зависимости от развития процесса.

Классическим примером такого процесса является развитие международной торговли. Конечно, каждая страна или фирма обычно знает, чего она хочет и что такое в той или иной ситуации хорошо, а что такое плохо. Но перевести эти цели и критерии на язык математики, как правило, невозможно. Поэтому проблему выбора стратегии в том сложном многостороннем процессе, который называется международной торговлей, нельзя свести к математической задаче.

К числу подобных процессов, изучение которых нельзя свести к математической задаче или цепочке задач, относятся, например, все задачи внешней политики, процессы социального развития, войны и т. д. Для описания и изучения подобных ситуаций и развивается метод имитационных моделей.

Модель представляет собой совокупность уравнений и соотношений, обуславливающих данный изучаемый процесс. Всякий процесс,

происходящий в обществе, как это уже было сказано, является управляемым. В модели присутствуют «управления». В моделях чисто экономических, например в модели международной торговли, это объем экспорта или импорта отдельных стран, уровень цен и т. д. К числу управляющих факторов можно отнести также и различные соглашения, которые могут заключать партнеры.

Если эти управляющие факторы (стратегии) известны, то имитационная модель позволит с большой степенью точности предсказать возможное течение процесса и финальное состояние изучаемой системы. Таким образом, в зависимости от задания варианта стратегий мы можем находить тот или иной исход изучаемого процесса.

Так вот, формально отличие имитационных моделей от оптимизационных и игровых состоит в том, что задание стратегий и сравнение результатов, к которым они приводят, производит эксперт.

В оптимизационных моделях, например, выбор стратегий — это результат решения некоторой математической задачи. Когда мы имеем дело с оптимизационными моделями, то мы формулируем некоторую гипотезу поведения — некоторую функцию, значение которой оценивает качество решения и правило его выбора. В имитационных моделях это правило поведения не может быть формализовано, так как принимает и оценивает решения человек. Это его творческий вклад в работу модели.

Таким образом, имитационные модели — это человеко-машинные системы, позволяющие объединить неформальное мышление и талант специалиста с методами математического моделирования.

Имитационные модели имеют широкое и разнообразное применение. Многократное «проигрывание» модели для разных стратегий позволяет понять роль и значение различных факторов для течения того или другого процесса. Это обстоятельство открывает огромные возможности для научного использования имитационных моделей в технике, экономике, социологии, юриспруденции и т. д.

Поскольку сама проблема сводится прежде всего к изучению того эффекта, к которому приведут наши управленческие акты, то имитационные модели найдут очень широкое применение в управлении и со временем, я в этом убежден, сделаются основой автоматизированных систем управления.

* * *

Имитационные модели уже начали использоваться. Первыми потребителями этих моделей стали, естественно, инженеры.

Приведем один пример, где использование имитационных моделей имеет огромное будущее. Любые проблемы автоматизации проектирования сводятся в конечном счете к построению моделей, имитирующих функционирование проектируемой конструкции. Сегодня в связи с техническим прогрессом человечество стало создавать конструкции все более и более сложные, завершение которых требует более чем десятка лет. Создание и ввод в эксплуатацию сверхзвукового пассажирского самолета типа Ту-144 или «Конкорд» требует 12–15 лет. При нынешних темпах развития научных знаний идеи, заложенные в его конструкцию в начале проектирования, будут уже безнадежно устаревшими в тот момент, когда этот самолет выйдет на линию. Имитационные модели, в которых имитируются функционирование самолета в разных условиях и процедуры его совершенствования, имитационные модели, которые снимают с коллектива конструкторов всю тяжесть рутинной работы по расчетам и предварительным экспериментам, имитационные модели, которые оставляют за конструктором только творческий акт анализа вариантов, их сравнения и оценки, — по моему мнению, единственный эффективный путь преодоления всех подобных трудностей. Работы в области автоматизации проектирования уже начались в разных странах, и достигнутые успехи связаны прежде всего с построением подобных человеко-машинных систем.

Другой не менее важный пример, где имитационные модели, вероятно, играют решающую роль, — это АСУ (автоматизированные системы управления). Мы научились относительно просто решать проблемы систем сбора и передачи информации, автоматизации документооборота, бухгалтерского учета и ряда других вспомогательных операций. Это, конечно, немало!

Но главное в автоматизированных системах — это процедура принятия решений. Их задача — обеспечивать правильность принимаемых решений и свести до минимума ошибочные варианты. И вот эта процедура пока еще не испытала существенного влияния автоматизации вспомогательных операций.

Здесь не место перечислять все причины этого явления — они достаточно сложны. Но на одно обстоятельство имеет смысл обратить внимание. Основное развитие экономической мысли начиная

с 1950-х годов связано с идеей оптимизации. Оптимальные планы, оптимальные решения — область применения этих понятий оказалась значительно уже, чем это предполагалось сначала.

Дело даже не в том, что обычно очень трудно определить критерий оптимальности. Не менее трудно бывает вычислить оптимальное решение. Так, например, одна из важных задач — задача календарного планирования, составление расписания работ. Процедура, которая позволяет найти оптимальное решение в этой задаче, хорошо известна. Однако если количество работ равно 1000, то самой быстроедействующей машине потребуется для этого время, равное времени существования Солнечной системы!

Но пусть, наконец, мы не только имеем критерий оптимальности, но и умеем легко находить оптимальное решение. Теперь нам остается реализовать его на практике. И вдруг обнаруживается, что мы или не учли некоторых важных обстоятельств, или произошло нечто неожиданное, не предусмотренное при составлении плана. Наш оптимальный план рушится, и лучшим может оказаться план далеко не оптимальный!

Вот почему на практике область концепций оптимальности оказывается ограниченной. Поэтому построение процедур принятия решений в АСУ на базе оптимизации не имеет особых успехов, и сами АСУ играют пока лишь вспомогательную роль.

Я думаю, что завершение создания АСУ на базе ЭВМ второго поколения без использования имитационных моделей, по существу требующих простого и быстрого общения человека с машиной — основной характеристики ЭВМ третьего поколения, — вряд ли возможно. Имитационные модели, позволяющие быстро оценить эффект того или другого решения, быстро просмотреть серию вариантов, назначаемых человеком, — единственное средство автоматизации процедур, принятия решения.

* * *

Перечисленными примерами далеко не исчерпывается область применения имитационных моделей. Более того, я думаю, что мы еще очень далеки от полного понимания мощи и значения этого средства научного анализа, которое приобрело человечество. Может быть, основное значение метода имитационных моделей лежит как раз не в технической и естественнонаучной сфере, а в области изучения и управления процессами, происходящими в обществе.

Я думаю, в частности, что метод имитации окажет неоценимые услуги в дальнейшем изучении развития человеческого общества и сделается со временем мощным инструментом исторического материализма, позволяющим глубже понять роль всех факторов, определяющих это развитие, понять, как изменяются возможности человечества влиять на собственную эволюцию, как они растут вместе с техническим прогрессом и совершенствованием организации человеческого общества.

Эти вопросы представляют не только теоретический интерес. Сегодня человечество распоряжается энергетическими ресурсами планетарного масштаба, оно способно совершать действия, изменяющие лик Земли, поставить себя на грань катастрофы или, наоборот, открыть двери такому прогрессу, о котором не могли мечтать и самые безответственные фантасты.

Всегда, когда человеческое общество действительно остро нуждалось в том или другом изобретении, это изобретение появлялось. Когда два века назад промышленное развитие, казалось, зашло в тупик, поскольку возможности ручного труда, мускульной энергии человека и животных оказались исчерпанными, появилась паровая машина. Когда в послевоенное время возникла реальная опасность информационного тупика, появилась электронная вычислительная техника.

Подобная ситуация складывается и в той области, о которой идет речь в этой статье. Идеи имитации открывают реальную возможность создать инструментарий предвидения, позволяющий заранее знать результаты наших действий, в том числе и глобального масштаба. Сегодня не стоит вопрос о том, нужны ли имитационные модели. Они уже начали создаваться и служат человеку. Мы можем только ускорить или замедлить этот объективный процесс.

Значительную часть нашей планеты теперь занимают социалистические государства, упорядоченность управления которых существенно выше, чем в капиталистическом обществе, значительная часть сил, неконтролируемых в условиях капитализма, теперь может быть использована для целенаправленного развития нашего общества.

Имитационные модели могут составить основу интегрированного планирования и управления. С их помощью могут изучаться и сравниваться различные варианты экономических, социальных, политических программ, программ развития регионов и отраслей.

Имитационные модели могут быть использованы и для решения других вопросов общегосударственного значения, таких, например, как выбор рациональной иерархии управления промышленностью и т. д. Вот почему идеи имитации имеют чрезвычайно большие перспективы именно в социалистических государствах, и поэтому именно у нас интерес к ним столь высок.

Использование имитационных моделей в государственном управлении — актуальнейшая проблема, стоящая перед нашими специалистами, которые занимаются использованием вычислительной техники.

* * *

Имитационные модели — это один из мощнейших инструментов научного анализа в общественных науках и прежде всего в истории. У нас, в Вычислительном центре Академии наук СССР, интенсивно обсуждается проблема построения имитационной модели какого-либо известного исторического процесса. Именно известного и хорошо изученного историками, с информационным обеспечением. На этом примере нам хочется отработать все особенности проектирования таких моделей. Зная причины, модель будет выдавать следствия. Если модельные следствия будут значительно отличаться от тех, которые на самом деле имели место, то значит, конструкция модели плоха, ее надо совершенствовать.

Выбор пал на XIX в., XX в. чересчур сложен. Очень сложной сделалась экономика, бесконечно усложнились связи между государствами, исчезли локальные исторические процессы. XVIII в. более прост для имитации, однако он недостаточно насыщен конкретной информацией. Начало XIX в., более точно — период наполеоновских войн, нам кажется самым удобным периодом для нашего эксперимента.

Вообще выбор объекта исследования очень важен. Ведь это первая попытка испытать новый инструмент. Историк, вероятно, было бы интереснее изучать более общие и глубокие процессы развития капитализма, например, построить модель развития капитализма в России в период от освобождения крестьян до отставки Витте.

Наши аргументы в пользу наполеоновской эпохи носят чисто технический характер. В эту эпоху большую роль играли войны и отдельные сражения. Их модели относительно просты, а сами

события хорошо изучены в деталях. Социальные модели России второй половины XIX в. значительно сложнее.

Изучение исторических процессов с помощью имитационных моделей может дать историку материал уникальной ценности. Экспериментируя с различными вариантами решений экономического, политического, военного, правового характера, мы можем быстро и наглядно выявить объективные закономерности, скрытые под грандиозным нагромождением фактов.

В истории много субъективного, многое зависит от решений, которые принимает та или другая историческая личность. Конечно, историю творят массы, но отдельные события могли бы произойти иначе, если бы иначе вел себя тот или другой деятель или военачальник. Наполеон мог проиграть Аустерлиц, но выиграть Ватерлоо. Насколько это повлияло бы на течение общего исторического процесса, на темпы развития общества, на характер развития капитализма во Франции и т. д.?

Имитационная модель любого процесса — это прежде всего совместное творчество специалистов, которые изучают этот процесс, и математиков, которые занимаются инструментом — конструированием модели. Именно совместное. Только слияние этих обеих «цивилизаций» позволит сделать не игрушку, а инструмент, полезный для человечества.

И теперь задача состоит в том, чтобы найти партнеров. Конечно, это не так-то просто найти партнера-историка. Этот партнер должен удовлетворять многим требованиям. Во-первых, он должен обладать высокой профессиональной подготовкой историка. И, во-вторых, он должен обладать мужеством настоящего таланта. Он должен шагнуть в неведомое ему будущее с риском, который неизбежен в настоящей науке, и на годы подчинить себя этой новой для гуманитариев и не очень понятной на первых порах форме деятельности.

Мы уже не раз сталкивались с трудностями подобного рода. В 1950-х годах создавалось взаимопонимание математика и экономиста. Сегодня они работают часто плечом к плечу, и многие трудные вопросы остались позади. Математики многому научились у экономистов, а экономисты восприняли наш инструментарий. Проблемы конкретной экономики, планирования решаются теперь совместно. Но чем дальше от конкретного числа, тем труднее и тернистее становится дорога. Естественно поэтому, что нам так трудно навести мосты с историками.

Но я оптимист и убежден, что созданное средство труда неизбежно будет использовано. Сегодня уже не только математики понимают, что главная сила математики — методы качественного анализа, без которого нельзя получить и надежной цифры. Имитационная техника в еще большей степени — инструмент качественного анализа.

Может быть, эта статья и поможет построить одну из опор того мостика, который позволит нашим друзьям-историкам понять, какие возможности сулит им новая техника анализа.

* * *

Наука об обществе и управление обществом сегодня неразделимы. Сложность принимаемых решений требует предварительного научного анализа. Значение общественных наук непрерывно растет. Растут требования к методам анализа, которые они используют. Практическая необходимость совершенствования инструментария этих дисциплин неизбежно вызовет к жизни новые научные направления или существенно изменит структуру некоторых традиционных.

Процесс математизации (так условно назовем синтез формальных и неформальных методов исследования), как уже говорилось, довольно глубоко проник в экономику. Но уже сегодня видно, что этим дело не кончится. Действительно, любые модели, например, должны быть насыщены социологическим материалом. Недавно, занимаясь совершенно конкретной проблемой построения оптимальной иерархической системы, я столкнулся с вопросом, как соизмерять моральные стимулы с материальными благами. Для социалистического общества это очень важный вопрос. Значит, для того чтобы понять принципы централизации и децентрализации в больших управляемых системах — один из генеральных вопросов современной теории управления, — математику нужна помощь социолога, и не просто социолога, который понимает существо проблем, стоящих перед математиком, но и методы, с помощью которых он будет проектировать конструкцию.

И еще один пример. Любой процесс развивается в определенных рамках. Эти ограничения бывают разной природы. В частности, изучая, например, функционирование отрасли или предприятия, мы сталкиваемся с определенными правовыми нормами. Как выработать оптимальную систему правовых норм, регламентирующих

функционирование современных предприятий, права и обязанности руководителей разных рангов и т. д.? Как связать эту систему с качеством функционирования предприятия? По-видимому, единственный путь решения подобных проблем — использование моделей, имитирующих реальный процесс производства. В самом деле, единственной альтернативой мог бы быть эксперимент. Но он дорог и затрагивает судьбы людей. А здесь мы можем провести на машине, причем многократно, анализ различных вариантов, которые предложат эксперты — юристы и экономисты, истинные знатоки своего дела. Сложность ситуации, огромный объем фактов, которые надо учитывать, не позволяет без модели сразу же увидеть результаты их решений — увидеть, как отразится изменение правовых норм на производственном процессе. А модель может это сделать, но она, в свою очередь, беспомощна в выработке решений, что остается прерогативой специалистов. Таким образом, не только социологи, но и юристы оказываются для нас, математиков, совершенно необходимыми партнерами.

Машинная имитация реальных процессов — это не просто новый и мощный инструмент научного анализа. Это инструмент совершенно нового типа, требующий совместных усилий естественников, с одной стороны, и специалистов в области гуманитарных наук, с другой; инструмент, который, как я убежден, позволит возникнуть на стыке естественных и гуманитарных наук новым научным направлениям.

* * *

Эту статью я начал с обсуждения причин, порождающих различие во взглядах, образе мышления и интересах у людей разных интеллектуальных профессий, и прежде всего — гуманитариев и естественников. Я позволил себе не согласиться с Ч.П. Сноу, который видит главную причину в дефектах традиционного образования. Но в чем я абсолютно согласен с ним — это в актуальности самой проблемы образования.

То, о чем писалось в этой статье, не входит сегодня в университетские курсы. Разрыв между гуманитарным и естественным образованием мешает сегодня же начать широкие исследования того типа, о которых говорилось в статье.

Проблема образования чересчур сложна, чтобы о ней можно было говорить между прочим. Но все же в заключение этой

статьи мне хотелось бы сделать два замечания, касающихся образования.

Естественники должны получать значительно более глубокое и более целеустремленное образование в области общественных наук. Это прежде всего касается курса политической экономии. Нам, естественникам, необходима демонстрация прямой преемственности общетеоретических положений этой науки и конкретных экономических приложений. Великие основоположники научной политической экономии были не только великими мыслителями, которых занимали абстрактные категории. Они были также великими учителями действия: наука, которую они создали, стала основой практической рецептуры. В этом и сила, и революционность марксистской политической экономии.

Сегодня тысячи и тысячи математиков и инженеров привлечены к созданию автоматизированных систем управления, т. е. к проблемам социального и экономического планирования, к изучению процессов, происходящих в обществе. Они сталкиваются со многими вопросами конкретной экономики, и — знаю по собственному опыту — многим и не приходит в голову, что о некоторых трудных вопросах уже говорилось в курсе политической экономии.

Маленький пример. Мы изучаем теорию стоимости. Нам говорят об абстрактных категориях, и мы их забываем. Такова судьба всех абстрактных категорий в головах практически мыслящих инженеров. И никогда нам не говорят о том, что стоимость товара может быть измерена числом! Именно числом! Трудно переоценить, как важно математику и инженеру понять, что человеческий труд, заложенный в данной вещи, может быть измерен, что абстрактный, казалось бы, закон стоимости может лечь и ложится в основу планирования. Очень важно правильно расставить акценты. Если бы мы все это знали, когда кончали университет!

Надо также научиться готовить специалистов в области общественных наук для эффективной работы бок о бок с математиками и инженерами. Сейчас делаются некоторые попытки найти решение. В программы экономических факультетов, например, включают небольшие курсы математики и т. д. Но мне кажется, что должен быть другой подход. В основе естественнонаучной «цивилизации» лежит метод моделей — модельного описания явлений. Гуманитариев надо учить физике, но не так, как учат сегодня физику на естественных и инженерных факультетах. Здесь не важны детали,

которые занимают основное место в традиционном преподавании физики. Физика как система моделей — вот, вероятно, чего не хватает сегодня специалистам в области общественных наук.

Это первые и необходимые мероприятия. Но их явно недостаточно для решения тех проблем, о которых говорилось в статье.

За последнее время появились факультеты и специальности химической физики, физической химии, молекулярной биологии и т. д., которые объединили физиков и химиков, физиков и биологов. Вероятно, настало время говорить о факультетах, где можно объединить физиков и экономистов, математиков и историков, юристов и т. д.

1973 г.

К теории оптимального управления на бесконечном интервале времени

1. Рассматривается задача отыскания функций $x(t)$ и $u(t)$, связанных соотношением

$$\dot{x} = f(x, u, t), \quad (x, u) \in G_u, \quad (1.1)$$

и доставляющих минимум функционалу

$$I = \int_{t_0}^T F(x, u, t) dt. \quad (1.2)$$

Подобные задачи имеют ряд интересных особенностей. Например, может оказаться, что замена интеграла (1.2) интегралом

$$I_T = \int_{t_0}^T F(x, u, t) dt$$

приводит к задаче, неразрешимой в классе конечных управлений ни для какого конечного T . Кроме того, задачи типа (1.1), (1.2) встречаются в разнообразных приложениях.

Приведем два примера.

а) Пусть $f(0, 0, t) = 0$. Будем говорить, что тривиальное решение уравнения (1.1) I -устойчиво, если $0 \leq T_0 < t_0$ и $x(t_0) = x_0$, удовлетворяющего условию $\|x_0\| \leq \epsilon$, можно указать такое $u(x_0, t) \in G_u$, которое переводит систему из состояния x_0 в начало координат за любое время $T > T_0$, причем так, что положительный функционал

$$I_T = \int_{t_0}^T F(x, u, t) dt \rightarrow 0.$$

Для I -устойчивости необходимо, чтобы для $\|x_0\| \leq \epsilon$ решение вариационной задачи (1.1), (1.2) существовало и

$$I = \int_{t_0}^{\infty} F(x, u, t) dt \rightarrow 0.$$

б) В теории экономического роста рассматривается задача отыскания $u(t)$, удовлетворяющего условию

$$0 \leq u(t) \leq x \quad (1.3)$$

и доставляющего минимум функционалу

$$I = \int_{t_0}^{\infty} (c, x) p(t) dt. \quad (1.4)$$

Функция $p(t) \geq 0$ при $t \geq t_0$ называется функцией дисконтирования.

2. Построение необходимых условий типа принципа максимума встречает трудности в связи с отсутствием априорной информации о поведении множителей Лагранжа в окрестности бесконечно удаленной точки и в связи с построением конусов допустимых вариаций на бесконечности. Кроме того, с вычислительной точки зрения необходимые условия было бы трудно использовать непосредственно, поскольку их проверка требовала бы решения краевой задачи на бесконечном интервале времени. Для эффективного решения подобной задачи было бы необходимо определить процедуру переноса граничных условий из бесконечно удаленной точки в конечную точку, а следовательно, построить некоторую асимптотику искомых функций.

В силу сказанного, кажется целесообразным вместо попытки отыскать необходимые условия попытаться найти управление, которое, будучи допустимым, дает асимптотическое представление (в силу естественной метрики, задаваемой функционалом (1.2)) оптимального управления.

Для построения асимптотики необходимо условиться о характере поведения управления в окрестности бесконечно удаленной точки. Будем говорить, что $u(t)$ принадлежит классу G^N , если при $t \rightarrow \infty$

$$u(t) = \frac{u_1}{t} + \frac{u_N}{t^N} + o\left(\frac{1}{t^N}\right),$$

где u_i — некоторые векторы. В дальнейшем мы ограничимся рассмотрением только управлений из класса G^N .

3. Мы будем использовать две теоремы, принадлежащие Хукухара [1]. Пусть дано линейное дифференциальное уравнение

$$\dot{x} = A(t)x + g(t), \quad (3.1)$$

где матрица $A(t)$ и вектор $g(t)$ имеют представления

$$A = A_0 + \frac{A_1}{t} + \frac{A_N}{t^N} + o\left(\frac{1}{t^N}\right), \quad g = \frac{g_1}{t} + \frac{g_N}{t^N} + o\left(\frac{1}{t^N}\right). \quad (3.2)$$

Справедливы следующие теоремы.

Теорема 1. Если все собственные значения матрицы A_0 попарно различны, то любое частное решение $y(s, t)$ однородного уравнения

$$\dot{x} = A(t)x, \quad (3.3)$$

соответствующее собственному значению λ_s , представимо в виде

$$y(s, t) = \exp\left\{\lambda_s t + o(\ln t)\right\} \{x_0(s) + x_1(s)t^{-1} + \dots + x_N(s)t^{-N} + o(t^{-N})\}. \quad (3.4)$$

Коэффициенты разложения — векторы $x_i(s)$ определяются подстановкой выражения (3.4) в (3.1) и сравнением коэффициентов при одинаковых степенях t .

Теорема 2. Если матрица $A(t)$ и вектор $g(t)$ удовлетворяют условиям (3.2), то уравнение (3.1) имеет, и притом единственное, частное решение, представимое с помощью следующего асимптотического разложения:

$$\tilde{x}(t) = \frac{\tilde{x}_1}{t} + \frac{\tilde{x}_N}{t^N} + o\left(\frac{1}{t^N}\right). \quad (3.5)$$

Коэффициенты разложения определяются формулами

$$\begin{aligned} A_0 \tilde{x}_1 &= g_1, & A_0 \tilde{x}_2 &= g_2 - (A_1 - E)\tilde{x}_1, \\ A_0 \tilde{x}_3 &= g_3 - (A_1 - 2E)\tilde{x}_2 - A_2 \tilde{x}_1, \end{aligned} \quad (3.6)$$

Заметим, что формулы (3.6) позволяют вычислить g_i по известным \tilde{x}_i .

Приведенные факты распространяются на нелинейные уравнения вида

$$\dot{x} = Ax + g(x, t), \quad (3.7)$$

где A и g — удовлетворяют условиям (3.2), а вектор-функция $g(x, t)$ имеет вид

$$\frac{g_1(x)}{t} + \frac{g_2(x)}{t^2}, \quad (3.8)$$

причем g_i для достаточно малых $\|x\|$ разлагаются в ряд Тейлора

$$g_i(x) = \Gamma_{i2}(x) + \Gamma_{i3}(x) + \dots,$$

где Γ_{ij} — однородные операторы порядка j .

4. Рассмотрим задачу (1.1), (1.2) при дополнительных условиях

$$x(t_0) = x_0, \quad x(\infty) = 0. \quad (4.1)$$

В силу условия (4.1), естественно предположить, что для достаточно больших t уравнение (1.1) имеет вид

$$\dot{x} = Ax + Bu + v(x), \quad (4.2)$$

где функция $v(x)$ представима в виде (3.8).

Точно так же естественно рассмотреть класс задач, в которых подынтегральная функция $F(x, u, t)$ представима в виде

$$F(x, u, t) = (x, R_{11}x) + (x, R_{12}u) + (u, R_{22}u) + o(\max\{\|x\|^2; \|u\|^2\}) \\ F_2(x, u, t) + o(\dots), \quad (4.3)$$

где $F_2(x, u, t)$ — неотрицательная функция x, u и $F_2(0, 0, t) = 0$. Предположению (4.3) удовлетворяют все функции, с помощью которых обычно оценивается качество управления.

Кроме того мы допустим, что управление принадлежит к асимптотическому классу G^N , а матрица A имеет представление (3.2).

Замечание 1. Эти последние предположения с прикладной точки зрения более ограничительны, поскольку они не охватывают классы задач, в которых управление носит релейный характер, а матрица A имеет существенную особенность на $\lambda = 0$.

Обозначим

$$S(t_1, x_1) = \min_{t_1} \int_{t_1} F(x, u, t) dt$$

при условиях (1.1) и (1.4) и

$$S(t_1, x_1; t_2, x_2) = \min_{t_1} \int_{t_1}^{t_2} F(x, u, t) dt$$

при условиях (1.1) и $x(t_1) = x_1, x(t_2) = x_2$. Таким образом,

$$S(t_0, x_0) = \min_{x_T} \{S(t_0, x_0; T, x_T) + S(T, x_T)\} \quad T \geq t_0. \quad (4.4)$$

Решение этой задачи обозначим через x^*, u^* .

Пусть теперь \hat{u} — некоторое допустимое управление, которое переводит систему из состояния (x_T, T) в начало координат, так что интеграл

$$R(x_T) = \int_T F(\hat{x}, \hat{u}, t) dt$$

конечен. Поскольку управление задано, то этот интеграл будет функцией только x_T .

Вместо (4.4) рассмотрим соотношение

$$(t_0, x_0) = \min_{x_T} \{S(t_0, x_0; T, x_T) + R(x_T)\}. \quad (4.5)$$

Значение x_T , реализующее этот минимум, обозначим через \hat{x}_T . Определение (t_0, x_0) сводится к следующей задаче Больца: найти управление $u(t) = u^{**}(t)$, доставляющее минимум функционалу

$$I_* = \int_{t_0}^T F(x, u, t) dt + R(x_T) \quad (4.6)$$

при ограничении (1.1).

Если $|R(x_T) + S(T, x_T)| = 0$ при $T = t_0$, то управление $U(t)$, заданное с помощью соотношений

$$U = \begin{cases} u^{**}(T), & t \in [t_0, T], \\ u(t), & t \in (T, \infty), \end{cases}$$

будет аппроксимировать (слабо, в смысле функционала I) решение исходной вариационной задачи u^* .

Замечание 2. При решении вариационной задачи (4.6) необходимо на выбор x_T наложить ограничение вида

$$x_T \in R(x_T), \quad (4.8)$$

где $R(x_T)$ — множество точек, из которого достижимо конечное состояние. Однако условие (4.8) при реализации численного расчета будет выполнено автоматически, поскольку если $x_T \in \bar{E}$, то $R(x_T) \neq \emptyset$.

5. Реализация обсуждаемой схемы требует построения допустимого управления \hat{u} и решения задачи со свободным концом (4.6). Вторая из этих задач является стандартной задачей оптимального управления. Построение \hat{u} — это самостоятельная и трудная задача, и в каждом случае она требует специального анализа. Для системы (4.2) и при условии, что $u \in G^N$, определение \hat{u} сводится к задаче нелинейного программирования.

Рассмотрим вспомогательную задачу определения минимума

$$\hat{I} = \int_T F(x, u, t) dt, \quad (5.1)$$

где x и u связаны уравнением (4.2) и удовлетворяют ограничениям

$$x(T) = x_T, \quad x(0) = 0. \quad (5.2)$$

Предположим, что управления представимы в виде

$$u = \frac{u_1}{t} + \frac{u_N}{t^N} + o\left(\frac{1}{t^N}\right). \quad (5.3)$$

Согласно теореме Хукухара, общее решение уравнения (4.1) для больших t представимо в виде

$$x(t) = C_s y(s, t) + \tilde{x} + o\left(\frac{1}{t^N}\right) = X(C_1, \dots, C_k, u_1, \dots, u_N, t) + o\left(\frac{1}{t^N}\right); \quad (5.4)$$

здесь $y(s, t)$ определяются формулами (3.4), а \tilde{x} — формулой (3.5), где $g_i = Bu_i$, \tilde{x} — произвольные постоянные.

Поскольку $x(0) = 0$, то сумма в выражении (5.4) должна быть распространена лишь на те индексы s , $s = 1, 2, \dots, k$, которым соответствуют

собственные значения λ_s , обладающие отрицательными действительными частями.

Из (5.4) следует, что

$$x_T = X(C_1, \dots, C_k, u_1, \dots, u_N, T) \approx \frac{1}{T^N}. \quad (5.5)$$

Равенство (5.5) определяет в E_n некоторое многообразие — множество достижимости.

Подставляя выражение (5.4) в (5.1) и проводя интегрирование, получим

$$\hat{I} = Z(C_1, \dots, C_k, u_1, \dots, u_N) \approx \frac{1}{T^{N-1}}. \quad (5.6)$$

Найдем теперь минимум \hat{I} при условии (5.5):

$$S(T, x_T) = \min Z(C_1, \dots, C_k, u_1, \dots, u_N) \approx \frac{1}{T^{N-1}}.$$

Из приведенных рассуждений сразу следует способ приближенного определения величины $S(T, x_T)$, которую мы обозначим через $R(T, x_T)$:

$$R(T, x_T) = \min Z(C_1, \dots, C_k, u_1, \dots, u_N) \quad (5.7)$$

при условии

$$x_T = X(C_1, \dots, C_k, u_1, \dots, u_N, T). \quad (5.8)$$

Задача (5.7), (5.8) — это задача нелинейного программирования. Очевидно, что

$$S(T, x_T) = R(x_T) \approx \frac{1}{T^{N-1}}.$$

Управление, реализующее значение $R(x_T)$, обозначим, следуя п. 4, через \hat{u} .

6. Рассмотрим снова величину $S(t_0, x_0)$. Ее можно представить в виде

$$S(t_0, x_0) = \min_{x_T \in X(t_0, x_0; T, x_T)} \{S(t_0, x_0; T, x_T) = S(T, x_T)\}.$$

Введем еще величину

$$S_1(t_0, x_0) = \min_{x_T \in X} \{S(t_0, x_0; T, x_T) - S(T, x_T)\}.$$

Если функция Беллмана непрерывна в окрестности оптимального решения, то

$$S(t_0, x_0) - S_1(t_0, x_0) = o\left(\frac{1}{T^N}\right). \quad (6.1)$$

И, далее,

$$S_1(t_0, x_0) = \min_{x_T \in X} \{S(t_0, x_0; T, x_T) - R(x_T)\} = o\left(\frac{1}{T^{N-1}}\right) \quad (6.2)$$

$$(t_0, x_0) = o\left(\frac{1}{T^{N-1}}\right),$$

где (t_0, x_0) определяется формулой (4.5).

Сопоставляя (6.1) и (6.2), получим

$$S_1(t_0, x_0) - (t_0, x_0) = o\left(\frac{1}{T^{N-1}}\right).$$

Таким образом, имеет место

Теорема 3. Если функция Беллмана в задаче отыскания минимума функционала (1.2) при условиях (4.1)—(4.3) для достаточно больших T в окрестности оптимальной траектории непрерывна, то управление U , определенное формулой (4.7), будет асимптотическим представлением оптимального управления $u^* \in G^N$ в том смысле, что

$$\int_{t_0}^{t_0} F(x^*, u^*, t) dt - \int_{t_0}^{t_0} F(x(U), U, t) dt = o\left(\frac{1}{T^{N-1}}\right),$$

где $x(U)$ — фазовая траектория, определяемая управлением U .

Замечание 3. Условие непрерывности функции Беллмана является существенным лишь в том случае, если множество достижимости не совпадает со всем пространством.

Итак, приближенным, асимптотическим решением исходной задачи с бесконечным временем является функция, реализующая минимум функционала

$$I_1 = \int_{t_0}^T F(x, u, t) dt - R(x_T).$$

Это задача со свободным концом, и для нее может быть сформулирован принцип максимума. Функция Гамильтона будет при этом иметь вид

$$H(x, u, t) = F^*(x, u, t),$$

где

$$F^*(x, u, t) = F(x, u, t) - \frac{dR}{dx} f,$$

а импульсы должны обращаться в нуль при $t = T$.

7. В качестве примера рассмотрим тот частный случай, когда уравнение (1.1) линейное:

$$\dot{x} = Ax + Bu,$$

все собственные значения матрицы A_0 из (3.2) имеют положительные действительные части, а матрица B постоянная и имеет обратную. В качестве функционала рассмотрим следующий:

$$I = \int_{t_0}^T (u, u) dt.$$

Согласно теореме Хукухара, единственное ограниченное решение уравнения (7.1) имеет вид

$$x = \frac{x_1}{t} + \frac{x_N}{t^N} + o\left(\frac{1}{t^N}\right),$$

где числа x_i определены формулами (3.6):

$$A_0 x_1 = Bu_1, \quad A_0 x_2 = Bu_2 - (A_1 - E)x_1,$$

Эти формулы позволяют определить величины u_i единственным образом через x_1, x_2, \dots . Они будут линейными функциями чисел x_1, \dots, x_N . Таким образом, задача определения управления $u = \hat{u}$ сведется к некоторой задаче квадратичного программирования: отыскать числа x_1, \dots, x_N , доставляющие минимум квадратичной форме

$$I(x_1, \dots, x_N) = \int_T (u, u) dt$$

при условии $x_T = x_1/T \dots x_N/T^N$. На величину x_T здесь никаких ограничений не накладывается.

8. Условие существования B^{-1} весьма сильное. В прикладных задачах, как правило, управление имеет меньшую размерность и входит не во все уравнения системы. Конечно, этот случай также вкладывается в общую схему п. 6. Тем не менее некоторые особенности имеет смысл пояснить на примере:

$$\dot{x}^{(1)} = x^{(1)} - \frac{1}{2}x^{(2)}, \quad \dot{x}^{(2)} = x^{(2)} - \frac{1}{2}x^{(1)} + u. \quad (8.1)$$

Оба характеристических корня положительные:

$$\lambda_{1,2} = 1 \pm \sqrt{3/2}.$$

Следовательно, отличие рассматриваемого примера от того, который был рассмотрен в п. 7, состоит только в том, что в системе (8.1) матрица B вырожденная и B^{-1} не существует.

Следуя общей схеме, ограниченное частное решение мы будем искать в виде

$$x^{(1)} = \frac{x_1^{(1)}}{t} \dots \frac{x_N^{(1)}}{t^N} + o\left(\frac{1}{t^N}\right), \quad x^{(2)} = \frac{x_1^{(2)}}{t} \dots \frac{x_N^{(2)}}{t^N} + o\left(\frac{1}{t^N}\right).$$

Связь между величинами u_i и $x_j^{(1)}$ будет следующей:

$$u_1 = x_1^{(2)} - \frac{x_1^{(1)}}{2}, \quad x_1^{(2)} = 2x_1^{(1)},$$

$$u_2 = x_2^{(2)} - x_1^{(2)} - \frac{x_2^{(1)}}{2}, \quad x_2^{(2)} = 2x_2^{(1)} - x_1^{(1)},$$

$$u_3 = x_3^{(2)} - 2x_2^{(2)} - \frac{x_3^{(1)}}{2}, \quad x_3^{(2)} = 2x_3^{(1)} - 4x_1^{(2)},$$

и т. д. Таким образом, не только величины u_i , определяющие управления, но и величины $x_i^{(2)}$ будут функциями $x_i^{(1)}$:

$$u_i = V_i(x_1^{(1)}, \dots, x_i^{(1)}), \quad x_i^{(2)} = X_i^{(2)}(x_1^{(1)}, \dots, x_i^{(1)}).$$

Следовательно, и функционал будет зависеть только от величины $x_i^{(1)}$. Задача отыскания управления \hat{u} снова будет задачей нелинейного программирования с ограничением типа равенства, которое в данном случае имеет вид

$$x_T^{(1)} = \frac{x_1^{(1)}}{T} \dots \frac{x_N^{(1)}}{T^N}.$$

Величина $x_T^{(2)}$ при этом не может быть задана произвольно. Она определится однозначно:

$$x_T^{(2)} = \frac{x_1^{(2)}}{T} \dots \frac{x_N^{(2)}}{T^N}.$$

Эта особенность рассматриваемой задачи — следствие того факта, что ограниченные решения системы (8.1) могут быть порождены только некоторым одномерным многообразием.

9. Рассмотрим теперь задачу экономического роста. Она состоит в отыскании минимума функционала (1.4) при ограничениях (1.1) и (1.3). Ограничения на правый конец траектории отсутствуют, и компоненты вектора x в общем случае неограниченно возрастают.

Условимся рассматривать тот случай (который обычно изучается в экономике), когда дифференциальные связи линейные

$$\dot{x} = A(t)x + B(t)u \tag{9.1}$$

и ограничения (1.3) имеют вид $0 \leq u_i \leq k_i x_i$, где $k_i \in [0, 1]$.

Условимся, наконец, что в окрестности бесконечно удаленной точки матрицы A и B имеют представления (3.2).

Вместо управления $u(t)$ введем новое $u = x$, где Λ — диагональная матрица, элементы которой удовлетворяют условиям

$$0 \leq \Lambda_i \leq k_i. \tag{9.2}$$

Уравнение (9.1) мы можем теперь представить в виде

$$\dot{x} = (A + B\Lambda)x. \tag{9.3}$$

Таким образом, еще одна из особенностей этой задачи состоит в том, что дифференциальная связь — это однородное уравнение. Изменение управления непосредственно влияет на величину характеристических показателей.

Условимся изучать класс управлений, обладающих следующей асимптотикой:

$$o \frac{1}{t} \dots \frac{1}{t^N} o \frac{1}{t^N} .$$

Более того, матрицу A_0 мы будем считать известной.

Замечание 4. Подобное предположение тесно связано с теорией магистралей: для достаточно больших t развитие экономики стремится иметь максимальный темп роста в каждый данный момент:

$$(c, x) \rightarrow \max.$$

При $t \rightarrow \infty$ это условие дает

$$(c, B_0^{-1} x) \rightarrow \max$$

и матрица B_0^{-1} определяется из решения задачи линейного программирования.

Предположив, что собственные значения матрицы $A_0 - B_0^{-1} c c^T$ не кратные, мы можем снова применить теорему Хукухара и представить общее решение уравнения (9.3) в форме

$$x = \sum_{s=1}^n C_s y(s, t),$$

где $y(s, t)$ имеют представления (3.4) и будут явно выражены через элементы матриц C_1, \dots, C_n , т. е. величины c_{ik} .

Следовательно, и функционал будет функцией этих величин:

$$I = \int_{t_0}^T C_s y(s, t) p(t) dt = I(C_1, \dots, C_n, c_{ik}). \quad (9.4)$$

Дальнейшая схема вычислений вполне очевидна. Отбрасывая члены порядка $o(1/t^N)$, мы определим допустимое управление \hat{u} как решение следующей задачи нелинейного программирования: определить

величины C_1, \dots, C_n и элементы ik , доставляющие минимум функции (9.4) при условиях (9.2) и условии типа равенства

$$x_T = \int_s C_s y(s, t). \quad (9.5)$$

Для рассматриваемого случая также может быть сформулирована теорема, аналогичная той, которая сформулирована в п. 6. Более того, так как условие (9.5) не накладывает никаких ограничений на выбор x_T (задача со свободным концом и любая точка x_T принадлежит множеству достижимости), то мы можем не требовать непрерывности функции Беллмана.

10. В основу данной работы была положена идея предварительной «асимптотической обработки» исходной задачи — ее замена некоторой более простой задачей с конечным временем. Идея такого предварительного анализа неоднократно использовалась в работах Вычислительного центра АН СССР. В частности, на ней были основаны численные методы расчета траекторий космических объектов, позволившие проводить численное интегрирование с большим шагом по независимой переменной [2], методы расчета колебаний вязкой жидкости [3] и т. д.

Такой подход к решению задач с особенностями позволяет, как правило, значительно упростить процедуру получения численных решений. В случае рассмотренных оптимизационных задач он позволил, в частности, избежать трудной вспомогательной задачи отыскания необходимых условий типа принципа максимума.

Конечно, предлагаемая схема не дает еще общей теории. Мы рассмотрели лишь некоторые частные случаи, и предлагаемая форма анализа не может быть непосредственно перенесена на общий класс задач, в которых управления обладают ограниченной энергетической нормой.

Литература

1. *Hukuhara M.* Sur les points singuliers des equations differentielles lineaires. J. Fac. Sci. Hokkaido Univ. Ser. I, 1934, 2, 13–88.
2. *Моисеев Н.Н.* Асимптотические методы нелинейной механики. — М., Наука, 1969.
3. *Моисеев Н.Н., Румянцев В.В.* Динамика твердого тела, содержащего жидкие массы. — М., Наука, 1967.

1974 г.

Оптимизация и управление (эволюция идей и перспективы)

Идеи оптимизации пронизывают все разделы теории управления, являясь по существу основой анализа и проектирования (синтеза) систем управления объектами самой различной природы. Сегодня опубликовано уже много первоклассных обзоров, содержащих детальный анализ основных результатов и концепций использования идей оптимизации и подробное обсуждение места и роли отдельных исследований. Эти обзоры снабжены к тому же обширной библиографией, и читателю, который нуждается в подобной информации, можно рекомендовать, например, обзор [1].

В настоящей статье рассказывается о том, каким образом идеи оптимизации сделали одним из основных инструментов развития теории управления, как в свою очередь теория управления оказала решающее влияние не только на развитие математических теорий, но и трансформировала само содержание понятия «оптимальность» в управлении. Настоящая работа, конечно, никак не претендует на полноту изложения и отражает лишь те субъективные представления, которые сформировались по этому вопросу у автора.

1. Возникновение теории управления обычно связывают с именами Максвелла и Вышнеградского. И для этого есть достаточно веские причины. Сам термин «управление» имеет смысл тогда, когда имеется цель, к которой мы стремимся, и определенные возможности (ресурс), которые мы можем использовать для ее достижения. К середине XIX в. уже были созданы простейшие регуляторы, и первые задачи возникающей теории состояли в том, чтобы так распорядиться, так выбрать параметры этих регуляторов (т. е. так его спроектировать), чтобы обеспечить устойчивость заданного стационарного режима работы машины. В работах Максвелла и Вышнеградского четко сформулирована цель управления — обеспечение устойчивости и предложены методы, позволяющие выбрать параметры системы управления. Затем на протяжении почти ста лет основные проблемы теории управления так

или иначе оказывались связанными с проблемами устойчивости. В 1920-е, а особенно в 1930-е годы в теории регулирования (тогдашний синоним теории управления) широкое распространение получили идеи и методы ляпуновской устойчивости, и в свою очередь теория регулирования непрерывно индуцировала интерес математиков к задачам устойчивости.

Значение идей оптимизации в теории управления легко объяснить. Методы теории управления должны давать возможность определять параметры (или функции), которые обеспечивают достижение определенной цели. А задача достижения любой цели (устойчивости, качества и т. д.) всегда может быть сформулирована на «оптимизационном» языке. Покажем это на примере проблемы устойчивости.

Задача формулировалась следующим образом. Рассматривается динамическая система

$$\dot{x} = f(x, t, c_1, \dots, c_k),$$

где c_1, c_2, \dots, c_k — некоторые скалярные параметры. Это — свободные параметры, или, как теперь говорят, управления. Если речь идет об автопилоте, то c_1, \dots, c_k — параметры, характеризующие обратную связь. Задача состоит в том, чтобы выбрать их таким образом, чтобы некоторое частное решение системы

$$\dot{x}^0 = f(x^0, t, c_1, \dots, c_k) \quad (1.1)$$

было устойчивым в смысле Ляпунова. Как правило, рассматривалось некоторое стационарное решение системы (1.1).

Введем теперь функцию этих параметров

$$J_0(c_1, \dots, c_k) = \begin{cases} 1, & \text{если траектория устойчива,} \\ 0, & \text{если траектория неустойчива.} \end{cases} \quad (1.2)$$

Тогда задача устойчивости может трактоваться как задача отыскания чисел c_1, \dots, c_k , доставляющих максимум функции (1.2). Если c_i стеснены условиями вида

$$(c_1, c_2, \dots, c_k) \in G_s(s = 1, 2, \dots, r), \quad (1.3)$$

где G_s — некоторые заданные множества в пространстве параметров, то задача отыскания максимума функции (1.2) будет некоторой специальной задачей нелинейного программирования.

2. Такая переформулировка задачи устойчивости сама по себе еще не является источником каких-либо новых методов. Она просто иллюстрирует тот почти очевидный факт, что практически любая задача теории управления — это задача отыскания некоторого решения, оптимального в том или ином смысле.

Задача (1.2), (1.3) не выделяет единственного решения. Условия, которым должны удовлетворять параметры системы для того, чтобы обеспечить ее устойчивость, определяют некоторое множество $\cap G_s$. Это означает, что конструктор, проектирующий систему, имеет еще достаточно произвола: обеспечив устойчивость, он может еще удовлетворить целому ряду дополнительных требований. Так рождается задача о «качестве управления» устойчивым движением. Как правило, эти дополнительные требования накладываются на характер переходного процесса. Их также легко сформулировать на языке оптимизации подобно задаче устойчивости. Таким образом, задача проектирования системы управления может быть сформулирована как некоторая многокритериальная задача: отыскать числа c_i , удовлетворяющие условиям (1.3) и доставляющие максимум (минимум) функциям $J_i(c_1, c_2, \dots, c_k)$ ($i = 0, 1, \dots, N$). В общем случае эта задача не имеет решения и проблема выбора c_i — проблема отыскания некоторого компромисса. Однако есть один специальный случай, когда проблема многокритериальной (или, как говорят, векторной) оптимизации сводится к задаче нелинейного программирования.

Обозначим через Ω_i множество значений параметров, доставляющих минимум функционалу J_i , и предположим, что $\bigcap_{i=0}^N \Omega_i \neq \emptyset$.

Тогда очевидно, что рассматриваемая задача векторной оптимизации сводится к последовательности задач нелинейного программирования

$$J_i \rightarrow \min, \quad (c_1, \dots, c_k) \in \bigcap_{j=1}^i \Omega_j. \quad (2.1)$$

В предвоенные годы появилось довольно много работ, посвященных качеству управления устойчивыми режимами. Все они допускают формулировку в терминах задачи (2.1).

Примечание. Ограничения также можно включить в число функционалов и оперировать только терминами теории векторной оптимизации. Однако для дальнейшего изложения подобные обобщения нам будут ненужны.

3. В послевоенные годы язык теории оптимизации во все большей степени начинает использоваться в теории управления. Для этого много причин, главная из которых — расширение фронта исследования, появление новых объектов, изучением которых начинает заниматься теория управления. Если в предвоенные годы изучались прежде всего стационарные режимы, устойчивость равномерного полета самолета, стабилизация оборотов турбины и т. д., то в послевоенные годы запросы техники требуют все большей степени анализа нестационарных процессов.

Одним из основных источников новых задач оказалась на первых порах ракетная техника. Управление объектами, перегрузка которых достигала десятков g , разумеется, никак не вкладывалось в традиционные рамки теории регулирования. Конечно, еще и сейчас в инженерной практике используются классические методы этой теории. Однако без анализа, который использует современные методы, никакие действительно сложные задачи проектирования не могут быть решены. В самом деле, весь активный участок ракеты, т. е. те немногие минуты, в течение которых ракета разгоняется до космической скорости, — это некоторый переходный процесс. И исследование ляпуновской устойчивости, конечно, оказывается здесь недостаточным для проектирования управления ракетой.

Место задачи устойчивости как основной задачи теории управления начинают занимать задачи отыскания оптимального управления. Ракетная техника была первой и, может быть, главной причиной создания и развития теории оптимального управления, новой главы общей теории управления. Этой теории было суждено превратиться в теорию управления из некоторого раздела механики (какой была теория регулирования) в обширную область человеческих знаний. Ее проблематика сегодня связана не только с техникой, экономикой, теорией социальных процессов, но и с любой областью человеческой деятельности, где имеет смысл использовать теорию принятия решений, одним из основных разделов которой и стала теория управления.

Началось все с небольшой работы Д.Е. Охоцимского [2], содержавшей решение первой задачи теории оптимального управления, которая ныне хорошо известна специалистам. Именно она открыла первую страницу новой теории.

Еще до войны теория управления уже имела дело с оптимизационными задачами. Однако тогда речь шла о минимизации функции, об отыскании параметров, доставляющих характеристике качества максимальное значение. Теперь же речь шла об отыскании функции, доставлявшей экстремум некоторому функционалу

$$J[u(t)] \min, \dot{x} \quad f(x, u, t). \quad (3.1)$$

Формально задача (3.1) напоминает постановку классической задачи вариационного исчисления. Однако она содержит одну особенность, исключаящую возможность непосредственно использовать аппарат вариационного исчисления. Дело в том, что технически правильно поставленные задачи всегда содержат ограничения на управление. Так, например, в работе Д.Е. Охоцимского управление — величина тяги или расход массы — естественно, не могло быть отрицательным или превосходить некоторую константу. Таким образом, к условиям (3.1) должно быть добавлено условие $u(y) \in G$, где G некоторое множество, которое может быть и замкнутым. Таким образом, вместе с задачами минимизации расхода топлива получен новый класс математических задач, требующий развития специфических методов.

4. На первых порах какие-либо специальные методы анализа задач теории оптимального управления отсутствовали. В [2] и потоке работ, который возник после нее в конце 1940-х годов в СССР и за рубежом, решение задачи опиралось на прямой анализ вариации функционала. Это означает, что каждая задача в то время требовала изобретения специальных приемов исследования. Общий формализм теории оптимального управления, получивший название принципа максимума, создан Л.С. Понтрягиным и его учениками В.Г. Болтянским, Р.В. Гамкрелидзе и Е.Ф. Мищенко в первой половине 1950-х годов [3]. Принцип максимума позволил редуцировать задачу отыскания оптимального управления к некоторой специальной образом поставленной краевой задаче.

Приведем ее формулировку. Пусть речь идет об отыскании минимума функционала

$$J[u(t)] = \int_{t_0}^T F(x, u, t) dt \quad (4.1)$$

при дифференциальных связях (3.1) и ограничении $u(t) \in G(t)$. Введем функцию $H(x, u, t, \lambda) = F(x, u, t) + \lambda^T (x - \dot{x})$. Здесь $\lambda(t)$ — вектор множителей Лагранжа. Пусть он удовлетворяет системе уравнений

$$\dot{\lambda} = -H_x(x, u, t, \lambda) - F_x(x, u, t). \quad (4.2)$$

Принцип максимума утверждает, что оптимальное управление для любого t должно быть решением задачи

$$u(t) \in G(t): H(x, u, t, \lambda) \rightarrow \max. \quad (4.3)$$

Сам по себе этот результат имеет выдающееся значение для математики. В самом деле, в классическом вариационном исчислении, которое изучает задачу (4.1) при условии, что множество G совпадает со всем пространством, т. е. задачу качественно более простую, утверждается только стационарность функции Гамильтона. Таким образом, Л.С. Понтрягину удалось в более сложной задаче получить результат значительно более сильный, чем аналогичный результат в классическом вариационном исчислении. Этот факт уже сам по себе замечателен — получить в середине XX в. новый результат в классическом вариационном исчислении — дисциплине, история которой насчитывает уже почти 300 лет, где работали Л. Эйлер, Ж.-Л. Лагранж, К.Г. Якоби, К. Вейерштрасс и десятки других выдающихся математиков, — достижение уникального научного значения*.

Результаты Л.С. Понтрягина имели и чисто прикладное значение. В самом деле, задача (4.3) позволяет определить

$$u^*(x, t) \in G(t). \quad (4.4)$$

Подставляя (4.4) в (3.1) и (4.2), получим систему $2n$ -го порядка, которая будет иметь вид

* Справедливости ради следует заметить, что уже после опубликования принципа максимума ряд лиц, в том числе В.А. Троицкий и автор этой статьи, независимо друг от друга показали, что для задач классического вариационного исчисления принцип максимума следует (и по существу эквивалентен) из классической теории Вейерштрасса.

$$\dot{x} = (x, \dots, t), \quad \dot{y} = (x, \dots, t). \quad (4.5)$$

Условия трансверсальности вместе с условиями на левом конце фазовой траектории $x(t)$ определяют $2n$ граничных условий для системы (4.5). Таким образом, из теоремы Понтрягина следует утверждение: оптимальные траектории необходимо должны содержаться среди решений краевой задачи (4.5).

Таким образом, принцип максимума позволил унифицировать формализм решения широкого класса задач управления, которые возникают в технике.

Работы школы Л.С. Понтрягина сразу же приобрели мировую известность, и на их базе возникли многочисленные алгоритмы численного решения прикладных задач (см., например, [4], где дано описание основных способов численного решения). Конечно, теория Л.С. Понтрягина была не единственным источником идей для эффективного отыскания оптимальных траекторий. Целый ряд идей и методов привлечен в теорию управления извне, когда в конце 1950-х годов теория управления перестала ограничиваться только техническими объектами и вторглась в сферу социальных процессов и прежде всего в экономику.

Примечание. Развитие идей теории оптимального управления повлияло на целый ряд постановок задач теории управления. Одним из ярких примеров плодотворности использования языка оптимизации явилась теория инвариантности. Эта теория имеет трудную и долгую историю [5]. Применение языка теории оптимальных управлений сделало проблему инвариантности совершенно прозрачной: найти условия независимости функции Гамильтона от возмущений, которые рассматриваются как некоторые управления. Такая переформулировка сводит задачу к анализу вариации функционала и, следовательно, открывает возможность использования всего арсенала средств теории оптимального управления. Эта точка зрения предложена Л.И. Розоноэром [6]. По существу, во всяком случае с точки зрения математики, проблема инвариантности сегодня закрыта в превосходном исследовании В.В. Величенко [7]. Им предложен практически исчерпывающий набор необходимых и достаточных условий не только сильной, но и слабой инвариантности.

5. Задачи теории оптимального управления — совершенно новые элементы в теории управления. Конечно, теория управления и раньше имела дело с оптимизационными задачами. Более того, и на это мы специально обратили внимание читателя, все традиционные задачи, которыми занималась теория, могли быть сформулированы на языке оптимизации. И тем не менее содержательный смысл их был совершенно иной.

Оптимизационные задачи, которые решались в предвоенные годы, имели своей целью определить структуру и коэффициенты обратной связи (коэффициенты усиления), которые гарантировали устойчивость и требуемое качество заданного, заранее установившегося программного движения. В задачах, которые стали решать инженеры в послевоенные годы, возникла проблема отыскания самого программного движения. И теория оптимального управления — это, собственно, и есть теория оптимальных программ. Однако появление новых задач никак не уменьшало значения традиционной проблемы синтеза оператора обратной связи.

При расчете программного движения всегда используется идеализованная схема движения и делаются некоторые предположения о природе конкретной реализации возмущающих воздействий. Реальность, разумеется, бывает иной. Управляемый объект никогда не движется точно вдоль программной траектории, поэтому чтобы управляемый объект достиг цели, всегда необходима еще некоторая система управления (автомат).

Обозначим через $\tilde{x}(t)$ программную траекторию. Тогда движение управляемого объекта может быть описано системой уравнений вида

$$\dot{x} = (\tilde{x}, x, v, t), \quad (5.1)$$

где v — дополнительное (корректирующее) управление, которым располагает система управления, (t) — случайный процесс, т. е. помехи, которые не дают возможности аппарату точно следовать расчетной траектории. Задача синтеза состоит в таком выборе управления v , чтобы в каком-то смысле наилучшим образом достичь цели управления. В классических задачах теории регулирования для этого оказывалось достаточным (с технической точки зрения) обеспечить устойчивость системы. Новый тип задачи — управление оптимальной траекторией на конечном интервале времени — предъявляет и другие требования. Приведем типичный пример.

Пусть речь идет о выводе космического аппарата на круговую орбиту. Предполагая помехи отсутствующими, мы определили программную траекторию, которая выводит нашу ракету на расчетную орбиту с минимальной затратой топлива. Задача автопилота — так распорядиться корректирующим управлением (и его ресурсом), чтобы наиболее точно были выполнены терминальные условия, т. е. условия выхода аппарата на орбиту. Поскольку x — теперь случайная вектор-функция, это требование состоит в минимизации некоторого функционала вида

$$J_1 \overline{(x(T), Rx(T))}, \quad (5.2)$$

где R — положительно определенная матрица, а черта означает знак осреднения.

Итак, задача синтеза сводится к определению такого $v \in V$, которое доставляет минимум функционала вида (5.2) при условии (5.1). Эта задача качественно отличается от задачи теории оптимального управления, в которой разыскиваются программа $x(t)$ и оптимальное управление $u(t)$ как функция времени. Более или менее очевидно, что в задаче (5.1), (5.2) речь должна идти не об отыскании $v(t)$, а об отыскании управления $v(x)$ как функции фазовых координат. Следовательно, задача отыскания оптимального корректирующего управления сводится к построению оператора обратной связи. Таким образом, постановка задачи синтеза аналогична классической задаче теории регулирования с той только разницей, что вместо функционала (1.2) рассматривается функционал (5.2).

Задача синтеза значительно сложнее задачи теории оптимального управления, и какие-либо стандартные методы ее численного решения отсутствуют (да и вряд ли могут быть созданы). Эти трудности носят принципиальный характер. Техническая постановка задачи синтеза требует учета случайных возмущений. Собственно говоря, для борьбы с ними и создается система управления (или коррекции). Следовательно, фазовый вектор $x(t)$ обязательно должен считаться случайной вектор-функцией. Представим себе, что нам задано $v(x)$. Уравнение (5.1) определяет случайный процесс $x(t)$, который зависит от природы случайного процесса $\theta(t)$. Предположим, что все статистические характеристики процесса $\theta(t)$ известны. Однако эффективно определить статистические характеристики процесса $x(t)$ можно только в некоторых исключительных

случаях, например, если f — линейная функция x и y . Единственным инструментом анализа общей нелинейной задачи является метод Монте-Карло. Его трудность для анализа случайных процессов общеизвестна. Она исключает возможность решения сколь-нибудь сложных задач.

Вот почему основные усилия были направлены на решение задач линейного синтеза. Этим термином называют задачи, в которых дифференциальные связи (5.1) линейны относительно x , y и u , и, кроме того, корректирующее управление v — линейная функция фазовых переменных: $v = B(t)x$. Для решения задач линейного синтеза разработан целый ряд эффективных инженерных методов. В том случае, когда матрица $B = \text{const}$, задача может быть сведена к некоторой задаче нелинейного программирования. К сожалению, она оказывается многоэкстремальной. Тем не менее В.М. Пономаревым предложен весьма эффективный метод, основанный на идее сглаживания [8]. Автором настоящей статьи показано, что для задачи линейного синтеза могут быть получены необходимые условия типа принципа максимума, а следовательно, и регулярные способы решения [4]. Что же касается задач нелинейного синтеза, то успехи здесь скромные. Те приближенные методы, которые имеют прикладную ценность, связаны главным образом с идеей статистической линеаризации И.Е. Казакова [9].

Примечание. Задача синтеза усложняется, если в реальных условиях измерение фазовых переменных производится с ошибкой и измеряются не сами фазовые переменные, а их некоторые линейные комбинации $y = Cx + \xi$, где C — некоторая $m \times n$ матрица, причем $m \leq n$, а ξ — случайный процесс. В этом случае управление может определяться не фазовыми переменными, а только их статистическими оценками \hat{x} : $v = B\hat{x}$.

Если синтез линеен, случайные процессы (t) и (t) — гауссовские, функционал качества квадратичен (в частности, имеет форму (5.2)), то задача обработки наблюдений — определение оценок \hat{x} и задача определения оптимальных коэффициентов обратной связи — элементов матрицы B разделяются. И в этом более сложном случае приходим к задаче измерений. Этот факт носит название фильтра Калмана—Бьюсси, хотя в силу своей очевидности он был известен и использовался специалистами задолго до работы Калмана.

6. Попробуем подвести некоторые итоги. Пусть речь идет о проектировании системы управления нестационарным движением, например, о выводе аппарата на околоземную орбиту. В этом случае сталкиваемся с решением по меньшей мере двух оптимизационных задач. Сначала находим оптимальную программу, а затем оператор обратной связи, без которого реальная траектория аппарата не может быть близкой к расчетной. Часть ресурса (горючее, энергия) расходуется на выполнение программного движения, остальное — на коррекцию (работу автопилота). Какой математический смысл имеет эта процедура?

Рассмотрим общую постановку этой задачи. Имеется некоторое движение, описываемое дифференциальным уравнением

$$\dot{x} = f(x, u, v, t, \xi), \quad (6.1)$$

содержащим управления u и v и случайные факторы $\xi(t)$. Мы хотим добиться некоторой цели управления. Эта цель состоит в минимизации функционала

$$J_1(u, v, x) = \min. \quad (6.2)$$

Здесь J_1 — дисперсия $\overline{(x - \bar{x}(T))^2}$. Кроме того, мы хотим добиться минимизации усилий на достижение этой цели

$$J_2(u, v, x) = \min. \quad (6.3)$$

Таким образом, по существу имеем многокритериальную задачу, с которой справляемся весьма фамильярно. Забывая, что функционал J_2 также должен носить характер математического ожидания, зачеркиваем в (6.1) величины ξ , v и решаем стандартную задачу оптимального управления с функционалом (6.3), полагая затем $x = \tilde{x}$, где $\tilde{x}(t)$ — оптимальная траектория; считая \tilde{x} известным, линеаризуем (6.1) относительно u , v и решаем задачу оптимального синтеза. В какой мере эта процедура отвечает той технической задаче, о которой идет речь? Ответить на этот вопрос нелегко.

Заметим прежде всего: поставленная задача одновременного отыскания минимума функционалов (6.2) и (6.3) непосредственного математического смысла не имеет. Математика одновременно

умеет разыскивать экстремум только одного функционала. Конечно, могут быть исключения. И с ними мы столкнулись в начале статьи. Если пересечение множеств решений вариационных задач, сформулированных для каждого функционала, непусто, то действительно можем реализовать некоторую процедуру последовательной оптимизации. Но такая ситуация — случай исключительный. Как правило, задачи минимизации J_1 и J_2 приводят к множествам решений $\{u, v_1\}$ и $\{u_2, v_2\}$ соответственно, которые не имеют общих точек (например, когда решения обеих задач единственные, не совпадающие друг с другом). В этом типичном случае принятие решения требует некоторого компромисса. Математически это означает замену функционалов (6.2) и (6.3) некоторым новым

$$J(J_1, J_2). \quad (6.4)$$

Формирование функционала (6.4) — процедура неформальная, и структура свертки целиком определяется физическим содержанием задачи и особенностью лица, принимающего решение. Но коль скоро функционал (6.4) сформулирован, то можно говорить о том, в какой мере справедлива та или другая процедура отыскания управления. В нашем случае она сводится к решению следующих задач:

$$I. J_2(x, u, 0) \min, \dot{x} = f(x, u, 0, 0),$$

$$II. J_1(y, v) \min, \dot{y} = Ay + Bv + C,$$

где $A = (df/dx)$, $B = (df/dv)$, $C = (df/d)$. Производные вычислены при $x = \tilde{x}$, $u = \tilde{u}$, $v = 0$, $\dot{x} = 0$. Здесь \tilde{x} и \tilde{u} — решения задачи I.

Обоснование указанной процедуры сводится к оценкам разности $J(u^*, v^*) - J(\tilde{u}, \tilde{v})$, где u^*, v^* — точное решение задачи о минимуме функционала J , а \tilde{v} — решение задачи II. Автору известен здесь лишь один результат [4], смысл которого можно пояснить следующим образом. Если затраты для достижения цели управления, т. е. значение функционала в задаче I, во много больше значения функционала в задаче II (для сравнения обе задачи должны быть безразмерными), а свертка имеет вид $J = c_1 J_1 + c_2 J_2$; $c_i \geq 0$; $c_1 + c_2 = 1$, то описанная процедура дает ошибку в определении оптимального J порядка отношения обоих функционалов.

Этот результат почти тривиален, однако какие-либо более глубокие факты, которые позволили бы оценить правомочность разделения задачи на расчет программного движения и синтез обратной связи, автору неизвестны.

Итак, выдвинутые техникой задачи управления на конечном интервале времени привели не только к новым задачам оптимизации. Оказалось, что эти задачи многокритериальные, что в свою очередь привело к необходимости ввести в теорию управления то понимание оптимальности, которое выработано в теории конфликтных ситуаций.

7. Усложнение технических объектов, для которых проектируются системы управления, и использование методов этой теории для управления сложными технологическими процессами привело к появлению теории систем, организованных по иерархическому принципу. Необходимость введения в системах управления иерархической структуры связана с процессами сбора и переработки информации с учетом времени и других ресурсов, которые затрачиваются на эти процессы. Поясним смысл этих обстоятельств на следующем примере. Предположим, что речь идет об управлении динамической системой

$$\begin{aligned} \dot{x} &= f(x, y, z, u, v, w, \delta, \epsilon, \eta), \\ \dot{y} &= g(x, y, z, u, v, w, \delta, \epsilon, \eta), \\ \dot{z} &= h(x, y, z, u, v, w, \delta, \epsilon, \eta), \end{aligned} \tag{7.1}$$

где x, y, z — фазовые переменные, δ, ϵ, η — помехи, и ставится задача об отыскании управлений u, v, w , зависящих от фазовых переменных и доставляющих максимум некоторому терминальному функционалу $J(x(T), y(T), z(T))$. Таким образом, речь идет о построении оптимального синтеза, т. е. определении $u(x, y, z), v(x, y, z), w(x, y, z)$. Предположим теперь, что измерение фазовых переменных и обработка этой информации (например, из-за размерности задачи, структуры системы связи и т. д.) требует времени τ . Тогда $u(t) = u(x(t - \tau), y(t - \tau), z(t - \tau))$ и т. д. Это значит, что решение, т. е. вектор управляющих воздействий в момент t будет приниматься по устаревшей информации в момент $t - \tau$. Если τ достаточно велико, то это решение будет приниматься тогда, когда уже в нем никакой необходимости не будет, т. е. процесс (7.1) оказывается неуправляемым. В подобных условиях естественной альтернативой будет

распараллеливание обработки информации и построение управлений по ограниченной информации, например, такой

$$u \quad u(x), v \quad v(y), w \quad w(z). \quad (7.2)$$

Управление вида (7.2) может в этих условиях оказаться более эффективным. В результате такой процедуры принятия решения назначение величины управляющих воздействий делегируется отдельным звеньям. В звене x назначается воздействие только по информации о состоянии звена x , в звене y — по информации о звене y и т. д.

Классическим примером такого рода иерархических систем являются роботы и, в частности, шагающая многоножка. Каждая ее нога делает свой шаг по ограниченной информации. Подобные системы естественно назвать рефлекторными. Это наиболее простой тип иерархических систем. Конструктор, создавая робот, закладывает определенное правило поведения, в зависимости от информации, доступной тому или другому звену. Выбор этого правила поведения, т. е. функций $u(x), v(y), w(z)$, определяется целями системы, т. е. некоторым функционалом $J(u, v, w)$, который явно или неявно всегда присутствует.

С точки зрения математики, рефлекторные иерархические системы не содержат каких-либо новых задач. Основная их особенность — блочная структура. Это позволяет обычно справляться с высокой размерностью подобных задач с помощью тех или других итерационных алгоритмов. Иерархическим системам рефлекторного типа посвящена огромная литература, начиная от работ Дж. Данцига и Т. Вульфа по блочному программированию и кончая многочисленными публикациями по так называемой теории систем.

8. Обширным полем приложения методов теории управления сделалась в последние годы экономическая наука. На грани 1940–1950-х годов многие задачи конкретной экономики начинают формулироваться на языке теории оптимальных решений. Это были прежде всего различные распределительные и балансовые задачи, в которых речь шла об отыскании такого распределения ресурса (в том числе инвестиций), которые доставляли бы максимум некоторой целевой функции.

На первых порах изучались только статические задачи. Однако уже и на этом этапе возникли глубокие связи между обеими

дисциплинами, которые в большой степени обогатили инструментарий теории управления.

Создание и развитие линейного, нелинейного и динамического программирования связано было главным образом с созданием методов математической экономики. Эффективный аппарат, созданный в начале 1950-х годов, начал, естественно, проникать и в другие области научного знания, где использовались идеи оптимизации. Поэтому первым актом возникшего альянса экономической науки и теории управления было широкое использование методов математического программирования в теории оптимального управления.

Поскольку для получения численного решения необходимо заменить дифференциальные уравнения их разностными аппроксимациями, то формально любая задача оптимального управления может трактоваться как некоторая задача математического программирования. Такая редукция во многих случаях оказалась полезной. Конечно, получающаяся при этом задача математического программирования обладает, как правило, очень высокой размерностью, в особенности если велик интервал времени, на котором изучается движение. Тем не менее редукция к задачам математического программирования иногда была единственным способом эффективного решения задачи. В самом деле, методы, использующие принцип максимума, т. е. методы, которые лежат в основе большинства алгоритмов, неприменимы в задачах с фазовыми ограничениями. Здесь показателен один эпизод. В 1962 г. на Конгрессе по астронавтике в Варне в качестве теста была предложена задача отыскания оптимальной траектории космического аппарата, снабженного двигателем малой тяги, который должен покинуть околоземное пространство, причем так, чтобы его траектория не пересекла поясов радиации. Эта задача была решена и докладывалась в следующем году Н.Я. Багаевой на Международной конференции по динамике спутников. Численный алгоритм Багаевой использовал идеи динамического программирования. Попытки решения этой задачи с помощью принципа максимума успеха не имели.

Таким образом, методы математического программирования, возникшие в экономике, сделались одним из важнейших инструментов решения оптимизационных задач управления. В дальнейшем этот альянс экономика — теория управления сделался еще

более глубоким, когда в экономических исследованиях широкое распространение получили динамические задачи. Оказалось, что задача планирования — это некоторая специальная задача оптимального управления. Идеи и методы теории оптимального управления начинают с этого момента широко использоваться в экономике. Принцип максимума, в частности, становится одним из основных инструментов численного анализа. Распространение сферы интересов теории управления на экономические процессы поставило целый ряд новых задач.

Учет лагов. Инвестиции в экономике обычно играют роль управляющих воздействий. Однако эффект этих управлений проявляется лишь временем управления. Строительство новых зданий, покупка, монтаж и ввод оборудования в строй требуют времени, которое называется лагом. Поэтому уравнения, описывающие экономические процессы на относительно небольшом интервале времени, имеют вид

$$\dot{x} = f(x, u(t - \tau), t). \quad (8.1)$$

Необходимость изучения систем вида (8.1) привела к целому ряду фундаментальных исследований по теории оптимального управления с запаздывающим аргументом.

Ограничения типа узких мест. При отсутствии экзогенных факторов ресурс, который должен использоваться для развития экономической системы, может быть создан лишь внутри самой системы. Существенная экзогенность ресурса и определяет специфику задач управления экономическими системами по сравнению с системами техническими. Ограничения в экономических задачах всегда имеют вид

$$(u) \leq (x). \quad (8.2)$$

Смысл ограничений состоит в следующем: объем продукта, который может быть использован на инвестиции (u) , т. е. на создание фондов x , ограничен производственными возможностями (x) . Ограничение (8.2) называется ограничением типа узких мест. Непосредственно задача (8.2) не вкладывается в схему Л.С. Понтрягина. Однако многочисленными исследованиями последних лет найден целый ряд возможностей построения численных решений

подобных задач, опирающихся либо на принцип максимума, либо на необходимые условия, ему аналогичные.

Задача управления на бесконечном интервале времени. Задача отыскания управления $u(t)$, максимизирующего (или минимизирующего) функционал

$$J = \int_{t_0}^{\infty} F(x, u, t) dt \quad (8.3)$$

с дифференциальными связями вида

$$\dot{x} = f(x, u, t), \quad (8.4)$$

при внимательном рассмотрении оказывается столь же естественной, как и задача ляпуновской устойчивости. Существуют многочисленные технические и экономические задачи, которые сводятся к (8.3), (8.4). В экономике они широко распространены и возникли одновременно с теорией полезности, в которой функция F имеет вид

$$F = (c, x)e^{-\rho t},$$

где $e^{-\rho t}$ — функция дисконтирования. Она показывает процесс обесценивания благ со временем. Возникновение в экономике задач с бесконечным временем связано с принципиальной невозможностью в рамках теории полезности назначать на правом конце какие-либо граничные условия. Для задач вида (8.3), (8.4) удалось доказать целый ряд теорем, носящих характер асимптотических оценок.

Например, пусть известен характер особенности $u(t)$ при $t \rightarrow \infty$ (это означает, что управление принадлежит заданному асимптотическому классу). Строится новое управление

$$\hat{u} = \begin{cases} u_1, & t \in [t_0, T], \\ u_2, & t \in T, \end{cases}$$

где u_2 — некоторое управление, принадлежащее тому же асимптотическому классу, что и допустимые управления, а u_1 — оптимальное решение следующей задачи Майера:

$$\hat{J} = \int_{t_0}^T F(x, u, t) dt \quad (x(T));$$

$$(x(T)) = \int_T F(x_2, u_2, t) dt, \quad \dot{x}_2 = f(x_2, u_2, t), \quad x_2(T) = x(T).$$

Теоремы дают оценки вида $J(\tilde{u}) - J(\hat{u}) = o(T^{-r})$, где \tilde{u} — оптимальное управление в зависимости от способа построения u_2 .

9. Идеи теории управления, сформировавшиеся в технических задачах, весьма плодотворны при изучении экономических механизмов. Выше уже говорилось о необходимости разбивать проблему управления нестационарным движением на два этапа. На первом решается задача отыскания программной траектории, на втором — задача об управлении, необходимом для ее фактической реализации. Задачи экономики всегда нестационарны, поэтому именно здесь особенно естественно использовать те способы анализа, которые сформировались в теории управления. Задача о выборе оптимальной траектории — известная задача оптимального планирования. Для осуществления плана необходима некоторая система обратной связи, которая реализуется экономическим механизмом. К сожалению, экономические системы настолько сложны, что обычная теория синтеза оказывается недостаточной. Экономические системы всегда организованы по иерархическому принципу.

Примечание. Необходимость введения иерархии в системе может служить мерой ее сложности.

Иерархические структуры социальных, в том числе экономических систем, имеют ряд качественных особенностей. Главная из них — подобные системы в принципе не являются рефлекторными. Поясним эти особенности на простом примере. Предположим, что имеется некоторое производственное объединение. Процесс производства будем описывать производственными функциями $p_i = p_i(x_i, u_i, \xi_i)$, где p_i — вектор продукта, вырабатываемого предприятием номера i , которое входит в объединение, x_i — его фонды, u_i — все прочие факторы, которые будем рассматривать как некоторые управления (количество рабочей силы, норма зарплаты и т. д.), ξ_i — случайный фактор, например, состояние погоды, оборудования, квалификации рабочей силы и т. д., ξ_i хорошо известные директору предприятия, но неизвестные генеральному директору объединения.

Пусть изменение фондов описывается уравнением $\dot{x}_i(x_i, v_i)$, где v_i — инвестиции.

Цель объединения — добиться максимизации некоторого функционала, например производства продукта в заданных пропорциях; его значение определяется результатом деятельности заводов, входящих в объединение (само объединение, т. е. его управление, никаких материальных ценностей не производит):

$$J_0 = \int_{t_0}^T F(p_1, p_2) dt. \quad (9.1)$$

Задача руководства объединения — так распорядиться управлениями u_i и v_i , чтобы максимизировать функционал (9.1). Но руководство объединения не знает величины x_i , а директора заводов знают. Значит, им не только проще назначать функции u_i и v_i , но они и лучше их могут назначить, поскольку располагают более полной информацией. Но директора заводов не знают целей объединения, т. е. функционала (9.1), и, кроме того, у них могут быть и самостоятельные цели (более или менее выгодная продукция и т. д.). Другими словами, генеральный директор не может назначить функцию $u_i(p_i)$ или $u_i(x_i)$. Он может только высказать пожелание или даже издать приказ. Действительное же решение, которое будет принято в нижнем звене директором завода, будет зависеть не только от приказа. Математически это означает, что нижние звенья будут принимать решения, стремясь максимизировать некоторые собственные функционалы $J_i(p_i)$. Но как же в этом случае будет обстоять дело с основной задачей объединения, которая состоит в максимизации функционала (9.1)?

Очевидно, объединениям необходимо иметь какие-то возможности воздействия на непосредственных производителей — на заводы. Для этого надо прежде всего часть управляющих воздействий, например инвестиции v_i , сохранить в распоряжении объединения. Кроме того, объединение может воздействовать на цели производителей с помощью системы штрафов и поощрений $y_i(p_i)$, зависящих от результатов деятельности производителей. Математически это будет означать, что производители, стремясь увеличить свой доход, будут максимизировать не функционал $J_i(p_i)$, а $J_i^* = J_i + y_i(p_i)$. В результате стратегия производителя будет функцией управления v_i , и функционалом от функции штрафа $y_i(p_i)$ и, конечно, x_i :

$$u_i = u_i(v_i, y_i, p_i). \quad (9.2)$$

Мы описали некоторую простейшую схему иерархической организации в экономике (более подробно см. [10]). Как видно, решения нижних звеньев системы (9.2) не могут быть точно определены ее конструктором. Решения принимаются нижними звеньями из условий, которые никогда точно не будут известны верхним звеньям. Центр, зная о существовании функционала J_i , никогда точно его знать не будет.

Как же в таких условиях объединение может определять величину своих управляющих воздействий — планировать свою деятельность? Для этого оно явно или неявно должно формулировать некоторое предположение о способах действий своего нижнего звена. Это предположение называется гипотезой поведения. Например, генеральный директор постулирует, что в ответ на штраф $y_i(p_i)$ завод номера i так распорядится своими возможностями, чтобы максимизировать некоторый функционал $J_i^*(v_i, y_i, p_i)$.

В этом случае он уже может определить управление u_i в форме (9.2). Задача управления нерелекторной системы сведена к релекторной. Теперь уже задача объединения сводится к такому назначению v_i и $y_i(p_i)$, которое максимизирует (9.1), разумеется, с учетом сохраняющейся при этом неопределенности информации о .

Сегодня математическая теория нерелекторных иерархических систем уже сделала первые шаги. Стал виден не только широкий спектр ее возможных приложений, но и методы ее исследования и основные проблемы, решение которых необходимо для создания эффективного аппарата исследования. Оказалось, в частности, что этот аппарат тесно связан с общей теорией игр с противоположными интересами, которая в последние годы развивается Ю.Б. Гермейером и его учениками [11].

Заметим, что объединению, чтобы управлять заводами, не нужна информация о целях нижних звеньев. Необходимо знать только функцию $u_i(v_i, y_i, p_i)$. Вот почему одна из центральных проблем быстро развивающейся теории иерархических систем — аппроксимация нерелекторных систем релекторными. Эта проблема может быть сформулирована на языке теории планирования эксперимента и теории адаптивных систем. Будущее покажет, какой из

этих путей будет реализован, но ее решение — это ключ к созданию эффективных методов расчета.

10. Автор старался показать, какое глубокое влияние оказали идеи оптимальности на эволюцию современной теории управления. На протяжении всех послевоенных лет теория управления, в свою очередь, оказывала глубокое стимулирующее влияние на развитие методов и идей оптимизации. Достаточно напомнить, что открытие принципа максимума было индуцировано прежде всего запросами теории управления.

Сегодня теория управления — это быстро прогрессирующая дисциплина. Непрерывно возникает большое количество новых задач, растут область ее приложения и сложность изучаемых объектов. Постепенно становится все более очевидной недостаточность традиционных методов анализа, требующих редукции практической задачи к четко поставленной математической задаче и эффективно-го алгоритма ее численного решения. Педантичное следование этой традиции резко ограничивает круг задач, которые имеет смысл изучать. Типична в проблемах управления социально-экономическими системами ситуация, где постановка задачи не совсем точная и размерность задачи не позволяет и надеяться получить точное решение, какова бы ни была мощность гипотетических ЭВМ будущего. Это не пессимизм, а реальность, с которой необходимо считаться. Вот одна из иллюстраций к сказанному. Задача о расписании 1000 работ проста с математической точки зрения и обычна для практики. Ее точное решение существует и алгоритм известен — прямой перебор. Однако даже если мы будем иметь ЭВМ в 10^6 раз более быстроедействующую, чем БЭСМ-6, то нам потребуются миллиарды лет для решения подобной задачи теории расписаний. Значит, надо искать другие, «нечеткие» способы ее решения. Человеческий разум как-то умеет справляться с подобными задачами. По-видимому, сталкиваясь с такими ситуациями, основное количество вариантов, которые заведомо плохи, наш разум умеет отбрасывать без анализа. А дальше с точностью, согласованной с нечеткостью своих целей, человек делает выборку уже с помощью формальных методов.

Сегодня эти обстоятельства интенсивно обсуждаются. Один из крупнейших современных представителей теории управления Л. Заде разрабатывает специальный аппарат для анализа таких нечетко поставленных задач, который опирается на понятие нечеткого

множества (fuzzy-set). Как ни заманчива и интересна подобная концепция, автор настоящей статьи полагает, что прагматическая основа развития теории управления в ближайшие годы будет связана с идеями имитации, т. е. построением вычислительных систем, объединяющих человека со способностью машины проделывать быстро и точно огромный объем логических операций. При этом надо, разумеется, рассматривать эксперта — биологическое звено имитационной системы — как звено некоторого алгоритма. Тогда формирование процедуры использования этих систем превращается в некоторую математическую задачу следующего инфрауровня.

Литература

1. *Красовский Н.Н., Моисеев Н.Н.* Теория оптимальных управляемых систем./Изв. АН СССР, Техническая кибернетика, 1967, № 5.
2. *Охоцимский Д.Е.* К теории движения ракет./Прикл. матем. и механ. 1940, 10, № 2.
3. *Понтрягин Л.С., Болтянский В.Б. и др.* Математическая теория оптимального уравнения. — М., Физматгиз, 1961.
4. *Моисеев Н.Н.* Численные методы теории оптимальных систем. — М., Наука, 1970.
5. *Петров Л.Н.* Инвариантные системы./В сб. «Техническая кибернетика». М., Наука, 1968.
6. *Розоноэр В.И.* Вариационный подход к проблеме инвариантности./Автоматика и телемеханика. 1963, № 6, 7.
7. *Величенко В.В.* О вариационном методе в проблеме инвариантности управляемых систем./Автоматика и телемеханика. 1972, № 4.
8. *Пономарев В.М.* Метод последовательной оптимизации в задачах управления./Изв. АН СССР, Техническая кибернетика, 1967, № 2.
9. *Казаков И.Е.* Статистические методы проектирования систем управления. — М., Машиностроение, 1969.
10. *Моисеев Н.Н.* Иерархические системы и теория игр./Кибернетика, 1973, № 12.
11. *Гермейер Ю.Б.* Игры с противоположными интересами (Теория принятия решений при неполном единстве). — М., Изд-во МГУ, 1972.

1974 г.

Кибернетическое описание эколого-экономических систем

Введение

В настоящее время особо актуальными становятся общие экономические проблемы концептуального характера, а не решение отдельных оптимизационных задач. Конкретное «экономизирование» может быть успешным только тогда, когда оно опирается на общие динамические концепции воспроизводства.

Необходимо исследовать общую перспективу эволюции экономических процессов. Для подобных утверждений есть много причин. Главная причина — в возросшем могуществе человека и в опасности принятия решений при ограниченной информации. За послевоенные годы возник целый ряд проблем, исследованием которых современная экономическая наука еще не занималась или занималась в недостаточной степени.

Способность человека оказывать влияние на окружающую среду заставляет рассматривать его производственную деятельность как естественную составляющую общих экологических процессов. Человек не против биоты, а вместе с ней. Под этим девизом должны вестись экономические исследования. Однако его реализация требует не только значительного расширения области экономических исследований, но и существенной трансформации наших представлений об их содержании, целях и аппарате.

Органическое единство планеты проявляется все больше. Взаимное влияние процессов, протекающих на разных широтах и меридианах, становится определяющим фактором эволюции человеческого общества. Эколого-экономические проблемы должны получить концептуальную основу научного анализа и методы отыскания рациональных коллективных решений.

Экономическая наука капиталистического общества не в состоянии объяснить и прогнозировать многие важнейшие явления современности, не в состоянии вскрыть причины инфляции и предотвратить ее. Кризис валютно-финансовой системы капитализма,

энергетический кризис, нарушение экологического равновесия ставят, казалось бы, неразрешимые проблемы. Вырождение буржуазной экономики, ее неспособность изучать объективные закономерности подчеркивает Дж. Гэлбрайт [1]: «...Она приводит к выводам, удобным для крупных корпораций, но болезненным для общества. Она культивирует утверждения (а с ними и поведение), которых требуют экономические интересы этих корпораций».

Величайшим преимуществом социалистического способа производства является плановое ведение хозяйства в масштабах всей страны. Вопросы совершенствования, организации и планирования социалистического производства находятся в центре внимания экономической политики нашей партии. В Отчетном докладе ЦК КПСС XXIV съезду Л.И. Брежнев говорил: «Центральное звено, сердцевина руководства народным хозяйством в условиях социализма — это планирование... Задачей первостепенного значения становится дальнейшее повышение его научного уровня. Назрела необходимость совершенствования методов планирования. Оно должно опираться на более точное изучение общественных потребностей, на научные прогнозы наших экономических возможностей, на всесторонний анализ и оценку различных вариантов решений, их непосредственных и долговременных последствий».

В современном производстве усложнение задач управления и планирования носит объективный характер и обусловлено в первую очередь количественными и качественными изменениями в объемах и структуре производства. Расширение современного производства, ускорение темпов научно-технического прогресса, совершенствование технологии, развитие специализации и кооперирования производства — вот факторы, заставившие качественно изменить отношение к вопросам планирования и управления. Существенное значение для повышения роли управления и планирования на современном этапе развития общества играет научно-техническая революция, обуславливающая единство науки и производства.

Перечисление подобных проблем можно продолжить. Все сказанное убеждает в необходимости фундаментальных исследований. Многообразие причин, связей, объемы информационных потоков, значимость следствий принимаемых решений делают многие проблемы неразрешимыми для традиционных методов экономического анализа. Нужны более общие подходы и методы, которые позволили бы в хаотическом нагромождении фактов увидеть общие

тенденции. Для этого прежде всего необходим язык, позволяющий дать единообразное описание различных процессов, в которых участвует человек, язык, отражающий в первую очередь динамический характер экономических процессов.

Обсуждению проблемы такого описания и некоторым соображениям по поводу методов анализа и посвящена настоящая работа. Примеры подобраны так, чтобы, не усложняя изложения, пояснить взгляды автора на эту проблему.

1. Кибернетические системы

Мы будем говорить о системах, состояние которых однозначно описывается конечномерным вектором x ; его компоненты x^j ($j = 1, 2, \dots, N$) являются функциями времени t и пространственной переменной r . Изменение состояния происходит благодаря действию случайных сил (t, r) , статистическое описание которых известно, воздействий (t, r) , заданных только своей мерой неопределенности G :

$$(t, r) \in G \quad t, r, \quad (1.1)$$

и управляющих воздействий u_i ($i = 1, 2, \dots, k$).

Будем говорить, что система имеет кибернетическое описание (является кибернетической), если эффективно построен оператор A , который позволяет определить состояние системы в любой момент времени t ($t_0, T(t_0)$) по значениям вектора $x(t_0)$ (t_1, t_0), при условии, что все внешние и управляющие воздействия фиксированы:

$$x(t) = A(x(t_0), t, u). \quad (1.2)$$

Те кибернетические системы, в которых управление является функцией фазовых переменных, возмущений и времени, называются рефлекторными, или рефлексными. Если управление $u(x, t)$ выбирается из условия максимизации единственного функционала $J(u)$, то такая система всегда является рефлексной. Все технические системы рефлексные. Биологические системы, которые управляются рефлексными, также рефлексные.

Таким образом, рефлексными мы называем все системы с четко запрограммированным управлением.

В отличие от технических и биологических систем экономические и социальные системы не являются рефлексными. В них участвуют люди и организации людей, обладающие свободой в принятии решений, которых будем называть субъектами системы.

В общем случае нерелекторных (нерелексных) систем управления u_i находятся в распоряжении субъектов и их выбор определяется стремлением к достижению ряда целей, которые без ограничений общности могут быть формализованы как стремление к максимизации некоторых функционалов I_i ($i = 1, 2, \dots, k$).

Примечание. Эти функционалы обычно выражаются через параметры, фазовые координаты системы. Но поскольку последние являются отображением множества управляющих воздействий, то без ограничения общности функционалы I_i могут считаться определенными в пространстве управлений, т. е. $I_i(u) = I_i^*(x)$.

Субъективным описанием системы будем называть ее описание с точки зрения отдельного субъекта. Для отыскания собственного управления каждый субъект должен иметь определенные гипотезы о выборе управлений остальными субъектами. Эти гипотезы будем формулировать в терминах принадлежности

$$u_i \in G_{ij}, \quad (1.3)$$

т. е. управление субъекта номера i , по мнению субъекта номера j , будет принадлежать к классу G_{ij} .

Исследование кибернетических систем может проводиться лишь с позиций определенного субъекта, основывается на его целях и его представлении о ситуации.

Интересы, выражаемые функционалами I_i , могут быть точно известными, более или менее известными (если пользоваться терминологией Л. Заде) и неизвестными. Но они всегда объективно существуют в нерелексных системах. Признание этого факта является краеугольным камнем теории.

2. Системы с иерархической структурой

Одной из важнейших задач анализа кибернетических систем является выделение классов систем, для которых можно разработать некоторые общие методы анализа. К таким классам относится так

называемая иерархическая система. Рассмотрим простейший пример такой системы — систему с двухуровневой иерархией. Предположим, что она состоит из $K + 1$ субъекта. Интересы субъектов описываются следующим образом:

$$J_0(u_0^1, u_0^N, u_1, u_N) \quad \max, \quad u_0^i \quad G_0, \quad (2.1)$$

$$J_i(u_0^i, u_i) \quad \max, \quad u_i \quad G_i, \quad (2.2)$$

где управление u_0^i находится в распоряжении субъекта номера 0, которого впредь будем именовать центром, а управления u_i — в распоряжении субъектов номеров 1, 2, ..., N , которых будем называть производителями.

В нашей системе субъекты неравноправны. Центр имеет право первого хода — он знает о том, что его выбор, т. е. векторы u_0^i , будет известен остальным производителям. Благодаря этому класс иерархических систем обладает рядом свойств, допускающих качественное и численное исследование. Поясним некоторые из этих свойств на примере, когда $N = 1$.

Во взаимоотношениях центра и производителя могут возникнуть следующие ситуации.

Ситуация Γ_1 . Центр сообщает производителю значения вектора u_0 , например количество ресурса, которое будет ему передано.

Условимся проводить все исследования с точки зрения центра, т. е. использовать его субъективное описание. Последнее означает, что центр обязан сформулировать определенную гипотезу о поведении своего производителя. Значит, центр будет предполагать, что производитель, получив ресурс u_0 , использует его так, чтобы

$$J_1(u_0, u_1) \quad \max. \quad (2.3)$$

Тогда на основе (2.3) центр ожидает, что управление производителя будет $u_1 = u_1(u_0)$ и, значит, его цель $J_0(u_0, u_1(u_0)) \quad \max$ будет достигнута таким выбором u_0 , который гарантирует

$$J_0(u_0, u_1(u_0)) = J_0^*(u_0) \quad \max.$$

Итак, рациональный выбор стратегий в описанной ситуации сводится к решению ряда оптимизационных задач.

Ситуация Γ_2 . Центр сообщает производителю функцию $u_0(u_1)$. Это означает, что величина воздействия центра будет зависеть от действия производителя. Возможны различные интерпретации этого факта. Например, $u_0(u_1)$ может быть поощрением или штрафом. Центр ожидает, что откликом производителя будет такой выбор u_1 , который удовлетворяет условию

$$J_1(u_0, (u_1), u_1) \quad \max. \quad (2.4)$$

(2.4) — это некоторая вариационная задача. Ее решением является функция

$$u_1 = u_1[u_0(u_1)], \quad (2.5)$$

определяющая отображение множества функций $u_0(u_1)$ в пространстве функций $u_1(u_0)$.

Имея возможность определить отклик производителя, центр может определить свое воздействие наилучшим для себя образом. Для этого ему достаточно выбрать функцию $u_0(u_1)$ так, чтобы она решала следующую вариационную задачу:

$$J_0(u_0, (u_1), u_1[u_0(u_1)]) = J_0^*[u_0(u_1)] \quad \max. \quad (2.6)$$

Ситуация Γ_3 . Продолжим рекурсию: центр может сообщить производителю функции и более сложного вида:

$$u_0 = u_0[u_1(u_0)]. \quad (2.7)$$

Экономическая интерпретация функции (2.7) может быть следующей: u_0 — объем ресурса, выделяемого центром, $u_1(u_0)$ — способ его использования производителем. Значит, величина выделяемого ресурса зависит от способа его использования. На основе своего субъективного представления центр предполагает, что производитель выберет величину u_1 из условия

$$J_1(u_0, [u_1(u_0)], u_1) \quad \max. \quad (2.8)$$

Вариационная задача (2.8) будет иметь решение вида

$$u_1 = u_1(u_0 [u_1(u_0)]). \quad (2.9)$$

Располагая выражением (2.9), центр делает выбор функции (2.7) из условия

$$J_0(u_0 [u_1(u_0)], u_1(u_0 [u_1(u_0)])) \rightarrow \max.$$

Кажется, что описанную рекурсию можно продолжать неограниченно. Однако оказалось, что здесь имеет место теорема, доказанная С. Кукушкиным. Обозначим через J_{0i} значения функционала J_0 в ситуации Γ_i . Тогда: 1) $J_{01} > J_{03} > J_{02}$; 2) $J_{02} > J_{04} > J_{06} > \dots$, $J_{03} > J_{05} > J_{07}$. И число возможных управлений, существенно отличающихся друг от друга, равно всего лишь трем.

Разумеется, в более общем случае количество различных ситуаций несколько больше, но оно тем не менее конечно. Это дает возможность провести исчерпывающее изучение систем с иерархической структурой.

3. Система Гермейера

Важный класс кибернетических систем составляют так называемые системы с иерархическим вектором интересов, которые впервые были изучены Ю.Б. Гермейером и И.А. Вателем [2]. В таких системах интересы субъекта номера i , в распоряжении которого находится управление u_i , представлены как «личным», так и «общественными» компонентами. При этом последние отражают групповые интересы всех тех сообществ, членом которых является данный субъект. Здесь мы рассмотрим случай, когда вектор интересов состоит только из двух компонентов — личного $f(u_i)$ и общественного $F(u_1, \dots, u_N)$, причем субъект номера i стремится к максимизации функционала

$$I_i = [f_i(u_i) F(u_1, \dots, u_N)], \quad (3.1)$$

где \square — некоторый оператор свертки векторного критерия, состоящего из компонентов f_i и F . Например,

$$I_i = \min\{f_i(u_i); \quad \square_i F_i(u_1, \dots, u_N)\}, \quad (3.1)$$

$$I_i = f_i(u_i) \quad \square_i F_i(u_1, \dots, u_N), \quad (3.2)$$

где α_i и β_i — весовые коэффициенты в соответствующих свертках критериев.

Кибернетические системы с критериями вида (3.1) в дальнейшем будем называть гермейеровскими.

Для кибернетических систем гермейеровского типа может быть формализовано понятие гомеостазиса. Границей гомеостазиса в пространстве переменных будем называть поверхность

$$F(u_1, \dots, u_N) = F_0, \quad (3.2)$$

выделяющую в этом пространстве область существования (гомеостазиса, стабильности) системы. Условимся, что области гомеостазиса отвечают значения функционала, превосходящие F_0 . Таким образом, максимизация функционала F_i дает формальное описание стремления субъекта находиться в возможно более стабильном состоянии.

Проблему определения множества функций u_i в пространстве управлений, удовлетворяющих условию

$$F(u_1, \dots, u_N) = F_0, \quad (3.3)$$

будем называть общей проблемой стабильности. Частной задачей стабильности (гомеостазиса) будем называть задачу определения отдельных точек области гомеостазиса.

Важным частным случаем кибернетических систем являются динамические системы со свободными функциями. Для таких систем проблема определения области гомеостазиса может быть переформулирована в терминах областей достижимости. Для того чтобы провести подобную редукцию, необходимо только использовать принципиальную возможность построения отображения множества управлений в пространство состояний.

Выбор управляющих воздействий — это акт принятия коллективных решений; следовательно, одной из важнейших проблем теории является выработка принципов принятия таких решений. Как известно, для общего случая кибернетических систем эта проблема практически неразрешима. Однако гермейеровские системы в этом отношении являются известным исключением.

Для статистических гермейеровских систем справедлива следующая теорема Гермейера—Вателя [2]. Пусть субъекты выделяют

из наличных ресурсов a_i долю u_i на увеличение общественного критерия и долю $(a_i - u_i)$ — на увеличение личного критерия, причем свертка критериев имеет вид (3.1), где f и F — монотонно возрастающие функции по своим аргументам. Тогда существуют устойчивые решения (точки равновесия по Нэшу), среди которых по меньшей мере одно является эффективным (т. е. принадлежащим множеству Парето).

Эта теорема позволяет надеяться, что и в более общем, чем (3.1), случае в качестве принципа принятия решения целесообразно взять принцип выбора устойчивого паретовского решения. Таким образом, определение условий (необходимых или достаточных), которым должен удовлетворять оператор A для того, чтобы теорема Гермейера—Вателя оставалась справедливой, оказывается одной из центральных задач теории.

Если рассматривается свертка (3.1), то максимизация $F(u_1, \dots, u_N)$ приведет к паретовскому компромиссу между $f_i(u_i)$ при метачели $F(u_1, \dots, u_N)$, который обладает устойчивостью по общему критерию F при учете интересов $f_i(u_i)$.

Для линейных систем можно показать, что такой компромисс дает результат, совпадающий с паретовским компромиссом для системы (3.1) при специальном выборе коэффициентов λ_i . Этот результат получен Ф.И. Ерешко.

4. Параметризация, базовые траектории, человеко-машинный диалог

Кибернетическое описание всегда является приближенным модельным описанием. В процессе исследования приходится упрощать отдельные модели, иногда даже заменять их конечными связями. Такой процесс называется параметризацией.

Идеи параметризации занимают очень большое место в физике. Например, сложнейшую модель взаимодействия молекул газа после параметризации удастся заменить конечным соотношением — уравнением состояния. Общая идея малого параметра А. Пуанкаре дает еще один класс методов параметризации и т. д.

Проблема параметризации заключается в построении такого упрощенного описания, которое обеспечивает близость решаемых задач для упрощенного и исходного описаний. Понятие близости понимается в смысле тех или иных слабых метрик, определяемых критериями задачи. Так, например, параметризация модели Навье—Стокса моделью погранслоя дает отличающуюся от реальности

топологию линий тока, но в то же время гарантирует близость значений такого функционала, как суммарное сопротивление.

В отличие от физических систем, где методы параметризации уже давно находятся в центре внимания математиков, в кибернетических системах подобные методы анализа практически отсутствуют. Естественно, что кибернетические системы, построенные для исследования реальных ситуаций, оказываются весьма сложными и требуют создания специальных численных методов. Однако и численное исследование часто бывает затруднено и ему должен необходимо предшествовать определенный качественный анализ.

Задачи параметризации в теории управляемых систем еще не являются даже строго поставленными математическими задачами.

Среди способов получения приближенных оценок особое место занимают различные варианты теории возмущений. Но для их построения нужно уметь определять некоторые опорные траектории.

В теории управления для анализа сложных стохастических систем используется так называемый метод двухэтапной оптимизации. Поясним его сущность на простейшем примере динамической системы. Пусть эволюция системы описывается уравнением

$$\dot{x} = f(x, u, \xi), \quad (4.1)$$

где $u \in G_u$ — управление, ξ — случайный процесс. Выбор u и $\xi(x, t)$ ограничен критериями J_1 и J_2 , один из которых имеет смысл энергии

$$J_1(x, u) = \min. \quad (4.2)$$

Если множество возможных траекторий $x(t)$ интерпретировать как множество траекторий аппарата, достигающих заданную орбиту за время T , то условие (4.2) означает достижение заданной орбиты с минимальной затратой топлива. Пусть вторым функционалом будет дисперсия

$$J_2 = \overline{(x_T - \hat{x}_T)^2} = \min, \quad (4.3)$$

где \hat{x}_T и x_T — соответственно желаемое и реально достигнутое конечное состояния системы. Этот критерий можно трактовать как характеристику точности достижения желаемого результата.

Таким образом, проблема состоит в том, чтобы выбрать управление, а следовательно, и траекторию так, чтобы наилучшим образом, с минимальными затратами, достичь цель управления. Для того чтобы сделать эту задачу математической, необходимо провести еще некоторый неформальный анализ проблемы свертки критериев (4.2) и (4.3). Оказывается, что при известных условиях для анализа этой конфликтной ситуации можно использовать теорию возмущений.

Положим 0. Вместо критерия (4.3) рассмотрим критерий

$$J_2 = \begin{cases} \dots, & \text{если } x_T = \hat{x}_T. \\ 0, & \text{если } x_T \neq \hat{x}_T. \end{cases} \quad (4.4)$$

Задача с критериями (4.2) и (4.4) уже разрешима. Обозначим ее решение через $\tilde{x}(t)$ и $\tilde{u}(t)$ и положим $x = \tilde{x} + y$, $u = \tilde{u} + v(y, t)$. Линеаризуя эту задачу относительно y , v и \dots , приходим к следующей задаче:

$$\dot{y} = By + Cv + D, \quad \overline{y_T^2} \rightarrow \min, \quad (4.5)$$

где B , C и D — некоторые матрицы. (4.5) — это задача оптимального синтеза. Ее решением является некоторая случайная функция $v = v(y, t)$.

При определенных ограничениях на случайный процесс $\dots(t)$ такая поэтапная оптимизация имеет смысл и дает приближенное решение для любой линейной свертки критериев J_1 и J_2 : $J = C_1 J_1 + C_2 J_2$, т. е. для любых весовых коэффициентов C_1 и C_2 (подробнее см. [3]). Конечно, центральной здесь является проблема выбора базового решения. Неправильный выбор может привести к ложным результатам.

Задачи, возникающие в практике, обычно имеют такую размерность, что получить строгий результат, даже с помощью ЭВМ гипотетической мощности, невозможно. В этом случае единственный путь — использовать диалог с помощью соответствующей имитационной системы. Такой подход тем более оправдан, так как реальные задачи бывают поставлены достаточно точно. В их постановке, а следовательно, и в результатах, непременно присутствует элемент «размытости» (если пользоваться терминологией Л. Заде).

Метод имитации по существу представляет собой способ вариантного анализа с оценкой результата с помощью эксперта. Однако организация процедур диалога всегда оказывается нетривиальной и требует специальных исследований.

Диалог — это тоже некоторый алгоритм оптимизации, в котором присутствует процедура расчета величин, характеризующих целевую функцию, хотя сама целевая функция может быть и неизвестной. И оценку результата счета, и выбор следующего варианта производит эксперт. Несмотря на большую роль различных эвристических соображений, указанную систему процедур часто удается привести к такой форме, которая значительно уменьшает количество перебираемых вариантов.

5. Несколько замечаний об эколого-экономических системах

Одним из важнейших классов нерелексных систем, изучение которых с каждым годом становится все более важным, являются системы, описывающие взаимное влияние экологических и производственных процессов. Они как правило, крайне сложны и имеют в своем составе модели весьма разной природы.

Практически любая система взаимодействия человека и окружающей среды имеет структуру, представленную на рис. 1. Если исключить модели человеческой активности, то система, изображенная на рис. 1, оказывается рефлексной.

Сложности ее анализа имеют неформальный характер и связаны с нерелексностью блока человеческой активности.

Количественный анализ задач конкретной практики должен сразу ориентироваться на использование методов имитации. Полная формализация задачи возможна, по-видимому, только в особо простых случаях.

Но даже если удастся преодолеть трудности формализации, анализ систем в каждом конкретном случае все равно будет весьма сложным. Основная посылка для его упрощения — различие в характерных временных масштабах. Характерное время изменения климатических условий T_k , например, существенно больше характерных времен других процессов. Поэтому мы имеем возможность изучать верхние модели независимо от модели климата. Обратное, разумеется, делать нельзя, так как антропогенные воздействия на климат и есть задача анализа.

Далее, процессы в биоте и процессы круговорота веществ в природе являются по своей природе рефлексными и, следовательно,

Следовательно, какой бы ни был набор, субъектов рассматриваемой кибернетической системы, среди критериев, которыми определяются субъективные решения, всегда присутствует критерий стабильности, общий для всех субъектов.

Но если система гермейеровская, то целесообразно рассматривать для нее проблему отыскания устойчивых коллективных решений, принадлежащих множеству Парето.

Таким образом, изложенная схема рассуждений дает определенные основания для построения теории, позволяющей исследовать допустимые типы решений в процессах взаимодействия производственной деятельности людей и окружающей среды. По мере своего развития такая теория может оказаться полезной для решения многочисленных проблем управления ресурсами и производственной деятельностью, вызванных необходимостью сосуществования.

6. Программный метод (случай рефлексных систем)

В технических системах управления, т. е. в теории рефлексных систем, возник так называемый метод программ, или программный метод. Поясним его сущность на примере системы (4.1). Пусть дана динамическая система

$$\dot{x} = f(x, u, \xi), \quad (6.1)$$

где $u \in G_u$ — управление, а $\xi(t)$ — случайный процесс. Предположим также, что условие $u \in G_u$ имеет характер интегрального неравенства

$$\int_{t_0}^T F(u, t) dt \leq Q. \quad (6.2)$$

Обозначим через x_T цель управления. Так как $x(t)$ является случайным процессом, то условие достижения x_T должно выражаться требованием минимума некоторой вероятностной характеристики этого процесса. Например, достижение цели управления будем характеризовать величиной функционала (4.3).

Тогда

$$J_2 = \overline{(x_T - x(T))^2} \rightarrow \min. \quad (6.3)$$

Достижение цели (6.3) — не единственная задача. Предположим, что мы стремимся достичь цель при условии

$$J_1 = \int_{t_0}^T (x, u, t) dt \quad \min. \quad (6.4)$$

Начальное состояние будем считать фиксированным

$$x(t_0) = x_0. \quad (6.5)$$

Примечание. В рамках излагаемой схемы могут рассматриваться и более сложные ситуации, когда на функционирование системы накладываются не два функционала J_1 и J_2 , а целый ряд ограничений

$$J_i = \min, \quad i = 3, 4 \quad (6.6)$$

Рассмотрим задачу (6.1)—(6.5) с позиций двухэтапной оптимизации. Положим

$$x_0(t). \quad (6.7)$$

$x_0(t)$ может быть прогнозом случайного процесса или известным значением его математического ожидания. Имея в виду переход к нерелективным системам, величину $x_0(t)$ целесообразно называть сценарием.

Если реальную внешнюю обстановку заменить сценарием, то процесс $x(t)$ становится детерминированным:

$$\dot{x} = f(x, u, x_0(t)) = f^*(x, u, t). \quad (6.1)$$

Условие (6.3) заменяется условием

$$x(T) = x_T. \quad (6.3)$$

Решение

$$x = x^0(t), \quad u = u^0(t) \quad (6.8)$$

будем называть соответственно программной траекторией и программным управлением.

Теперь для того чтобы преодолеть влияние нерасчетных факторов, нужно построить некоторую систему обратной связи. Это можно сделать различными способами. Опишем два из них.

1. Положим $u = u^0 + v$. Тогда условие (6.2) примет вид

$$\int_{t_0}^T F(u^0 + v, t) dt = Q,$$

или

$$\int_{t_0}^T F^*(v, t) dt = Q. \quad (6.9)$$

Положим далее $x = x^0 + y$, $u = u^0 + v$. Новая фазовая переменная y будет удовлетворять уравнению

$$\dot{y} = f(y, t, v), \quad (6.10)$$

где

$$f(x^0 + y, u^0 + v, t) = f(x^0, u^0, t) + f(y, v, t),$$

а цель управления (6.3) примет вид

$$J_2 = \overline{y(T)^2} \rightarrow \min. \quad (6.11)$$

Итак, мы пришли к некоторой задаче синтеза: определить функцию $v = v(y, t)$, доставляющую минимум функционалу (6.11) при условиях (6.9), (6.10) и условию

$$y(t_0) = 0. \quad (6.12)$$

Эта задача хорошо известна, существуют многочисленные приближенные способы ее решения. Основная трудность состоит в том, что управление v , по существу, должно быть случайным процессом, поскольку оно компенсирует случайные возмущения. Однако ограничение (6.9) (т. е. величина Q) всегда является детерминированной величиной. Теория управляемых случайных процессов не дает хороших рекомендаций для решения этой задачи. Поэтому в инженерной практике условие (6.9) часто заменяют рандомизированным условием, например

$$\int_{t_0}^T F^*(v, t) dt = Q. \quad (6.13)$$

2. В последние годы получил распространение способ организации обратной связи, который удобно назвать способом «водохранилища» (скользящего плана).

Задачи управления процессами иригации имеют следующую особенность. Количество воды, которое должно быть направлено на поля — величина Q^* , существенно зависит от погодных условий лета (случайный процесс (t)). Следовательно, Q^* сама должна быть случайным процессом. Но то количество воды, которое скопилось в водохранилище к концу паводка, — величина Q — практически не зависит от летних погодных условий. Таким образом, задача управления состоит в том, чтобы максимизировать урожай при не известных заранее внешних условиях, располагая вполне определенным запасом воды.

Опишем схему организации процесса управления.

Сначала по прогнозу x^0 определяем программу u и u^0 и в течение некоторого времени полагаем $u = u^0(t)$. Так как в общем случае $x(t) \neq x^0(t)$, то при $t_1 > t_0$ $x(t_1) \neq x^0(t_1)$.

При $t = t_1$ строим новый прогноз $x^1(t)$ с учетом информации о случайном процессе $x(t)$, который мы наблюдаем в течение времени $t - t_1$. Затем строим новую программу и т. д.

Метод водохранилища позволяет строить эффективную обратную связь, удобен для практического использования, однако строго математического обоснования не имеет. Неизвестно, например, какое отношение он имеет к оптимальному управлению. Очень важно было бы выяснить, при каких условиях управление, полученное с помощью этого метода, дает оптимальное решение задачи и в каком смысле.

7. Программный метод (общий случай нерефлексных систем)

Рассмотрим кибернетическую систему достаточно общего вида

$$\dot{x} = f(x, u_1, \dots, u_N, \dots). \quad (7.1)$$

Как уже было отмечено, исследование можно проводить только с субъективных позиций. Будем вести анализ с позиций субъекта номера 1. Для того чтобы легче было использовать содержательные понятия, рассмотрим частный случай системы (7.1):

$$\dot{x}_i = f_i(x_i, u_1, \dots, u_T), \quad (7.2)$$

где x_i — вектор, описывающий состояние системы номера i .

Первый этап реализации программного метода состоит в назначении цели, т. е. функционала J_1 , который описывает стремление субъекта номера 1, и в описании сценария, т. е. в формулировке гипотезы поведения остальных субъектов — в назначении (прогнозе) управлений u_2, u_3, \dots, u_N и прогнозе внешних условий (t) .

Разработка сценария требует специальных процедур. Некоторые особенности этой проблемы были показаны на примере иерархических систем и систем Гермейера.

Если в результате исследования этого этапа мы сможем тем или иным образом задать величины управлений и внешних воздействий $u_i = u_{i0}(t), \dots, (t), i = 2, 3, \dots$, то придем к системе (7.2), которая содержит только одно управляющее воздействие u_1 . Располаясь этим управлением, можно реализовать всю схему программного метода, которая изложена в предыдущем разделе. В этой схеме однако, если речь идет о нерелфесных системах, есть одна особенность. В рефлексных системах проблема цели (или целей) не возникает. Иное дело — общие кибернетические системы. В разделе 1 было сказано, что каждый субъект системы обладает определенными целями и что любые действия целенаправленны. В действительности же все бывает сложнее — назвать эти цели далеко не просто. В действительности каждый субъект функционирует в условиях, которые ему точно не известны и которые непрерывно изменяются. Конечно, он ставит себе определенные цели, которые отражают его субъективное представление о действиях, необходимых или достаточных для сохранения гомеостазиса. Например, в экономике в качестве целевой функции часто используют прибыль, доход. Действительно, в условиях традиционной рыночной экономики максимизация прибыли является наилучшей гарантией стабильности. В условиях же огромной концентрации капиталов, как утверждает Дж. Гэлбрайт, такое условие не является ни необходимым, ни достаточным. Более важно, например, иметь гарантированные контракты.

Таким образом, назначение целей типа

$$J_1 = \max \quad (7.3)$$

и фиксация времени их достижения $t \in T$ означают, по существу, организацию специальной петли обратной связи в системе управления, сохраняющей гомеостазис субъекта. Реализуя программную траекторию (7.3), мы в некоторый момент $t \in T$ придем к новой ситуации, которая будет характеризоваться новыми внешними условиями (t) и новыми состояниями x_i других субъектов системы. Это приведет к назначению новых целей и т. д.

Именно по такой схеме и реализуется программный метод управления и планирования в централизованной социалистической экономике.

8. Простейший пример анализа нерефлексной кибернетической системы

Изучение проблемы построения кибернетических систем, описывающих взаимоотношения человека и окружающей среды, показывает, что целесообразно использовать теорию воспроизводства и аппарата производственных функций, которые дают общий язык для описания как процессов производства, так и процессов, происходящих в биоте.

Рассмотрим двухсекторную модель экономики. Введем следующие обозначения: F_i — объем фондов в подразделении номера i , L_i — количество людей, занятых в производственном процессе, Y_i — «уровень жизни», k_i — «коэффициенты амортизации», Y_i — инвестиции, P_i — поток продукта, Q_i — запасы. Уравнения модели примут следующий вид:

$$\begin{aligned} \dot{F}_i &= Y_i - k_i F_i, \quad i = 1, 2, \\ \dot{Q}_1 &= P_1 - Y_1 - Y_2, \\ \dot{Q}_2 &= P_2 - (L_1 + L_2), \\ P_i &= F_i \cdot f_i(L_i). \end{aligned} \tag{8.1}$$

Здесь F_i — производственные функции. Все величины, входящие в уравнения (8.1), неотрицательны.

Кроме фазовых переменных F_i и Q_i в уравнения входят величины Y_i , L_i и P_i , которые в рамках этой модели следует считать управляемыми. В действительности они связаны с функционированием других моделей — модели распределения и модели инвестиций. Выбор этих величин отражает интересы и цели государства, общества и, в частности, интегрированные интересы потребителей и производителей. Для того чтобы их описать, нужно добавить еще

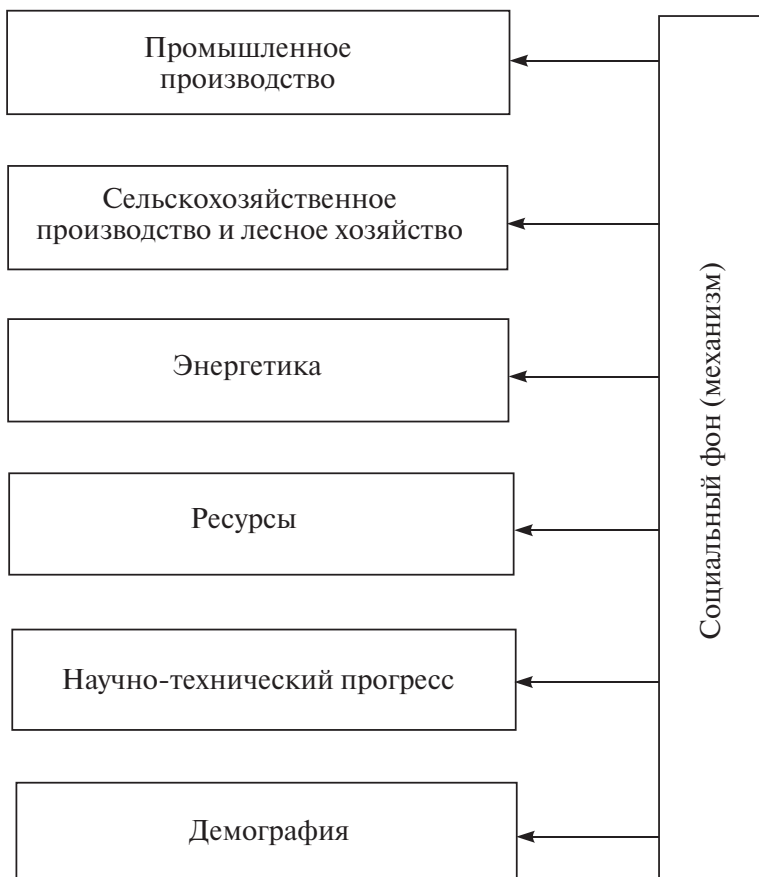


Рис. 2. Общая схема блока человеческой активности

целый ряд моделей и прежде всего модель функционирования денег. Уравнения, описывающие производственную деятельность, — это законы сохранения. Они не зависят от социальной природы общества. Что же касается других моделей, то они уже не являются стандартными и связаны с функционированием социального механизма общества. В рыночной экономике эти модели будут иными, чем в плановой. Таким образом, для полного анализа даже двухсекторной модели экономики необходимо привлечь целый ряд других моделей. Примерная их схема приведена на рис. 2.

Итак, даже в простейшем случае приходится рассматривать кибернетическую систему весьма большой размерности. Поэтому если нужно подробно исследовать блок воспроизводства, то следует

каким-то образом параметризовать модели. Параметризация дает на данном уровне агрегированное описание механизмов, действующих в изучаемом экономическом организме.

Рассмотрим, например, возможности инвестиционной политики. В этой ситуации естественно считать заданным уровень жизни t . Задание t по существу представляет собой параметризацию некоторой системы моделей, которая в настоящей работе включена в единый блок социального фона, объединяющий, в свою очередь, ряд других моделей. Величина t в капиталистической экономике является результатом конфликта между профсоюзами и предпринимателями, а в плановой экономике должна выражать концепцию повышения жизненного уровня. Общее количество работающих

$$L(t) = L_1(t) + L_2(t) \quad (8.2)$$

и отношение

$$(t) = \frac{L_2(t)}{L_1(t)} \quad (8.3)$$

также условимся считать заданными.

Следующим этапом анализа является выбор базовой траектории. Предположим, что речь идет о рыночной экономике. Тогда в качестве базовой траектории можно рассматривать, например, некоторый равновесный режим — режим «наибольшего социального благополучия». И условие полной занятости, по-видимому, следует отнести к числу подобных условий. Кроме того, рассматриваемый режим будет характеризоваться некоторым заданным ростом потребления.

На равновесном режиме с заданным ростом потребления t следует принять $Q_i = 0$. В самом деле, растущие или убывающие запасы окажутся дополнительным стимулом человеческой активности и приведут в действие различные механизмы, которые существуют в обществе. Таким образом, второе уравнение относительно Q_2 дает

$$F_2(\hat{}_2, L_2) = L, \quad (8.2)$$

где $\hat{}_2$ обозначает базовое решение, и позволяет определить необходимый объем фондов $\hat{}_2$ во втором подразделении. Но тогда однозначно определяются и инвестиции \hat{Y}_2 во второе подразделение:

$$\hat{Y}_2 = \hat{k}_2 \hat{Y}_2. \quad (8.3)$$

После этого первое уравнение относительно запасов дает

$$\hat{Y}_1 = F_1(\hat{Y}_1, L_1) - \hat{Y}_2,$$

и приходим к следующему дифференциальному уравнению относительно фондов:

$$\dot{\hat{Y}}_1 = F_1(\hat{Y}_1, L_2) - \hat{Y}_2 - k_1 \hat{Y}_1. \quad (8.4)$$

Таким образом, в качестве базовой траектории можно принять любое из решений задачи Коши для уравнения (8.4).

Далее строим теорию возмущений. В данном случае мы просто линеаризуем задачу относительно базовой равновесной траектории: $\hat{Y}_i = \hat{Y}_i + y_i$; $Y_i = \hat{Y}_i + y_i$; $Q_i = \hat{Q}_i + q_i$. В результате приходим к системе четырех уравнений первого порядка. Если принять, что (t) , $L = L(t)$, (t) , то возмущенная система будет линейной системой с переменными коэффициентами, а величины y_1 и y_2 в ней — свободными функциями. Эти величины должны реализовать обратную связь по наблюдаемым параметрам системы. Если речь идет о рыночной экономике, то такими наблюдаемыми параметрами являются только избыток или недостаток товаров на рынке, т. е. величины q_1 и q_2 . Поэтому в данном случае естественно принять

$$y_i = -c_i q_i, \quad (8.5)$$

где c_i — некоторые положительные коэффициенты усиления.

Примечание. Таким образом, рыночный механизм параметризован нами в форме отрицательной обратной связи (8.5). В рамках данной схемы определить значения c_i нельзя. Для этого нужен более детальный анализ модели инвестирования. Но, во всяком случае качественно, отрицательная обратная связь вида (8.5) соответствует стремлению производителей максимизировать ожидаемую прибыль. Приняв y_1 и y_2 в форме (8.5), получим замкнутую систему уравнений, относительно которой можно доказать следующее утверждение.

Решение любой задачи Коши с неотрицательными начальными данными для уравнений (8.4) в том случае, когда Y_1 и L являются монотонно возрастающими, неустойчиво, причем переменные, характеризующие второе подразделение, имеют колебательный характер, а переменные, характеризующие первое подразделение, — колебательно неустойчивы с экспоненциальным ростом амплитуды.

Этот результат показывает, что в условиях рыночной экономики при одновременном росте потребления и сохранении занятости стабилизация роста за счет изменений инвестиций невозможна.

В плановой экономике в качестве базовой должна быть принята траектория, реализующая некоторую концепцию развития производства. Во многих работах ее формулируют как максимизацию потребления, например:

$$J(Y_1, Y_2) = \int_0^T L(t) dt \quad \max. \quad (8.6)$$

Из-за линейности задач относительно управлений оптимальные управления всегда носят релейный характер:

$$\begin{aligned} Y_1 &= F_1, t \in [0, t^*]; Y_1 = 0, t \in [t^*, T]; \\ Y_2 &= 0, t \in [0, t^*]; Y_2 = F_1, t \in [t^*, T]. \end{aligned} \quad (8.7)$$

При этом значение t^* зависит от интервала планирования T .

Экономика не терпит резких изменений, и решения типа (8.7) нельзя использовать на практике. Более того, их, вероятно, нельзя и реализовать. Дефект подобного анализа состоит в том, что формулировка задачи (8.6) не учитывает многообразие критериев, один из которых требует минимизации структурных перестроек. Этот факт может быть формализован следующим образом.

Вместо Y_1 введем новое управление u :

$$Y_1 = u F_1, Y_2 = (1 - u) F_1, u \in [0; 1].$$

Тогда минимизацию перестройки можно охарактеризовать следующим образом:

$$J^*(u) = \int_0^T (u - u_0)^2 dt \quad \min, \quad (8.8)$$

где u_0 — распределение инвестиций в последний год предыдущего планового периода.

Следовательно, базовая траектория — это решение конфликтной задачи с функционалами (8.6) и (8.8), т. е. ее выбор, помимо решения серии оптимизационных задач (необходимых для построения множества Парето), требует еще и неформального анализа. В самом деле, формальный анализ может дать только зависимость $J(J^*)$, которая изображена на рис. 3. Какую же точку взять на этой кривой, т. е. какое из оптимальных управлений, изображенных на рис. 4, следует принять в качестве базового? На этот вопрос с помощью формальной теории в рамках данной модели ответить нельзя.

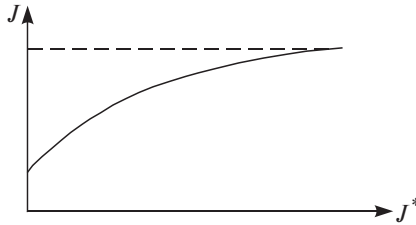


Рис. 3

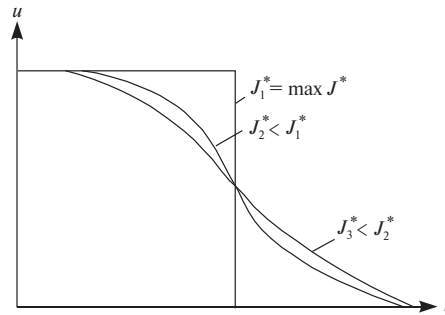


Рис. 4

9. Модель назначения штрафов за загрязнение окружающей среды

Еще один поучительный пример кибернетической системы со многими субъектами дают проблемы функционирования региона, на территории которого расположены предприятия, загрязняющие окружающую среду (например, воду).

Обозначим через y_i ($i = 1, 2, \dots, N$) фонды этих предприятий. Их изменение будем описывать следующими уравнениями:

$$\dot{y}_i = Y_i - k_i y_i, \quad (9.1)$$

где Y_i — инвестиции, а k_i — коэффициенты амортизации.

Каждое предприятие в единицу времени выдает продукцию

$$P_i = F_i(x_i), \quad (9.2)$$

где F_i — производственная функция.

Предприятия производят не только полезный, но и вредный продукт. Обозначим через v_i поток загрязняющих веществ:

$$v_i = f_i(P_i, V_i), \quad (9.3)$$

где V_i — затраты предприятий на усовершенствование технологии или очистку в заводских условиях.

Будем рассматривать ситуацию, когда в регионе кроме производителей существует еще один субъект кибернетической системы — региональное управление. Последнее имеет право облагать остальных субъектов штрафом

$$w_i = c \cdot v_i \quad (9.4)$$

и распоряжается величиной c . Все остальные управления Y_i и V_i находятся в распоряжении производителей.

Условимся, что инвестирование происходит только за счет внутренних средств предприятий и все свободные капиталы тратятся на инвестиции. Это значит, что

$$F_i(x_i) = Y_i - V_i - w_i. \quad (9.5)$$

Рассмотрим субъективное описание системы с точки зрения регионального управления. Чтобы принять какое-либо решение, последнее должно сделать определенные предположения о поведении остальных субъектов. Можно принять следующую гипотезу. Выбор величин V_i и w_i определяется условием

$$J_i = V_i - w_i \rightarrow \min. \quad (9.6)$$

Поскольку региональное управление обладает правом первого хода, т. е. может заранее сообщить остальным субъектам размеры штрафа, то гипотеза (9.6) позволяет определить реакцию производителей. В самом деле, так как

$$J_i = V_i - c_i - V_i \cdot cf_i(F_i(c_i), V_i),$$

то условие (9.6) позволит определить объем капиталовложений на усовершенствование технологии и внутризаводскую очистку:

$$V_i = V_i(c_i, j). \quad (9.7)$$

Предположим, что средства, полученные региональным управлением за счет штрафов, расходуются на централизованную очистку среды. Тогда уравнение, характеризующее изменение уровня загрязнения среды, можно записать в следующем виде:

$$\dot{x} = W - f(x) - W(w_i), \quad (9.8)$$

где W — количество вредных отходов, уничтожаемых централизованной очисткой. Положим $W = W(w_i)$, $f(x)$ — естественная очистка. Предполагая, что производители будут знать его политику штрафов, региональное управление в соответствии с проведенным анализом должно принять, что

$$w_i = cf_i[F_i(c_i), V_i(c_i, j)] = cf_i^*(c_i, j).$$

Тогда уравнение (9.8) переписывается в следующем виде:

$$\dot{x} = f_i^*(c_i, j)(1 - c) - f(x). \quad (9.9)$$

И, наконец, обсудим мотивы, которыми должно руководствоваться региональное управление при назначении штрафов. Прежде всего, оно должно стремиться к тому, чтобы качество среды не ухудшалось, т. е. чтобы

$$\dot{x} \geq 0 \text{ или } x = \min. \quad (9.10)$$

Но критерий (9.10) не единственный. Региональное управление заинтересовано также в промышленном развитии региона. Этот критерий можно формализовать различным образом, например, можно принять, что региональное управление стремится иметь устойчивую траекторию постоянного роста:

$$\frac{d}{dt} \dots_i \leq 0 \text{ или } \frac{d}{dt} \dots_i \geq \max. \quad (9.11)$$

Итак, мы имеем систему дифференциальных уравнений, зависящих от параметра c :

$$\frac{dx}{dt} = X(x_1, \dots, x_N, c), \quad (9.12)$$

$$\frac{d}{dt} \dots_i = \dots_i(x_1, \dots, x_N, c).$$

Параметр c нужно выбрать так, чтобы

$$\frac{dx}{dt} \geq \min, \quad (9.13)$$

$$\frac{d}{dt} \dots_i \geq \max. \quad (9.14)$$

Если $f(x)$ — вогнутая функция, а f_i — выпуклая функция V_i и вогнутая функция P_i , то характер зависимости \dots_i и x от c будет таким, как это изображено на рис. 5. Если $\hat{c} = c^*$, то любое значение c из интервала (\hat{c}, c^*) удовлетворяет условиям (9.10) и (9.11). Их дальнейший выбор носит уже неформальный характер, поскольку критерии (9.13) и (9.14) несовместимы.

Если $\hat{c} < c^*$, то удовлетворить одновременно условиям (9.10) и (9.11) невозможно.

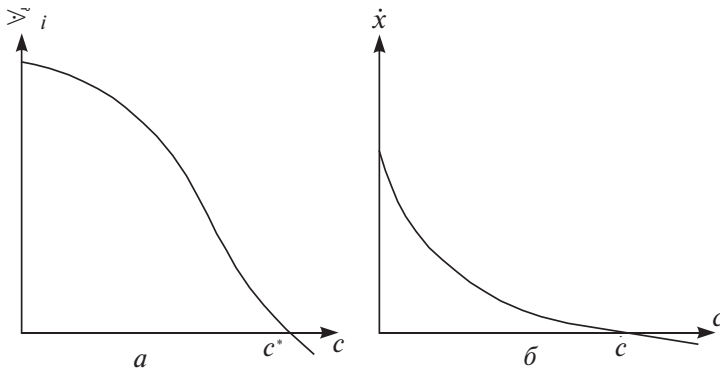


Рис. 5.

10. Минимальная модель эколого-экономической системы

Говорить об эколого-экономических системах имеет смысл лишь в том случае, когда речь идет об анализе их эволюции в течение достаточно больших интервалов времени. Это связано с тем, что характерные времена изменения таких систем (времена, через которые параметры экосистем могут заметно изменяться вследствие человеческой деятельности) обычно имеют порядок десятилетий, т. е. далеко выходят за рамки пятилетних планов. Это накладывает определенные требования на структуру моделей блока человеческой активности. Поэтому следует говорить о «минимальной модели» блока, т. е. перечислить те основные факторы, игнорирование которых может качественно исказить характер эволюции изучаемых величин.

Как уже было отмечено, в блоке человеческой активности целесообразно отделить те формы деятельности, описание которых основывается главным образом на законах сохранения и, следовательно, носит достаточно универсальный характер, от социальных механизмов, которые всегда индивидуализированы и отражают социальную природу общества.

К первой группе относятся модели: промышленного производства, сельскохозяйственного производства, эксплуатируемых биоценозов (лес, рыба), энергетики, ресурсов, демографии, научно-технического прогресса.

При построении минимальной модели можно использовать упрощенные, агрегированные модели блоков, но все блоки должны присутствовать.

Ко второй группе относятся следующие описывающие социальный фон модели: инвестирования, распределения благ, финансов, спроса и предложения рабочей силы и т. д.

Если речь идет о капиталистической экономике, то инвестирование — это рынок капитала, тесно связанный с моделями финансов и рынком рабочей силы. В модели социалистической экономики процесс инвестиции должен описывать некоторую программу и, следовательно, эта модель должна быть связана некоторой обратной связью с моделями, которые служат для определения целей и выработки программ.

Внешнюю среду в минимальной модели естественно задать в форме критериев ее качества I_1, I_2, \dots, I_s . Тогда совокупность критических значений параметров I_i^* определяет границу гомеостаза. Функция

$$F \min_i \frac{I_i}{I_i^*} \quad (10.1)$$

описывает расстояние до границы гомеостаза. Функционал (10.1) характеризует воздействие активности людей, и стремление $F \max$ является общим для всех субъектов системы.

Заключение

Цель данной работы — наметить пути создания языка описания нерефлексных, и в частности экономических, систем и в этом контексте представить содержание и место математических исследований.

Предложенная формализация достаточно универсальна. Она описывает ситуацию, которая возникает при оценке долгосрочных планов регионального развития и при решении разнообразных проблем международного экономического сотрудничества, а также, возможно, закладывает основу для построения математической модели сосуществования. Она позволяет также включить в единую систему процессы, происходящие в биоте и в экономическом организме.

Анализ нерефлексных систем необходимо требует эвристических подходов. Даже сама редукция исходной нерефлексной системы к рефлексной невозможна без дополнительных гипотез, т. е. без субъективного описания. Значит, анализ эволюции системы и процедуры принятия решений непременно предполагает человеко-машинный диалог.

Существует представление (возникшее прежде всего в США) о том, что возможности прямого общения «человек — машина», которое позволяет реализовывать машинные эксперименты, упрощает используемые математические методы. Мы с этим утверждением решительно не согласны. Человеко-машинный диалог не упрощает, а усложняет необходимый аппарат исследования. Он просто переводит эти исследования в другую сферу. Диалог — это тоже алгоритм. Алгоритм, снимающий неопределенность, выявляющий те объективные интересы и те представления, ради которых в конечном счете и принимаются решения. Задача организации диалога состоит в том, чтобы решения наилучшим образом соответствовали этим интересам, хотя их четкая формулировка нам недоступна. Но

человек может ответить лишь на относительно простые вопросы. Следовательно, математик должен так вести анализ, чтобы эксперт мог давать однозначные ответы. Если сделать это не удастся, то значит, модель требует расширения, учета новых фактов или аппарат, используемый при анализе, недостаточен для тех целей, которые ставит исследователь.

Литература

1. *Galbraith J.K.* Economics and Public Purpose. Mifflin, 1977.
2. *Гермейер Ю.Б., Ватель И.А.* Игры с иерархическим вектором интересов./Изв. АН СССР, Техническая кибернетика, 1974, № 3.
3. *Мусеев Н.Н.* Элементы теории оптимальных систем. — М., Наука, 1975.

1977 г.

Научное издание

**Никита
Николаевич
МОИСЕЕВ**

**Избранные
труды**

Гидродинамика и механика

**Оптимизация,
исследование операций
и теория управления**

Редактор Т. Репина
Корректор Е. Андреева
Технический редактор
О. Капелькин
Компьютерная верстка
Т. Максимова

Лицензия ИД № 05504 от 1 августа 2001 г.

ООО «Тайдекс Ко»
117311, Москва, ул. Строителей, д. 4, корп. 3

Подписано в печать 23.07.03. Формат 60 90/16. Бумага офсетная №1.
Гарнитура Ньютон. Печать офсетная. Печ. л. 23,5.
Заказ

Отпечатано с готовых диапозитивов в ФГУП ордена «Знак Почета»
Смоленской областной типографии им. В.И. Смирнова.
214000, Смоленск, пр-т им. Ю. Гагарина, 2.