

С 345Т

А-659

4915 / 79

АНДРИАНОВ С. Н. ИДР.

БИ-9-12851.



ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

БИ-9-12851

ДЕПОНИРОВАННАЯ ПУБЛИКАЦИЯ

Дубна 1979

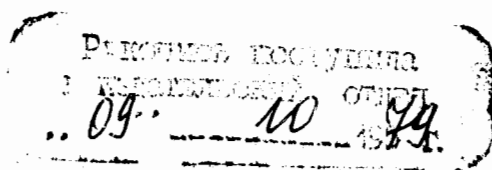
ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

Лаборатория нейтронной физики

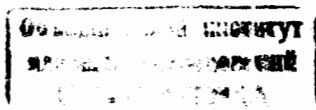
С.Н.Андреианов[†], А.Д.Дымников[†], Г.М.Осетинский[†]

Б1-9-12851

ОБ ОПТИМАЛЬНОМ УПРАВЛЕНИИ ПУЧКАМИ ЗАРЯЖЕННЫХ
ЧАСТИЦ



Дубна, 1979



А Н Н О Т А Ц И Я

Сформулирована задача оптимального управления пучками частиц, приведены теоремы существования оптимального управления. Задача оптимального управления решается сведением к задаче математического программирования. В качестве примера рассмотрена задача оптимизации нелинейного протонного микрозонда.

§ I. Основные понятия и утверждения

Прежде, чем сформулировать задачу оптимального управления (ОУ) пучками, введем некоторые вспомогательные понятия. Предварительно дадим пояснения, не прибегая к математической строгости, отсылая интересующихся к обширной литературе по функциональному анализу.

1° При формализации реальной физической ситуации остается некоторая неопределенность, которая должна быть уточнена математиками. К таким уточнениям следует отнести конкретизацию класса функций, с которым работает исследователь. С одной стороны этот класс должен определяться возможностью осуществления в физических системах, с другой стороны выбранный класс функций должен быть удобным для работы и позволять доказывать интересные утверждения. К такому классу функций, включающему в себя непрерывные, кусочно-непрерывные, ступенчатые и другие функции, относятся так называемые измеримые функции. Этот класс достаточно широк как с точки зрения приложений, так и с точки зрения математической теории. Важным свойством множества измеримых функций является его замкнутость относительно предельного перехода (то есть предел последовательности измеримых функций является также измеримой). Не вдаваясь в подробности можно сказать, что класс измеримых функций состоит из ступенчатых функций и пределов последовательностей, состоящих из ступенчатых функций. Об измеримой функции $f(t)$, заданной на интервале \mathcal{T} , будем говорить, что она задана почти всюду (п.в.) на \mathcal{T} .

Большую роль в функциональном анализе играет понятие функционального пространства, элементами которого служат функции, обладающие теми или иными свойствами. Например, пространство непрерывных функций, действующих из \mathcal{T} в евклидово пространство $R^n - C(\mathcal{T}, R^n)$, пространство интегрируемых с квадратом функций из \mathcal{T} в $R^n - L_2(\mathcal{T}, R^n)$ и т.д. В этих пространствах можно ввести метрику (расстояние между элементами), превращая их тем самым в метрические пространства. В пространстве $C(\mathcal{T}, R^n)$ вводится так называемая равномерная метрика:

$$\|f - g\|_C \stackrel{\text{def}}{=} \max_{t \in \mathcal{T}} |f(t) - g(t)|, \quad \forall f, g \in C(\mathcal{T}, R^n)$$

а в пространстве $L_2(\mathcal{T}, R^n)$:

$$\|f-g\| \stackrel{\text{дф}}{=} \int_{\mathcal{T}} \{f(t)-g(t)\}^2 dt, \quad \forall f, g \in L_2(\mathcal{T}, R^n)$$

Очень важным понятием является свойство компактности. В метрическом пространстве множество называется компактным, если любая последовательность элементов этого множества содержит сходящуюся последовательность. Сформулируем без доказательства две леммы.

Лемма I.1. Множество \mathcal{X} всех измеримых функций $f(t)$, действующих из \mathcal{T} в R^n и удовлетворяющих п.в. условию $|f(t)| < K, K > 0$, является компактным множеством в метрическом пространстве

Лемма I.2. Семейство непрерывных, кусочно-непрерывных, ступенчатых функций, действующих из \mathcal{T} в ограниченное множество из R^n , образуют всюду плотное подмножество из \mathcal{X} .

Замечание: Множество \mathcal{A} называется всюду плотным в \mathcal{B} , если его замыкание $\bar{\mathcal{A}}$ совпадает с \mathcal{B} , то есть $\bar{\mathcal{A}} = \mathcal{B}$.

2° Уравнение движения частиц в электромагнитном поле относится к дифференциальным уравнениям вида

$$x' = f(x, u, t) \quad (\text{I.1})$$

где x - n мерный вектор фазовых координат, u - τ мерный вектор управляющих воздействий (управлений), t - независимая переменная ("время").

- Условие I.1
1. Функция $f(x, u, t)$ удовлетворяет условию Липшица по переменным x, u и измерима по t .
 2. Существуют множества $V \subset R^n$ и $U \subset R^\tau$ и интегрируемая положительная функция $\Psi(t)$ такие, что $|f(x, u, t)| < \Psi(t), \forall x \in V, \forall u \in U$ п.в. на \mathcal{T} .

Замечание: Функция $f(t)$ удовлетворяет условию Липшица $\forall t \in \mathcal{T}$, если существует такая константа $L > 0$, что $|f(t_1) - f(t_2)| \leq L |t_1 - t_2|, \forall t_1, t_2 \in \mathcal{T}$.

Известно, что, если имеет место условие I.1, то решение задачи Коши для уравнения (I.1) непрерывно зависит от u равномерно по x_0 . Введем множество \mathcal{K} всех решений задач Коши для уравнения (I.1) при $x_0 \in M_0 \subset R^n$, где M_0 - начальное фазовое множество:

$$\mathcal{K} \stackrel{\text{дф}}{=} \left\{ x(\cdot) \in C(\mathcal{T}, R^n) \mid x(t) = x_0 + \int_{\mathcal{T}} f(x(\tau), u(\tau), \tau) d\tau, \forall x_0 \in M_0 \right\} \quad (\text{I.2})$$

Имеет место следующая лемма.

Лемма 1.3 Если функция $f(x, u, t)$ удовлетворяет условию I.1, M_0 — компактное множество в R^n , то X является полным метрическим пространством относительно равномерной метрики.

Замечание: Полным метрическим пространством называется пространство, в котором любая фундаментальная последовательность его элементов сходится к элементу этого пространства. Под фундаментальной последовательностью понимается последовательность, удовлетворяющая условию $|x_n - x_m| \rightarrow 0$, при $n, m \rightarrow \infty$

§ 2. Задача оптимального управления

I⁰ [Сформулируем задачу оптимального управления (OU) в общем виде. Уравнение

$$\dot{x} = F(x, u) \quad (2.1)$$

будем называть уравнением движения, здесь $F: X \times U \rightarrow X$, $X = C(\mathcal{T}, R^n)$, а U — некоторое, так называемое допустимое множество управлений, задаваемое с помощью следующего определения:

Определение 2.1 Семейство управляющих функций u , действующих из \mathcal{T} в R^2 называется допустимым, если

1. Для почти всех $t \in \mathcal{T}$ значение $u(t)$ принадлежат выпуклому, ограниченному и замкнутому множеству $U \in R^2$.
2. После подстановки u из U в правую часть (2.1) решение существует и единственно.
3. Семейство U является всюду плотным подмножеством в компактном множестве $\mathcal{L} \subset L_2(\mathcal{T}, R^2)$.

Следует отметить, что в качестве функций из U можно взять функции, упомянутые в формулировке леммы I.2.

[Определим функции $g_0: X \rightarrow R^1$, $g_1: X \times \mathcal{T} \rightarrow R^{m_1}$, $g_2: X \rightarrow R^{m_2}$, которые будем называть функционалом (или целевой функцией) и функциями, задающими ограничения (или просто ограничениями) соответственно. Введем три множества:

$$B(U) \stackrel{\text{def}}{=} \{ (x, u) \in X \times U \mid \dot{x} = F(x, u) \} \quad (2.2)$$

$$B_1(U) \stackrel{\text{def}}{=} \{ (x, u) \in B(U) \mid g_1(x(t), t) \in \mathcal{N}(t), g_2(x(t)) = 0 \} \quad (2.3)$$

$$B_2(U) \stackrel{\text{def}}{=} \{ (x, u) \in B(U) \mid g_1(x(t), t) \in \mathcal{N}_\delta(t), |g_2(x(t))| < \varepsilon \} \quad (2.4)$$

Множество $\mathcal{B}(\mathcal{U})$ представляет собой множество пар (x, u) , удовлетворяющих уравнению (2.1), а множества $\mathcal{B}_1(\mathcal{U})$ и $\mathcal{B}_2(\mathcal{U})$ множества пар (x, u) из $\mathcal{B}(\mathcal{U})$, удовлетворяющих указанным ограничениям. Множества $\mathcal{M}(t)$ и $\mathcal{N}_\delta(t)$ непрерывно зависят от $t \in \mathcal{T}$.

Задача оптимального управления Задачей ОУ будем называть задачу минимизации функционала g_0 на множествах $\mathcal{B}_1(\mathcal{U})$ или $\mathcal{B}_2(\mathcal{U})$.

Определение 2.2 Будем говорить, что задача ОУ обладает оптимальным \mathcal{U} -решением, если существует пара $(x_*, u_*) \in \mathcal{B}(\mathcal{U})$, которая минимизирует g_0 на множестве $\mathcal{B}_1(\mathcal{U})$:

$$g_0(x_*, u_*) = \inf g_0(\mathcal{B}_1(\mathcal{U})) \quad (2.5)$$

Если заменить \mathcal{U} на \mathcal{X} , то будем говорить об оптимальном \mathcal{X} -решении. Элементы $u_* \in \mathcal{U}$ и $e_* \in \mathcal{X}$ будем называть оптимальными \mathcal{U} - и \mathcal{X} -управлениями.

Определение 2.3 Последовательность $\{(x_k, u_k)\} \in \mathcal{B}(\mathcal{U})$, будем называть приближенным \mathcal{U} -решением, если, начиная с некоторого номера все элементы этой последовательности будут принадлежать $\mathcal{B}_2(\mathcal{U})$. Приближенное \mathcal{U} -решение (x_{*k}, u_{*k}) будем называть оптимальным приближенным \mathcal{U} -решением, если существует $\lim_{k \rightarrow \infty} g_0(x_{*k}, u_{*k})$

и
$$\lim_{k \rightarrow \infty} g_0(x_{*k}, u_{*k}) \leq \lim_{k \rightarrow \infty} g_0(x_k, u_k)$$

для любого приближенного \mathcal{U} -решения $\{(x_k, u_k)\}$.

Без доказательства приведем утверждения, дающие условия существования оптимального \mathcal{X} -решения и оптимального приближенного \mathcal{U} -решения.

Лемма 2.1 Пусть множество $\mathcal{N}(t)$ замкнуто $\forall t \in \mathcal{T}$, $\mathcal{B}_1(\mathcal{X})$ не пусто, $\mathcal{B}_1(\mathcal{X})$ компактно, сужение функций F, g_i ($i = \overline{0, 2}$) на множество $\mathcal{B}_1(\mathcal{X})$ непрерывны. Тогда $\mathcal{B}_1(\mathcal{X})$ компактно и существует оптимальное \mathcal{X} -решение (x_*, e_*) .

Заметим, что если в формулировке леммы 2.1 заменить $\mathcal{N}(t)$ на $\mathcal{N}_\delta(t)$ и $\mathcal{B}_1(\mathcal{X})$ на $\mathcal{B}_2(\mathcal{X})$, то множество $\mathcal{B}_2(\mathcal{X})$ также будет компактным.

Лемма 2.2 Предположим, что существует оптимальное \mathcal{X} -решение (x_*, e_*) , \mathcal{U} - допустимое множество и x_* единственное решение уравнения $x = F(x, e_*)$. Тогда существует оптимальное приближенное \mathcal{U} -решение $\{(x_{*k}, u_{*k})\}$.

Приведенные леммы дают условия существования оптимальных управлений для уравнения движения (2.1). Эти утверждения являются вспомогательными для теории существования оптимальных управлений пучками частиц.

§ 3. Оптимальное управление пучками частиц

Обычная теория ОУ имеет дело с оптимизацией вдоль отдельной траектории фазовой точки. В физике пучков приходится иметь дело с ансамблем частиц, каждая из которых описывает свою траекторию в фазовом пространстве. Поэтому возникает вопрос, при каких условиях может существовать оптимальное управление для всего пучка в целом. В данном параграфе будут приведены утверждения, отвечающие на этот вопрос.

1° Возвращаясь к уравнению движения частиц в электромагнитном поле (I.1), запишем его в интегральном виде и получим

$$F(x, u)(t) \stackrel{\text{df}}{=} x_0 + \int_{t_0}^t f(x(\tau), u(\tau), \tau) \cdot d\tau, \quad \forall t \in \mathcal{T} \quad (3.1)$$

Пусть M_0 выпуклое, замкнутое и ограниченное множество в R^n (следовательно и компактное), тогда в силу уравнения (I.1) в каждый момент времени $t \in \mathcal{T}$ будет определено множество

$$M(M_0, u, t) \stackrel{\text{df}}{=} \left\{ x(t) \mid x(t) = x_0 + \int_{t_0}^t f(x(\tau), u(\tau), \tau) d\tau, x_0 \in M_0 \right\}, \quad (3.2)$$

где $u \in \mathcal{U}$ — допустимое множество. Определим расстояние между множествами как

$$P_*(M_1, M_2) \stackrel{\text{df}}{=} \max \left\{ \sup_{x \in M_1} \rho(x, M_2), \sup_{x \in M_2} \rho(M_1, x) \right\}, \quad (3.3)$$

здесь $\rho(x, M) \stackrel{\text{df}}{=} \inf_{x_1 \in M} \rho(x, x_1)$

Имеет место следующая теорема

Теорема 3.1 Пусть $f(x, u, t)$ удовлетворяет условию I.1, тогда $M(M_0, u, t)$ непрерывно зависит от своих аргументов.

Доказательство теоремы 3.1 основано на непрерывной зависимости решений уравнения (I.1) от начальных условий и параметров при предположениях теоремы.

2° Введем так называемое отношение эквивалентности в пространстве состояний $X = C(\mathcal{T}, R^n)$.

Определение 3.1 Будем говорить, что $X_1(t)$ эквивалентно $X_2(t)$ $\forall t \in \mathcal{T}$, если $X_1(t), X_2(t) \in \mathcal{M}(M_0, u, t)$, $u \in \mathcal{U}$.

Иными словами мы отождествляем все траектории, выходящие из одного и того же начального множества M_0 и при одном управлении $u \in \mathcal{U}$. Все такие траектории образуют класс эквивалентности, который мы будем называть пучком. Множество всех таких классов эквивалентности, при пробегании u всего допустимого множества \mathcal{U} , снабженное метрикой

$$\hat{\rho}(y_1, y_2) \stackrel{\text{def}}{=} \sup_{t \in \mathcal{T}} \rho_*(M_1(t), M_2(t)) \quad (3.4)$$

будем называть фактор-пространством или пространством пучков \mathcal{Y} .
Здесь y_i элементы пространства \mathcal{Y} , им однозначно соответствует множество M_i (в (3.4) зависимость от M_0 и \mathcal{U} опущена). Можно показать, что отображение F , задаваемое формулой (3.1), индуцирует в фактор-пространстве \mathcal{Y} непрерывное отображение $\mathcal{F}: \mathcal{Y} \times \mathcal{U} \rightarrow \mathcal{Y}$, таким образом в пространстве пучков \mathcal{Y} мы получаем уравнение движения, аналогичное уравнению (2.1):

$$\dot{y} = \mathcal{F}(y, u) \quad (3.5)$$

с непрерывной правой частью. Введен функционал и ограничения, действующие уже в новом пространстве \mathcal{Y} , следующим образом:

$$\tilde{g}_0(y(t)) = \int_{M(t)} g_0(x) dx \quad (3.6)$$

$$\tilde{g}_1(y(t), t) = \sup_{x \in M(t)} g_1(x, t), \quad \tilde{g}_2(y(t)) = \sup_{x \in M(t)} g_2(x) \quad (3.7)$$

Здесь элементу $y \in \mathcal{Y}$ соответствует множество M . С помощью этих функций можно определить множества $\mathcal{B}(u), \tilde{\mathcal{B}}_1(u), \tilde{\mathcal{B}}_2(u)$, которые задаются также как и множества $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2$, если в формулах (2.2)-(2.4) заменить пространство \mathcal{X} на \mathcal{Y} и функции F, g_i на функции \mathcal{F}, \tilde{g}_i ($i=0, 2$). При сделанных предположениях справедливы следующие теоремы.

Теорема 3.2 Пусть множество $\mathcal{N}(t)$ замкнуто $\forall t \in \mathcal{T}$, $\tilde{\mathcal{B}}_2(\mathcal{X})$ не пусто, функция $f(x, u, t)$ удовлетворяет условию I.1, \tilde{g}_i непрерывны по своим аргументам ($i=0, 2$), \mathcal{F} - непрерывное отображение, индуцируемое отображением (3.1). Тогда множество $\tilde{\mathcal{B}}(\mathcal{X})$ компактно и существует оптимальное \mathcal{X} -решение (y_*, u_*) .

Здесь y_* - оптимальный элемент в пространстве пучков \mathcal{Y} , что соответствует существованию оптимального пучка в пространстве состояний \mathcal{X} .

Если заменить в формулировке теоремы 3.2 множество $\mathcal{N}(t)$ на $\mathcal{N}_\delta(t)$ и множество $\tilde{\mathcal{B}}(\mathcal{X})$ на $\tilde{\mathcal{B}}_2(\mathcal{X})$, то множество $\tilde{\mathcal{B}}_2(\mathcal{X})$ будет также компактным.

Теорема 3.3 Предположим, что существует оптимальное \mathcal{X} -решение (y_*, e_*) , \mathcal{U} -допустимое множество, y_* - единственное решение уравнения $y = F(y, e_*)$, тогда существует оптимальное приближенное \mathcal{U} -решение $\{(y_{*k}, u_{*k})\}$.

Доказательство этих двух теорем вытекает из лемм 2.1 и 2.2, а также теоремы 3.1.

Предположения об условиях существования оптимального управления пучком (теоремы 3.2 и 3.3) имеют место для большинства практических задач.

§ 4. Решение задач ОУ методом сведения их к задачам математического программирования

Прежде, чем описать способ сведения задач ОУ к задачам математического программирования (МП), сформулируем задачу МП в общем виде.

Задача математического программирования Минимизировать функцию $\Phi_0(w)$ при условиях $\Phi_1(w) \leq 0, \Phi_2(w) = 0$, или кратко

$$\inf \{ \Phi_0(w) \mid \Phi_1(w) \leq 0, \Phi_2(w) = 0 \} \quad (4.1)$$

Очевидно, что поиск \inf не ограничивает общности задач, так как изменяя знак у функции Φ_0 , мы получаем задачу поиска \sup . То же самое относится и к ограничениям в виде неравенств $\Phi_1(w) \leq 0$ ($-\Phi_1(w) \geq 0$). Если хотя бы одна из функций Φ_i ($i = \overline{0, 2}$) является нелинейной, то соответствующая задача МП называется задачей нелинейного программирования.

1° [В основе метода, позволяющего сформулировать задачу ОУ как задачу МП, лежит известный метод Рунца [1], который заключается в аппроксимации исходной задачи ОУ последовательностью задач минимизации функционала, зависящего от конечного числа параметров. А именно, управление $u \in \mathcal{U}$ ищется в виде комбинаций функций $\{z^k\}$, составляющих так называемую полную систему в допустимом множестве \mathcal{U} .] В качестве

такой системы могут быть выбраны функции, указанные в формулировке леммы 1.2. Если в качестве функций $\{z^k\}$ выбрать кусочно-постоянные функции, то вместо искомой функции мы будем иметь конечно-мерный вектор варьируемых параметров w , первые j компонент которого определяют высоту ступенек, а вторые $j-1$ длину этих ступенек (полная длина T интервала \mathcal{T} фиксирована), где j - число разбиений интервала \mathcal{T} . Оказывается, что при увеличении $j \rightarrow \infty$ мы получаем сходимость последовательности решений, получаемых конечномерных задач к решению исходной задачи ОУ. В случае наличия ограничений g_1 и g_2 существенную помощь при реализации задачи МП оказывает то обстоятельство, что в действительности необходимо проверять выполнение ограничений не для каждого $t \in \mathcal{T}$, а, например, в некоторые фиксированные моменты $t_k \in \mathcal{T}$. Выберем разбиение интервала \mathcal{T} таким образом, чтобы моменты t_k являлись одними из узловых точек этого разбиения, ограничения в этом случае примут вид

$$g_1(x(t_k), t_k) \leq 0, \quad g_2(x(t_k)) = 0, \quad (4.2)$$

что позволяет написать ограничения на параметры $\Phi_1(w) \leq 0, \Phi_2(w) = 0$. Аналогичную процедуру можно проделать с ограничениями на весь пучок (3.7). Итак, метод Ритца позволяет строить сходящуюся последовательность решений задач МП. Исходя из специфики физической задачи, часто оказывается возможным получить хорошее приближение к оптимальному решению задачи ОУ при не очень большом числе разбиений интервала \mathcal{T} . Выбор функций, составляющих полную систему, определяется возможностью их реализации в физической модели задачи.

2° В общем виде функционал \tilde{J}_0 запишем следующим образом:

$$\tilde{J}_0[y] = \alpha_0 \int_{\mathcal{T}_1} \tilde{G}_0(y(t), t) dt + \sum_{k=1}^m \alpha_k \tilde{G}_k(y(t_k)) \quad (4.3)$$

где α_i - весовые коэффициенты, $\mathcal{T}_1 \subseteq \mathcal{T}$ и $y \in \mathcal{Y}$. Здесь первое слагаемое представляет собой интегральный критерий качества, а второе - сумму функциональных критериев, заданных в фиксированных точках интервала \mathcal{T} . В соответствии с (3.6) функции \tilde{G}_0, \tilde{G}_k имеют вид:

$$\tilde{G}_0(y(t), t) = \int_{\mathcal{M}(\mathcal{M}_0, \mathcal{U}, t)} G_0(x, t) dt \quad (4.4)$$

$$\tilde{G}_k(y(t_k)) = \int_{\mathcal{M}(\mathcal{M}_0, \mathcal{U}, t_k)} G_k(x) dx \quad (4.5)$$

Здесь G_0, G_k - положительные, интегрируемые функции относительно аргумента X , G_0 по t - измерима на \mathcal{T} .

Например, для задачи фокусировки функционал (4.3) принимает простой вид

$$\tilde{q}_0[y] = \int_{M_{\text{вых}}} G(x) dx, \quad (4.6)$$

где $M_{\text{вых}}$ - выходное фазовое множество, занимаемое пучком и $G(x)$ функция фазового распределения частиц. Для задачи согласования фазовых множеств функционал \tilde{q}_0 принимает вид (4.6), но функция $G(x)$ имеет другой смысл, а именно

$$G(x) = \begin{cases} 1 & \forall x \in M_{\text{вых}} \setminus M_* \\ 0 & \forall x \in M_{\text{вых}} \cap M_* \end{cases}, \quad (4.7)$$

где M_* - множество, согласования с которым необходимо достигнуть. Иными словами мы хотим добиться выполнения условия $M_{\text{вых}} \subset M_*$.

Если предположить, что функции G_0, G_k удовлетворяют условию Липшица, то есть

$\|G_0(x_1, t) - G_0(x_2, t)\|_c \leq L_0 \cdot \|x_1 - x_2\|_c$ $\|G_k(x_1) - G_k(x_2)\| \leq L_k \cdot \|x_1 - x_2\|_c$,
то для функционала $q_0[u] \stackrel{df}{=} \tilde{q}_0[y(u)]$, порождаемого (4.3), справедливо неравенство

$$\|q_0(u_1) - q_0(u_2)\|_c \leq \left\{ \alpha \cdot L_0 + \sum_{k=1}^m \alpha_k \cdot L_k \right\} \cdot K \cdot \|u_1 - u_2\|_{L_2} \quad (4.8)$$

где $K = \frac{L_u}{\sqrt{L_x}} \left\{ \exp(L_x T) \frac{\sqrt{\pi}}{2} \right\}$, $u_i \in \mathcal{U}$, L_x, L_u - постоянные Липшица для функции $f(x, u, t)$. Аналогичные неравенства имеют место и для ограничений, задаваемых функциями (3.7)

$$q_1(u, t) \stackrel{df}{=} \tilde{q}_1(y(u), t), \quad q_2(u) \stackrel{df}{=} \tilde{q}_2(y(u)) \quad (4.9)$$

При переходе к задаче МП вместо функционала q_0 и ограничений q_i мы получаем целевую функцию Φ_0 и ограничения Φ_i , зависящих от конечномерного вектора варьируемых параметров w . Так как имеет место неравенство $\|u_1 - u_2\|_{L_2} \leq Q^{-1} |w_1 - w_2|$, $Q > 0$ то очевидно, что и целевая функция Φ_0 и ограничения Φ_i удовлетворяют условию Липшица с некоторыми константами.

[2]

§ 5. Метод скользящего допуска

К настоящему времени разработано и исследовано на сходимость довольно много методов минимизации функций многих переменных при

наличии ограничений. При оценке метода важное значение имеет не столько скорость сходимости метода, сколько общий объем вычислений, требуемый для получения решений с заданной точностью. А общий объем вычислений зависит не только от числа итераций, но и от трудоемкости вычислений на каждой итерации. Часто бывает более выгодно применять метод, который сходится медленнее (требует большего числа итераций), но расчет каждого шага осуществляется просто и требует небольшого числа машинных операций. Кроме того, необходимо учитывать устойчивость метода по отношению к ошибкам, требуемый объем памяти, удобство реализации на ЭВМ и т.д. Все методы минимизации можно разделить на две группы: использующие вычисления производных (градиентные) и не использующие (прямые методы).

В данной работе выбран метод скользящего допуска (СД) при реализации которого используется один из методов минимизации, не использующий вычисления производных, так называемый метод деформируемого многогранника (МДМ). Выбор метода СД связан с тем, что он позволяет улучшать значения целевой функции за счет информации, получаемой не только в допустимых точках пространства параметров, но и в точках, лежащих вне допустимой области, однако лежащих в некотором смысле близко к ней. Поскольку на первых шагах поиска ограничения задачи должны выполняться весьма приблизительно, и лишь при поиске в непосредственной близости от искомого решения требуется большая точность, то полный объем вычислений в процессе оптимизации по сравнению с другими методами существенно сокращается.

Г⁰ Изложим вкратце метод СД и сформулируем условия его сходимости. Введем функцию $\Psi(w)$:

$$\Psi(w) \stackrel{\text{def}}{=} \left\{ (\Phi_2, \Phi_2) + \sum_{k=1}^{m_1} \chi(\Phi_1^k) \right\}, \quad (5.1)$$

здесь (Φ_2, Φ_2) скалярное произведение вектор-функции Φ_2 , а χ функция Хевисайда:

$$\chi(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

и Φ_1^k - k-ая компонента вектор-функции Φ_1 . Очевидно, что условие $\Psi(w) = 0$ и условия $\Phi_1(w) \leq 0$, $\Phi_2(w) = 0$ эквивалентны. Введем функцию $\Psi(w, \varepsilon) = \Psi(w) - \varepsilon$, тогда условие $\Psi(w, \varepsilon) = 0$ будет эквивалентно ослабленным ограничениям вида

$$\Phi_1(w) \leq \delta_1(\varepsilon), \quad |\Phi_2(w)| \leq \delta_2(\varepsilon).$$

Пусть задана последовательность $\{\varepsilon_k\} \xrightarrow{k \rightarrow \infty} 0$, ограничение $\Psi(w) = 0$, задающее множество \mathcal{W} в пространстве параметров, тогда ограничения $\Psi(w, \varepsilon) = 0$ задают множества $\mathcal{W}_{\varepsilon_k}$, при этом имеет место последовательность включений:

$$\mathcal{W}_{\varepsilon_0} \supset \mathcal{W}_{\varepsilon_1} \supset \dots \supset \mathcal{W}_{\varepsilon_k} \supset \dots \supset \mathcal{W}$$

Заменим исходную задачу минимизации функции $\Phi_0(w)$ на множестве \mathcal{W} на последовательность задач минимизации $\Phi_0(w)$ на множествах $\mathcal{W}_{\varepsilon_k}$. Пусть $\{w_{*k}\}$ последовательность решений таких задач, то есть

$$\Phi_0(w_{*k}) = \inf_{w \in \mathcal{W}_{\varepsilon_k}} \Phi_0(w) \quad , \quad \lim_{k \rightarrow \infty} \varepsilon_k = 0 \quad (5.2)$$

Метод построения последовательности $\{w_{*k}\}$ и носит название метода скользящего допуска. Можно показать, что если функции Φ_i удовлетворяют условиям Липшица, то и $\Psi(w)$ обладает тем же свойством.

Тогда имеет место следующая теорема (без доказательства):

Теорема 5.1 Пусть $\mathcal{W}_{\varepsilon_0}$ компактное множество в пространстве варьируемых параметров R^p (p – размерность вектора w), Φ_0, Φ_1, Φ_2 удовлетворяет условию Липшица, тогда метод скользящего допуска сходится, то есть

$$\lim_{k \rightarrow \infty} \Phi_0(w_{*k}) = \Phi_0^* = \inf_{w \in \mathcal{W}} \Phi_0(w) \quad , \quad \lim_{k \rightarrow \infty} w_{*k} = w_*$$

где

$$\Phi_0(w_*) = \inf_{w \in \mathcal{W}} \Phi_0(w)$$

Множество $\mathcal{W}_{\varepsilon_0}$ можно определить как множество в R^p , полученное при помощи линейных и квадратичных ограничений на пределы изменения значений вектора параметров w .

2^o Рассмотрим вопрос о построении последовательности $\{w_{*k}\}$. Процесс построения последовательности $\{w_k\}$ будем называть релаксационным, если на множестве \mathcal{W} имеет место неравенство:

$$\Phi_0(w_{k+1}) \leq \Phi_0(w_k) \quad , \quad w_k, w_{k+1} \in \mathcal{W} \quad , \quad k = 0, 1, 2, \dots$$

В качестве метода, решающего задачу минимизации на множествах $\mathcal{W}_{\varepsilon_k}$ выбран метод деформируемого многогранника, который является релаксационным методом. Известно, что для ограниченной снизу функции $\Phi_0(w)$ любой релаксационный метод сходится, следовательно, МДМ является сходящимся методом для ограниченных снизу $\Phi_0(w)$. В процессе решения задачи минимизации на некотором множестве $\mathcal{W}_{\varepsilon_k}$ может возникнуть ситуация, когда текущее значение параметра $\tilde{w} \in \mathcal{W}_{\varepsilon_k}$,

то есть $\Psi(\tilde{w}, \varepsilon_k) > 0$. Такие точки будем называть недопустимыми. Для продолжения процесса поиска минимума необходимо из недопустимой точки перейти к почти допустимой (то есть принадлежащей $\mathcal{W}_{\varepsilon_k}$), либо к допустимой (принадлежащей \mathcal{W}). Можно доказать следующее утверждение.

Лемма 5.1 Если множество $\mathcal{W}_{\varepsilon}$ не пусто и удовлетворяет условию регулярности Слейтера, то МДМ за конечное число шагов из недопустимой точки приводит к почти допустимой точке из $\mathcal{W}_{\varepsilon}$.

Замечание: Говорят, что множество $\mathcal{W}_{\varepsilon}$ удовлетворяет условию регулярности Слейтера, если существует хотя бы одна точка $w \in \mathcal{W}_{\varepsilon}$ такая, что $\Psi(w, \varepsilon) < 0$.

Так как согласно лемме 5.1 за конечное число шагов мы из недопустимой точки переходим к почти допустимой, то не теряя общности можно предположить, что вся последовательность $\{w_k^i\} \in \mathcal{W}_{\varepsilon_k}$. Таким образом, метод ДМ позволяет строить последовательность $\{w_k^i\}$ такую, что

$$\lim_{i \rightarrow \infty} w_k^i = w_{*k} \quad \text{и} \quad \Phi_0(w_{*k}) = \lim_{i \rightarrow \infty} \Phi_0(w_k^i) = \inf_{w \in \mathcal{W}_{\varepsilon_k}} \Phi_0(w)$$

Пусть m_1 - число ограничений в виде равенств, m_2 - в виде неравенств, тогда, если задать ε_k формулой

$$\varepsilon_k = \min \left\{ \varepsilon_{k-1}, \frac{m_1+1}{m_2-m_1+1} \sum_{j=1}^{m_2-m_1+1} \|w_{k-1}^j - w_k^{y.t.}\| \right\} \quad (5.3)$$

здесь $w_k^{y.t.}$ - точка в R^p , соответствующая центру тяжести многогранника, рассматриваемого в методе ДМ, w_{k-1}^j - координаты j -ой вершины этого многогранника на $k-1$ шаге и положить

$$\varepsilon_0 = 2(m_1+1)d \quad (5.4)$$

где d - параметр, определяющий размеры начального многогранника, то мы получим последовательность

$$\varepsilon_0 \geq \varepsilon_1 \geq \dots \geq \varepsilon_k \geq \dots \geq 0$$

При этом оказывается очень удобно использовать величину ε_k как критерий окончания поиска, то есть задавая $\delta > 0$, процесс будем заканчивать при выполнении неравенства $\varepsilon_k < \delta$, в этом случае выполняются следующие условия:

$$\begin{aligned} \Phi_0(w_k^j) &\leq \Phi_0(w_* \pm \sqrt{\delta}) \\ \Psi(w) &< \delta \end{aligned}$$

Таким образом величина δ будет определять как точность достижения минимума целевой функции, так и точность выполнения ограничений. Следует отметить, что задание последовательности $\{\epsilon_k\}$ формулой (5.3) обеспечивает улучшение (уменьшение) значения целевой функции только при приближении к допустимой области.

Заметим, что метод деформируемого многогранника не является единственным методом, позволяющим осуществлять стратегию метода СД. Но простота его реализации, не чувствительность к мелким неровностям рельефа целевой функции, достаточно хорошая сходимость, отсутствие необходимости вычисления производных (что сильно осложнило бы вычисления для большинства практических задач) и другие его свойства позволяют считать его одним из наиболее эффективных.

§ 6. Метод погружения в пространство фазовых моментов

Введем обозначение

$$f_{,k}(x,t) \stackrel{\text{df}}{=} \frac{\partial^{|\kappa|} f(x,t)}{\partial x_1^{\kappa_1} \dots \partial x_n^{\kappa_n}}, \quad (6.1)$$

где $\kappa = \{\kappa_1, \dots, \kappa_n\}$ мультииндекс, $|\kappa| = \sum_{i=1}^n \kappa_i$, n -размерность фазового вектора X .

Условие 6.1 Функция $f(x,t)$ бесконечно дифференцируема по переменной X , измерима по t , существуют $V \subset R^n$ и положительная суммируемая функция $\varphi(t)$ такая, что $|f_{,k}(x,t)| < \varphi(t)$, почти всюду на \mathcal{T} и $\forall x \in V \subset R^n$.

Обозначим $(\kappa)! \stackrel{\text{df}}{=} \kappa_1! \dots \kappa_n!$. Имеет место следующая лемма.

Лемма 6.1 Если $f(x,t)$ удовлетворяет условию 6.1, $f(0,t) \equiv 0$, то $\forall x \in V \subset R^n$ имеет место разложение

$$f(x,t) = \sum_{\kappa=1}^{\infty} \frac{1}{(\kappa)!} (x, \nabla)^{\kappa} f(0,t) \quad (6.2)$$

Здесь ∇ - градиент в пространстве R^n и (\quad, \quad) - скалярное произведение в R^n .

Дано дифференциальное уравнение

$$x' = f(x,t) \quad (6.3)$$

Следствием леммы 1.3 является утверждение следующей леммы.

Лемма 6.2 Множество решений уравнения (6.3) с правой частью

удовлетворяющей условию 6.1, образует полное метрическое пространство относительно равномерной метрики.

Согласно утверждению леммы 6.1 имеет место разложение

$$f(x, t) = \sum_{|k|=1}^{\infty} \frac{(x)^{|k|}}{(k)!} f_{,k}(0, t) = \sum_{|k|=1}^{\infty} \frac{x_1^{k_1} \dots x_n^{k_n}}{k_1! \dots k_n!} f_{,k}(0, t) \quad (6.4)$$

Введем обозначение

$$f_m(x, t) = \sum_{|k|=1}^m \frac{x_1^{k_1} \dots x_n^{k_n}}{k_1! \dots k_n!} f_{,k}(0, t) = \sum_{k=1}^m \Pi_{1k}(t) z_k, \quad (6.5)$$

здесь $\Pi_{1j}(t)$ матрица $n \times q(k)$, составленная из элементов вида $\frac{1}{k_1! \dots k_n!} f_{,k}(0, t)$, а вектор z_k — из членов разложения $(x_1 + \dots + x_n)^{|k|}$.
Здесь $q(k) = \binom{n}{k_1, \dots, k_n}$ — число членов в разложении $(x_1 + \dots + x_n)^{|k|}$. Принимая во внимание (6.5), уравнение (6.1) запишется в виде

$$x' = \sum_{k=1}^N \Pi_{1k}(t) z_k + \overline{\Xi}_{N+1}(x, t), \quad (6.6)$$

где $\overline{\Xi}_{N+1}(x, t)$ остаточный член в разложении (6.4), содержит члены $N+1$ -го порядка по x и выше. Вектор z_k будем называть вектором фазовых моментов k -го порядка. Вычислим производную z_k по t , получим

$$\frac{d}{dt} z_k = \sum_{j=k}^N \Pi_{kj}(t) z_j + \overline{\Xi}_{N+1}'(x, t) \quad (6.7)$$

Вводя вектор $Y_N = \begin{Bmatrix} z_1 \\ \vdots \\ z_N \end{Bmatrix}$ и объединяя (6.6) и (6.7), получим

$$Y_N'(t) = P_N(t) Y_N(t) + \overline{\Xi}_{N+1}^2(x, t) \quad (6.8)$$

где $\{P_N(t)\}_{ij} = \Pi_{ij}(t)$, $i, j = 1, \dots, N$, $\Pi_{ij}(t) \equiv 0$, для $i > j$

Итак, вместо нелинейного уравнения (6.3) мы получили последовательность линейных уравнений

$$Y_N'(t) = P_N(t) Y_N(t), \quad (6.9)$$

эквивалентных уравнению (6.3) с точностью до членов N -го порядка относительно переменной x . Уравнения (6.9) действуют в пространстве фазовых моментов, размерность которого равна $\sum_{j=1}^N q(j)$.

Решение уравнения (6.9) запишем с помощью матрицанта

$$Y_N(t) = R_N(t|t_0) Y_N(t_0), \quad (6.10)$$

где $R_N(t_0|t_0) = E$ - единичная матрица порядка $\sum_{j=1}^N q(j)$. Матрицант R_N имеет верхнедиагональный вид:

$$R_N = \begin{Bmatrix} R^{11} & R^{12} & \dots & R^{1N} \\ 0 & R^{22} & \dots & R^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & R^{NN} \end{Bmatrix} \quad (6.11)$$

здесь $R^{ij}(t|t_0)$ - матрицант линейного уравнения $x' = \Pi_{ij}(t)x$. После громоздких выкладок можно показать, что для диагональных блоч-матриц R^{jj} ($j = \overline{1, N}$) имеет место выражение

$$\left\{ R^{jj} \right\}_{k\ell} = \begin{cases} \sum_{i=1}^j \{ R^{ii} \}_{k_1 \ell_1} \dots \{ R^{ii} \}, & \text{в случае, если среди компонент} \\ \{ R^{ii} \}_{k_1 \ell_1} \dots \{ R^{ii} \}_{k_n \ell_n}, & \text{мультииндекса } \ell \text{ есть различные} \\ & \text{все компоненты мультииндекса } \ell \\ & \text{равны друг другу} \end{cases} \quad (6.12)$$

Для недиагональных R^{ij} , $i < j$ можно записать:

$$R^{ij}(t|t_0) = \sum_{\ell=i+1}^j \int_{t_0}^t R^{ii}(t|\tau) P^{i\ell}(\tau) R^{\ell j}(\tau|t_0) d\tau \quad (6.13)$$

Вводя обозначение

$$Q_N(t|t_0) = \{ R^{11}(t|t_0), \dots, R^{NN}(t|t_0) \}, \quad (6.14)$$

можно записать

$$X_N(t) = Q_N(t|t_0) Y_N(t_0), \quad (6.15)$$

где $X_N(t)$ - вектор, составленный из первых n компонент вектора $Y_N(t)$. Описанный метод построения последовательности $X_N(t)$ и составляет суть метода погружения в пространство фазовых моментов. Имеет место следующая теорема.

Теорема 6.1 Пусть для функции $f(x, t)$ справедливо условие 6.1, тогда последовательность $X_N(t)$, определяемая равенством (6.15) сходится к решению уравнения (6.3) равномерно по $t \in \mathcal{T}$ и $x_0 \in \mathcal{U}_0$.

Доказательство:

Доказывать утверждение теоремы будем в два этапа. На первом этапе докажем, что последовательность X_N является фундаментальной,

а на втором, применяя лемму 6.2, докажем утверждение теоремы. Введем обозначение

$$L \stackrel{\text{df}}{=} \max_{t, \tau \in \mathcal{T}} \max_{1 \leq i \leq n} \sum_{j=1}^n |\{R''(t|\tau)\}_{ij}| < +\infty \quad (6.16)$$

Следует отметить, что все величины безразмерны (все длины отнесены к некоторой эталонной длине, в качестве которой можно, например, выбрать длину системы). Из формул (6.12), учитывая обозначение (6.16) получим оценку

$$\|R^{jj}(t|t_0)\| \leq j \cdot L^j \quad (6.17)$$

Из определения матриц $\Pi_{ij}(t)$, а также из условия 6.1 следует, что

$$\|\Pi_{ij}(t)\| \leq \frac{\varphi(t)}{(j-1)!} \quad (6.18)$$

Вводя обозначение

$$M \stackrel{\text{df}}{=} \int_{\mathcal{T}} \varphi(t) dt, \quad J(L) \stackrel{\text{df}}{=} \prod_{k=1}^{\infty} \left(\frac{k+1}{k!} L^k + 1 \right) \quad (6.19)$$

и используя неравенства (6.17), (6.18), с помощью метода математической индукции получим

$$\|R^{ij}(t|t_0)\| \leq \frac{j}{(j-1)!} L^{j+1} \cdot M A^j \cdot A^{-2} \cdot J(L), \quad (6.20)$$

где

$$A = \begin{cases} 1, & \text{при } ML \leq 1 \\ \frac{1}{ML}, & \text{при } ML \geq 1 \end{cases}$$

Из (6.14), (6.15) и (6.20) получим ($N_2 > N_1$)

$$\|X_{N_1} - X_{N_2}\| \leq \sum_{j=N_1+1}^{N_2} \|R^{ij}(t|t_0)\| \cdot \|Z_j(t_0)\| \leq A^2 \cdot M L \cdot J(L) \cdot \sum_{j=N_1+1}^{N_2} \frac{j}{(j-1)!} (A L \cdot \tau)^j \quad (6.21)$$

здесь $\|X_0\| \leq \tau$ (в силу ограниченности \mathcal{M}_0). Так как при $N_1, N_2 \rightarrow \infty$ из (6.21) следует, что $\|X_{N_1} - X_{N_2}\| \rightarrow 0$, то последовательность X_N является фундаментальной (заметим, что сходимость к нулю является равномерной относительно $X_0 \in \mathcal{M}_0$ и $t \in \mathcal{T}$).

Так как X_N является решением укороченного уравнения

$$X' = \sum_{j=1}^N \Pi_{ij}(t) \cdot Z_j = f_N(X, t) \quad (6.22)$$

и функция $f_N(x, t)$ удовлетворяет условию 6.1, то X_N принадлежит пространству решений уравнения (6.3), которое согласно лемме 6.2 является полным метрическим пространством. Следовательно, последовательность X_N сходится. Пусть

$$\lim_{N \rightarrow \infty} X_N(t) = \bar{X}(t) \quad (6.23)$$

Покажем, что $\bar{X}(t)$ является решением уравнения (6.3). Действительно, записывая (6.22) в интегральной форме

$$X_N(t) = X_0 + \int_{t_0}^t f_N(X_N(\tau), \tau) d\tau \quad (6.24)$$

и переходя к пределу при $N \rightarrow \infty$, в силу непрерывности f_N и (6.23), получим:

$$\bar{X}(t) = X_0 + \int_{t_0}^t f(\bar{X}(\tau), \tau) d\tau \quad (6.25)$$

где

$$f(\bar{X}(t), t) = \lim_{N \rightarrow \infty} f_N(\bar{X}(t), t)$$

Так как (6.23) эквивалентно дифференциальному уравнению (6.3) для $\bar{X}(t)$, то тем самым мы доказали утверждение теоремы.

Следствие 6.1 Имеет место следующая оценка скорости сходимости последовательности X_N к решению \bar{X} :

$$\|X_N - \bar{X}\| \leq B \cdot \left\{ e^p - \sum_{k=0}^{N-2} \frac{p^k}{k!} \right\} p^2 \quad (6.26)$$

здесь

$$B = 2M \cdot L \cdot A^{-2} \cdot J(L) \quad \text{и} \quad p = LZA$$

Оценка (6.26) вытекает из (6.21) при переходе к пределу $N_2 \rightarrow \infty$.

§ 7. Оптимизация микрозонда

Γ^0 В качестве примера практической реализации вышеизложенной теории рассмотрим задачу оптимизации нелинейного протонного микрозонда. Сформулируем кратко физическую задачу и ее математическую модель. Итак, дана система из четырех квадрупольных линз одинаковой длины, начальное множество, занимаемое пучком, определяется двумя последовательными диафрагмами. Требуется найти такие параметры системы (градиенты линз, их длину, местоположение линз и диафрагм, размеры диафрагм и т.д.); которые обеспечивали бы минимальные

размеры изображения пучка на мишени. С учетом третьего порядка дифференциальные уравнения движения пучка в системе квадрупольных линз принимает вид:

$$X'' + KX = \frac{K''}{12} X(X^2 + 3y^2) - \frac{K}{2} X(3X'^2 + y'^2) + K'xyy' + Kx'y'y' + O(5) \quad (7.1)$$

$$y'' - Ky = -\frac{K''}{12} y(y^2 + 3x^2) + \frac{K}{2} y(3y'^2 + x'^2) - K'yxx' - Ky'xx' + O(5) \quad (7.2)$$

для магнитных линз и

$$X'' + KX = \frac{K''}{6} X^3 + K^2 X(y^2 - x^2) + \frac{K'}{2} (x^2 - y^2) X - K(x'^2 + y'^2) X + O(5) \quad (7.3)$$

$$y'' - Ky = -\frac{K''}{6} y^3 + K^2 y(x^2 - y^2) - \frac{K'}{2} (x^2 - y^2) y - K(x'^2 + y'^2) y + O(5) \quad (7.4)$$

для электростатических линз. Здесь $K = K(S)$ — безразмерный градиент линзы, $K > 0$ для фокусировки линзы, $K < 0$ для дефокусирующей и $K = 0$ для свободного промежутка. В качестве независимого параметра выбрана безразмерная длина, измеряемая вдоль оптической оси системы — S .

I^0 Следуя изложенному в § 6 методу погружения в пространство фазовых моментов, введем вектор Y_3 , состоящий из трех векторов:

$$Z_1 = \begin{Bmatrix} X \\ X' \end{Bmatrix}, \quad Z_3^1 = \begin{Bmatrix} X^3 \\ X^2 X' \\ X X'^2 \\ X'^3 \end{Bmatrix}, \quad Z_3^2 = \begin{Bmatrix} Xy^2 \\ Xyy' \\ Xy'^2 \\ X'y^2 \\ X'y'y' \\ X'y'^2 \end{Bmatrix}, \quad Y_3 = \begin{Bmatrix} Z_1 \\ Z_3^1 \\ Z_3^2 \end{Bmatrix} \quad (5.5)$$

Тогда линейное уравнение, эквивалентное с точностью вплоть до третьего порядка уравнениям (5.1), (5.3), примет вид

$$Y_3' = P_3(S) Y_3 = \begin{Bmatrix} P_3'' & P_3^{12} & P_3^{13} \\ 0 & P_3^{22} & 0 \\ 0 & 0 & P_3^{33} \end{Bmatrix} \quad (5.6)$$

Для магнитных линз (уравнение (5.1))

$$P_3'' = \begin{Bmatrix} 0 & 1 \\ -K & 0 \end{Bmatrix}, \quad P_3^{12} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ \frac{K''}{12} & 0 & -\frac{3}{2}K & 0 \end{Bmatrix}, \quad P_3^{13} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{K''}{4} & K' & -\frac{K}{2} & 0 & K & 0 \end{Bmatrix}$$

$$P_3^{22} = \begin{Bmatrix} 0 & 2 & 0 & 0 \\ -K & 0 & 2 & 0 \\ 0 & -2K & 0 & 1 \\ 0 & 0 & -3K & 0 \end{Bmatrix}, \quad P_3^{33} = \begin{Bmatrix} 0 & 2 & 0 & 1 & 0 & 0 \\ K & 0 & 1 & 0 & 1 & 1 \\ 0 & 2K & 0 & 0 & 0 & 0 \\ -K & 0 & 0 & 0 & 2 & 0 \\ c & -K & 0 & K & 0 & 1 \\ 0 & 0 & -K & 0 & 2K & 0 \end{Bmatrix} \quad (5.7)$$

В случае электростатических линз (уравнение (5.3)) изменяются только P_3^{12} и P_3^{13} .

$$P_3^{12} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ -K^2 + \frac{K''}{6} & \frac{K'}{2} & -K & 0 \end{Bmatrix}, \quad P_3^{13} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ K^2 & 0 & -K & -\frac{K'}{2} & 0 & 0 \end{Bmatrix} \quad (5.8)$$

Чтобы получить уравнения, эквивалентные (5.2) и (5.4), необходимо при определении вектора Y_3 заменить X на Y , Y на X , а в выражениях для блок-матриц P_3^{ij} заменить K на $-K$.

2° Введем матрицанты уравнений вида (5.6) в двух плоскостях X и Y :

$$Y_3^X(s) = H_3(s/s_0) Y_3^X(s_0), \quad Y_3^Y(s) = V_3(s/s_0) Y_3^Y(s_0) \quad (5.9)$$

Блок-матрица H_3 , V_3 для прямоугольной модели линзы (то есть $K = \text{const}$ в линзе и $K = 0$ вне ее) имеют вид:

$$H_3''(s/s_0) = \begin{Bmatrix} \cos \varphi & \frac{\sin \varphi}{\sqrt{K}} \\ -\sqrt{K} \cdot \sin \varphi & \cos \varphi \end{Bmatrix}, \quad V_3''(s/s_0) = \begin{Bmatrix} \text{ch} \varphi & \frac{\text{sh} \varphi}{\sqrt{K}} \\ \sqrt{K} \text{sh} \varphi & \text{ch} \varphi \end{Bmatrix} \quad (5.10)$$

где $\varphi = \sqrt{K}(s - s_0)$. Блок-матрицы H_3^{22} , H_3^{33} задаются выражениями (для обоих типов линз):

$$H_3^{22} = \begin{Bmatrix} h_{11}^3 & 3h_{11}^2 h_{22} & 3h_{11} h_{22}^2 & h_{22}^3 \\ h_{11}^2 h_{21} & h_{11}(h_{11} h_{22} + 2h_{12} h_{21}) & h_{12}(2h_{11} h_{22} + h_{12} h_{21}) & h_{12}^2 h_{22} \\ h_{11} h_{21}^2 & h_{21}(2h_{11} h_{22} + h_{12} h_{21}) & h_{22}(2h_{12} h_{21} + h_{11} h_{22}) & h_{12} h_{22}^2 \\ h_{21}^3 & 3h_{21}^2 h_{22} & 3h_{21} h_{22}^2 & h_{22}^3 \end{Bmatrix} \quad (5.11)$$

$$\mathbb{H}_3^{33} = \left\{ \begin{array}{lll}
 h_{11} v_{11}^2 & 2 h_{11} v_{11} v_{12} & h_{11} v_{12}^2 \\
 h_{11} v_{11} v_{21} & h_{11} (v_{11} v_{22} + v_{12} v_{21}) & h_{11} v_{12} v_{22} \\
 h_{11} v_{21}^2 & 2 h_{11} v_{21} v_{22} & h_{11} v_{22}^2 \\
 h_{21} v_{11}^2 & 2 h_{21} v_{11} v_{12} & h_{21} v_{12}^2 \\
 h_{21} v_{11} v_{21} & h_{21} (v_{11} v_{22} + v_{12} v_{21}) & h_{21} v_{12} v_{22} \\
 h_{21} v_{21}^2 & 2 h_{21} v_{21} v_{22} & h_{21} v_{22}^2 \\
 \\
 h_{12} v_{11}^2 & 2 h_{12} v_{11} v_{12} & h_{12} v_{12}^2 \\
 h_{12} v_{11} v_{21} & h_{12} (v_{11} v_{22} + v_{12} v_{21}) & h_{12} v_{12} v_{22} \\
 h_{12} v_{21}^2 & 2 h_{12} v_{21} v_{22} & h_{12} v_{22}^2 \\
 h_{22} v_{11}^2 & 2 h_{22} v_{11} v_{12} & h_{22} v_{12}^2 \\
 h_{22} v_{11} v_{21} & h_{22} (v_{11} v_{22} + v_{12} v_{21}) & h_{22} v_{12} v_{22} \\
 h_{22} v_{21}^2 & 2 h_{22} v_{21} v_{22} & h_{22} v_{22}^2
 \end{array} \right. \quad (5.12)$$

Матрицы \mathbb{V}_3^{22} и \mathbb{V}_3^{33} получаются из (5.11)–(5.12) взаимной заменой h_{ik} и v_{ik} . Здесь h_{ik} , v_{ik} элементы матриц \mathbb{H}_3'' и \mathbb{V}_3'' соответственно. Для прямоугольной модели выражение для элементов матриц \mathbb{H}_3^{12} , \mathbb{H}_3^{13} в случае магнитных линз приведены в [4] (формулы табл. 5 и 6).

Для электростатических линз элементы этих матриц принимают вид (h_{ij}^{12} , h_{ij}^{13} – элементы матриц \mathbb{H}_3^{12} и \mathbb{H}_3^{13}):

$$\begin{aligned}
 h_{11}^{12} &= -K(12\varphi \cos\varphi + \cos\varphi - \cos 3\varphi) / 12 \\
 h_{12}^{12} &= -\sqrt{K}(4\varphi \cos\varphi - 3 \sin\varphi + \sin 3\varphi) / 8 \\
 h_{13}^{12} &= (\cos\varphi - \cos 3\varphi + 4\varphi \sin\varphi) / 8 \\
 h_{14}^{12} &= (15 \sin\varphi - 12\varphi \cos\varphi - \sin 3\varphi) / 24 \sqrt{K} \\
 h_{21}^{12} &= K\sqrt{K}(\sin 3\varphi - \sin\varphi - 3\varphi \cos\varphi) / 6 \\
 h_{22}^{12} &= K(\cos\varphi - \cos 3\varphi - 2\varphi \sin\varphi) / 4 \\
 h_{23}^{12} &= \sqrt{K}(12\varphi \cos\varphi - 13 \sin\varphi - 3 \sin\varphi) / 12 \\
 h_{24}^{12} &= (6\varphi \sin\varphi + \cos 3\varphi - \cos\varphi) / 12
 \end{aligned} \quad (5.13)$$

$$\begin{aligned}
 h_{11}^{13} &= K \vartheta \sin \vartheta / 2 & h_{21}^{13} &= K \sqrt{K} (3 \sin \vartheta + 2 \vartheta \cos \vartheta + \sin \vartheta \operatorname{ch} 2 \vartheta) / 4 \\
 h_{12}^{13} &= 0 & h_{22}^{13} &= K \sin \vartheta \cdot \operatorname{sh} 2 \vartheta / 2 \\
 h_{13}^{13} &= -\vartheta \cdot \sin \vartheta / 2 & h_{23}^{13} &= \sqrt{K} (-3 \sin \vartheta - 2 \vartheta \cos \vartheta + \sin \vartheta \operatorname{ch} 2 \vartheta) / 4 \\
 h_{14}^{13} &= \sqrt{K} (2 \sin \vartheta - \vartheta \cdot \cos \vartheta) / 2 & h_{24}^{13} &= K (2 \vartheta \sin \vartheta - \cos \vartheta (\operatorname{ch} 2 \vartheta - 1)) / 4 \quad (5.14) \\
 h_{15}^{13} &= 0 & h_{25}^{13} &= -\sqrt{K} \cos \vartheta \operatorname{sh} 2 \vartheta / 2 \\
 h_{16}^{13} &= (\vartheta \cos \vartheta - \sin \vartheta) / 2 \sqrt{K} & h_{26}^{13} &= -(2 \vartheta \sin \vartheta + \cos \vartheta (\operatorname{ch} 2 \vartheta - 1)) / 4
 \end{aligned}$$

Выражения для элементов матриц \mathbb{V}_3^{12} и \mathbb{V}_3^{13} получаются из (5.13), (5.14) заменой \sqrt{K} на $i\sqrt{K}$, $\sin \vartheta$ на $i \operatorname{sh} \vartheta$, $\cos \vartheta$ на $\operatorname{ch} \vartheta$ и наоборот. Для свободных промежутков длиной ℓ получим

$$\mathbb{H}_3'' = \mathbb{V}_3'' = \begin{Bmatrix} 1 & \ell \\ 0 & 1 \end{Bmatrix} \quad \text{и} \quad \mathbb{H}_3^{12} = \mathbb{V}_3^{12} = 0, \quad \mathbb{H}_3^{13} = \mathbb{V}_3^{13} = 0 \quad (5.15)$$

3⁰ Используя групповое свойство матрицанта, запишем

$$\mathbb{R}(S_{\text{tot}} | S_0) = \prod_{j=1}^{2n+1} \mathbb{R}(S_j | S_{j-1}), \quad (5.16)$$

здесь n - число участков постоянства $K(S)$, $S_j = S_0 + (j-1) \Delta S_j$ (ΔS_j - длина j -го участка постоянства), $S_{2n+1} = S_{\text{tot}}$. Здесь под \mathbb{R} понимаются матрицанты \mathbb{H}_3 и \mathbb{V}_3 . В целях экономии памяти ЭВМ и увеличения скорости вычислений при организации вычисления произведения (5.16) можно использовать блочную структуру матрицантов, так как можно заменить умножение полных матрицантов на вычисление блок-матриц, их составляющих, согласно следующим рекуррентным формулам:

$$\begin{aligned}
 \mathbb{R}''(S_j | S_0) &= \mathbb{R}''(S_j | S_{j-1}) \mathbb{R}''(S_{j-1} | S_0) & (5.17) \\
 \mathbb{R}^{12}(S_j | S_0) &= \mathbb{R}''(S_j | S_{j-1}) \mathbb{R}^{12}(S_{j-1} | S_0) + \mathbb{R}^{12}(S_j | S_{j-1}) \mathbb{R}^{22}(S_{j-1} | S_0) \\
 \mathbb{R}^{13}(S_j | S_0) &= \mathbb{R}''(S_j | S_{j-1}) \mathbb{R}^{13}(S_{j-1} | S_0) + \mathbb{R}^{13}(S_j | S_{j-1}) \mathbb{R}^{33}(S_{j-1} | S_0)
 \end{aligned}$$

Вычисляя матрицанты полной системы и задавая начальное множество (например, в виде матрицы начальных данных), мы можем получить его образ на мишени. Так как необходимо максимальным образом сжать пучок в реальном пространстве, то в качестве минимизируемого функционала могут выступать следующие выражения:

$$\vartheta_0 = \max \left\{ \sup_{x \in \mathcal{M}_{\text{tot}}} x, \sup_{y \in \mathcal{M}_{\text{tot}}} y \right\} \quad - \text{МИНИМИЗАЦИЯ МАКСИМАЛЬНОГО ОТКЛОНЕНИЯ}$$

$$g_0 = \sup_{x, y \in M_{\text{tot}}} \sqrt{x^2 + y^2}$$

- минимизация радиуса окружности, в которую вписано конечное множество

Ограничения задачи имеют вид:

$$\underline{w}_i \leq w_i \leq \overline{w}_i \quad (5.18)$$

$$\max\{|x(s)|, |y(s)|\} < a(s) \quad (5.19)$$

$$m_{11}(2|1) = m_{22}(2|1) \quad (5.20)$$

Ограничения (5.18) представляют собой ограничения на пределы изменения компонент вектора параметров w , (5.19) - апертурные ограничения ($a(s)$ - переменная апертура системы). Ограничения в виде равенств (5.20) связаны с требованием равномерного сжатия пучка в обеих плоскостях [5], [6].

Здесь $m_{ij}(2|1)$ - элементы матрицанта системы 4-х квадрупольных линз без учета начального и конечного свободных промежутков (квадруплет вращения). В некоторых случаях эти ограничения опускаются.

4° По приведенным выше формулам составлен пакет подпрограмм **ZOND**, с помощью которого вычисляются значения функционала g_0 и ограничений для некоторого вектора варьируемых параметров w . С помощью оптимизационного пакета подпрограмм **MIMIM**, составленного на основе метода скользящего допущения, проводилась оптимизация нелинейного (то есть с учетом нелинейных эффектов 3-го порядка) микрозонда. Следует отметить, что как оценки, полученные численным методом (сравнение с решением полного нелинейного уравнения), так и оценки, полученные с помощью формулы (6.26), показывают хорошие соответствие решения, полученного с помощью метода погружения и решения полного нелинейного уравнения (численно сравнение проводилось для нескольких траекторий).

ЛИТЕРАТУРА

1. С.Г.Михлин " Численная реализация вариационных методов", М., "Наука", 1966.
2. Д.М.Химмельблау " Прикладное нелинейное программирование", М., "Мир"; 1975.
3. А.Д.Дымников, сб. " Программирование и методы решения физических задач", Дубна, 1978, стр. 300-304.
4. G.E. Lee - Whiting, *Nucl. Instz. and Meth.*, 83, 232, 1970.
5. А.Д.Дымников, С.Я.Явор, *ЖТФ*, 33, 851, 1963.
6. А.Д.Дымников, Т.Я.Фишкова, С.Я.Явор, *ЖТФ*, 35, 431, 1965.

Handwritten signature:
Филиппов
Филиппов