**ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ**

**7.1 Общие понятия**

Один из методов решения экстремальных задач, связанных с оптимизацией управления производственными процессами, – динамическое программирование. В отличие от других видов программирования оно не имеет однозначной математической формулировки, а выражает своеобразный подход к решению, методику его построения. Это не значит, что динамическое программирование вообще не имеет математических формулировок, но они, как правило, следствие исследования процесса. Использующиеся в динамическом программировании зависимости могут быть заданы любым способом вплоть до графиков и таблиц. Характерные особенности задач динамического программирования:

* многовариантность решения;
* этапность решения;
* аддитивность критерия (общий критерий равен сумме критериев на этапах).

С помощью динамического программирования решаются задачи, связанные с процессами, которые можно разделить на некоторое число этапов (шагов). Оптимизация управления на каждом этапе в отдельности не обеспечивает оптимизации процесса в целом. Если число этапов и возможных решений на каждом этапе (управлений) ограничено, то оптимальное решение в целом (оптимальную стратегию) можно найти перебором всех возможных вариантов. Однако во многих случаях такой путь неприемлем вследствие очень большого числа вариантов. Динамическое программирование позволяет, не нарушая строгости решения, сократить число рассматриваемых вариантов. Идея заключается в том, что отыскание экстремального значения функции многих переменных заменяется многократным отысканием экстремальных значений функции одного или небольшого числа переменных. Для этого вычислительный процесс делится на этапы. Выбирают такое решение задачи, которое позволяет оптимизировать данный этап. Однако оно должно учитывать не только условия этого этапа, но и весь последующий ход процесса, для чего необходимо знать все решения задачи на последующих этапах.

Поскольку процесс заканчивается на последнем этапе, оптимальное решение не должно учитывать последующего хода. Если найти его для всех ситуаций, которые могут сложиться к началу последнего этапа, то можно найти оптимальное решение и для предпоследнего этапа, т. е. будет получена оптимальная стратегия для двух последних шагов. Если она отразит ситуации, которые могут сложиться к началу предпоследнего этапа, то можно приступить к поиску оптимального решения также на предшествующем ему этапе. Таким образом, процесс вычисления протекает в обратном направлении, от конца к началу. На каждом этапе находят некоторое множество условно-оптимальных решений, выбор из которых зависит от решения уже на предыдущем этапе. Непрерывная последовательность таких решений на всех этапах выражает одно из возможных решений задачи в целом (допустимую стратегию). Сопоставляя допустимые стратегии, выбирают действительно оптимальную в зависимости от начального состояния системы (процесса) и, повторяя процесс вычисления уже от начала к концу, выделяют на каждом этапе оптимальные решения, как составляющие оптимальной стратегии.

Принцип оптимальности впервые сформулирован и доказан Беллманом: оптимальная стратегия, начиная с любого этапа, зависит не от предыдущей стратегии, а лишь от состояния системы на данном этапе и последующей стратегии, т. е. от решений на последующих этапах.

Возможна и геометрическая интерпретация задачи динамического программирования (рис. 7.1).

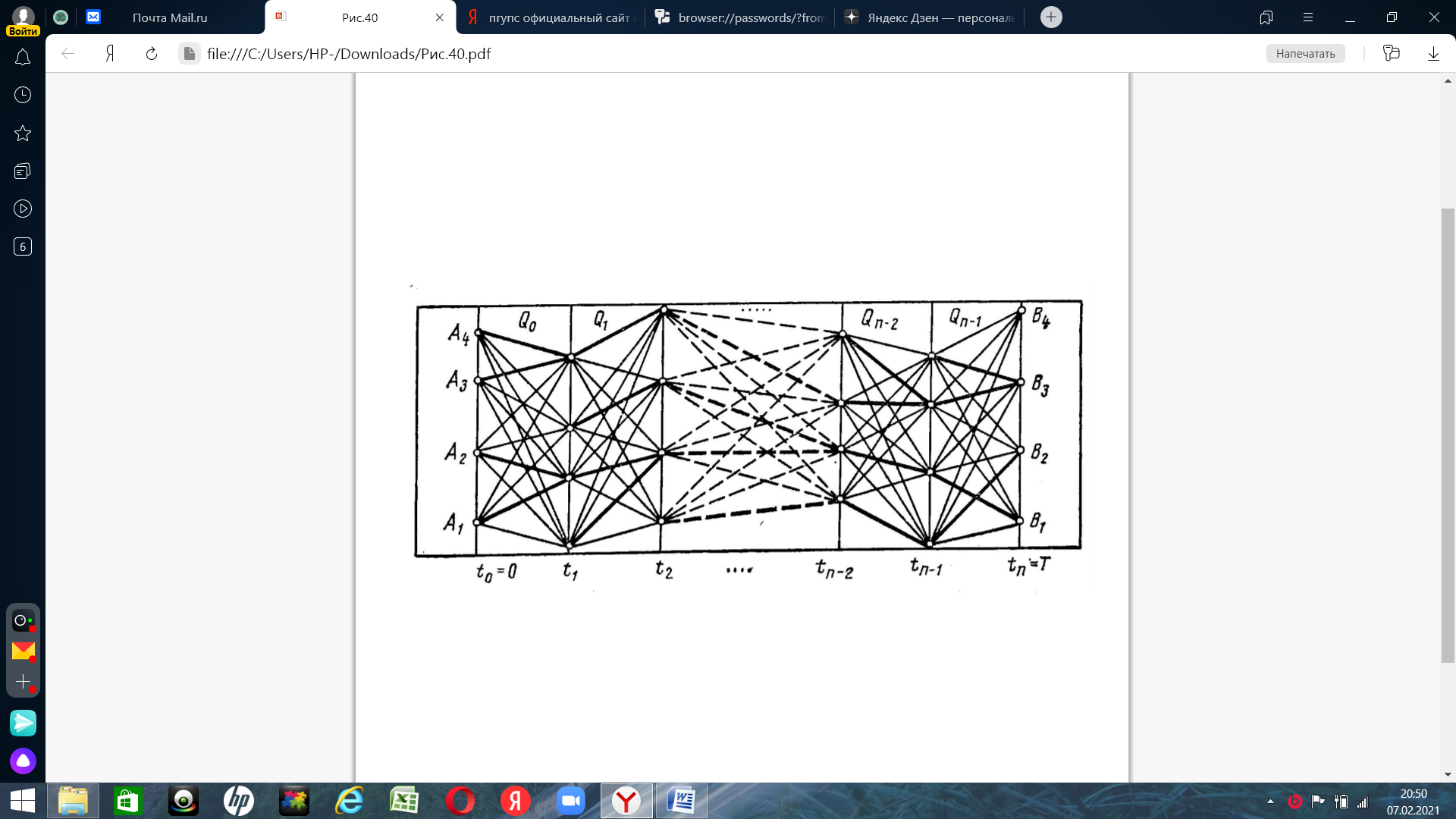


Рисунок 7.1 Геометрическая интерпретация задачи

Вертикальным линиям соответствуют моменты времени, в которые рассматривают исследуемую задачу. В начальный момент *t*0=0 процесс (система) находится в одном из возможных начальных состояний, множеству которых соответствует множество точек *Ai*. Начальное состояние может быть задано либо областью возможных состояний, либо одним конкретным значением, в данном случае четырьмя *A*1, *A*2, *A*3 и *A*4. Будем также считать для простоты, что в каждый момент времени система находится также в одном из четырех возможных состояний, которые показаны точками на соответствующих вертикалях. Конечное состояние системы – ода из четырех точек *В*1, *В*2, *В*3 и *В*4.

Система переводится из начального состояния в следующее с помощью функции перехода, которую еще называют управлением системы на данном этапе. Для каждого из возможных состояний существует своя функция перехода (или некоторое множество их), которая переводит систему в некоторое множество состояний в следующий момент времени. Эта функция – количественная характеристика перехода в следующее состояние в зависимости от предыдущего – выражает либо выигрыш, либо затраты. Поскольку значение функции перехода зависит от предыдущего *x*(*i*) и от последующего *x*(*i*+1) состояний системы, то ее можно записать в общем случае так: .

Каждая допустимая стратегия выражается ломаной линией, соединяющей вертикаль *t*0=0 с вертикалью *t*n=T. Состоит она из набора уравнений на каждом этапе, т. е. ей можно сопоставить число . Оптимальной стратегии соответствует ломаная с наименьшим значением *F*. Следовательно, исходную задачу можно сформулировать в следующем виде: требуется из всех допустимых ломаных, соединяющих вертикаль *t*0=0 с вертикалью *t*n=T, выбрать такую, которой соответствует наименьшее значение *F*.

Решают задачу в следующем порядке. Для всех возможных состояний системы в начале последнего этапа *x*(*n-*1) определяют оптимальное управление – выбирают функцию перехода в одно из конечных состояний с минимальным значением. Переходы, соответствующие минимальному значению  для каждого состояния (*n-*1), показаны на рис. 7.1 жирной линией. Таким образом, в какой бы точке не оказалась система в начале последнего этапа, всегда можно предложить оптимальную стратегию для перевода ее в конечное состояние, получить ряд условно-оптимальных решений. Условие оптимальности каждого такого решения – состояние системы в начале рассматриваемого периода.

Теперь для каждого состояния системы в начале предпоследнего этапа *x*(*n-*2) можно определить условно-оптимальные стратегии для перевода в одно из конечных состояний уже по общему минимуму функций перехода на двух последних этапах: min(+). При этом значения  уже известны в результате предыдущих вычислений. Затем аналогично определяют условно-оптимальные стратегии на трех последних этапах по условию min(++), причем сумма + уже известна. Расчеты продолжают до тех пор, пока не будет пройден весь процесс в обратном направлении.

Каждая из полученных ломаных (жирная линия) соответствует условно-оптимальной стратегии для всего процесса. Поскольку множеству начальных состояний системы соответствует множество точек на вертикали *t*0, каждой условно-оптимальной стратегии соответствует свое начальное состояние системы (точка, из которой она выходит). Таким образом, условно-оптимальная стратегия будет оптимальной при условии, что начальное состояние системы находится в соответствующей точке. Каждая условно-оптимальная стратегия оценивается значением . По нему можно выбрать начальное состояние системы и в зависимости от него окончательно определить оптимальную стратегию, т. е. пройдя процесс уже от начала к концу, установить на каждом этапе оптимальные решения.

Принцип оптимальности Беллмана в этой интерпретации задачи динамического программирования означает следующее: оптимальный путь из любой точки, отражающей состояние системы в какой-либо момент времени, не зависит от траектории, ведущей в эту точку. Поэтому для определения оптимального решения в целом необходимо всегда находить оптимальное продолжение процесса относительно состояния, достигнутого в результате решения на предшествующем этапе.

**7.2 Математическая постановка задачи**

Состояние некоторой развивающей системы в каждый дискретный момент времени *t*0=0, ..., *T* характеризуется множеством точек , , ..., , совокупность которых – вектор , , ..., , называемый вектором состояния системы. Обозначим множество всех состояний системы (вектор состояний) в момент времени *t*=*m* через *xm*; состояние ее в начальный момент *t*=0 считается заданным *х*(0)=*х*0. Развитие системы состоит в последовательном переходе из одного состояния в другое. На него можно воздействовать в каждый момент времени *t* определенным управлением *u*(*t*), выбранным из множества возможных управлений. Таким образом, состояние системы *х*(*t*+1) определяется с одной стороны, вектором *х*(*t*) и, с другой стороны, управлением *u*(*t*)

*х*(*t*+1)=*f*{*х*(*t*); *u*(*t*)}.

Функция *f* задает правило перехода от состояния *х*(*t*) в состояние *х*(*t*+1) в зависимости от управления *u*(*t*). Множество управлений, каждое из которых можно выбрать в момент *t*=*m*, обозначим через *Um*. Развитие системы определяется последовательностью *х*={*х*(0), *х*(1), ..., *х*(*T*)}, где  – вектор состояния системы при *t*=*m*. Эта последовательность называется стратегией. Стратегия допустимая, если существуют управления, позволяющие делать переход из любого ее состояния в следующее. Каждая стратегия оценивается функцией цели *F*(*x*). Таким образом, развитие системы описывается:

* множеством допустимых состояний системы *Xm* ;
* множеством допустимых управлений *Um*;
* правилами перехода из одного состояния в другое по выбранному управлению *х*(*m*+1)=*f*{*х*(*m*); *u*(*m*)};
* функцией цели *F*(*x*).

Необходимо найти допустимую стратегию, обеспечивающую минимум (максимум) функции цели. Последнюю в общем случае задают суммой оценочных функций *Qm*{*x*(*m*), *х*(*m*+1)}, получаемых при каждом переходе из состояний *х*(*m*) в состояние *х*(*m*+1),



Функцию цели можно считать функцией управления, поскольку любую допустимую стратегию *х* полностью определяет последовательность допустимых управлений *u=*(*u*0, *u*1, ...,  *u*Т-1). Таким образом, задачу динамического программирования можно сформулировать так: необходимо определить последовательность управлений



минимизирующую функцию цели

 (7.1)

при условиях:

 (7.2)

Обозначим через *Fk*{*x*(*k*)} максимальное значение функционала  для {*x*(*m*)}; {*u*(*m*)}; *m*=*k*+1, ..., *T*, удовлетворяющих условиям задачи. Тогда, применяя последовательно принцип оптимальности к функциям *F*1, *F*2 и т. д., можно написать систему уравнений:

 (7.3)

*Fi*{*x*(*i*)} называются функциями Беллмана. Решение задачи динамического программирования сводится к решению данной системы функциональных уравнений, как уже было сказано, начиная с последнего уравнения. Задачи динамического программирования, как правило, требуют большого объема вычислений и обычно их решают, используя вычислительную технику. Далее рассмотрим некоторые конкретные задачи, максимально упростив их для иллюстрации механизма вычислений.

**Пример.** Необходимо определить такой режим ведения поезда на любом отрезке пути, который обеспечил бы минимальные приведенные расходы на передвижение по участку в целом. Расходы на каждом отрезке пути зависят от скорости поезда и от сопротивления движению, а последнее – от профиля пути. Поэтому в качестве вычислительного этапа можно рассматривать элемент профиля пути с постоянным уклоном.

На участке с восемью элементами профиля (рис. 7.2) движение поезда начинается с первого. В зависимости от режима движения по нему кривая скорости может попасть в некоторое число точек на вертикали 1 (на ней фиксируются все возможные значения скорости через какой-либо интервал, например через 1 км/ч). Рассмотрим только три. Из каждой из них можно попасть в каждую их трех точек по вертикали 2, соблюдая определенный режим движения и реализуя определенные затраты (показаны на линиях, соединяющих точки). Последние подсчитывают заблаговременно для каждого элемента профиля во всех вариантах движения по формуле приведенных затрат в тяговых расчетах.

Решаем задачу с конца поэтапно. На последний элемент поезд может вступить с тремя значениями скорости (вертикаль 7). Каждому из них однозначно соответствуют приведенные расходы на передвижение по данному элементу до остановки. Выделяем три условно-оптимальных стратегии (показано жирными линиями): ; ; .

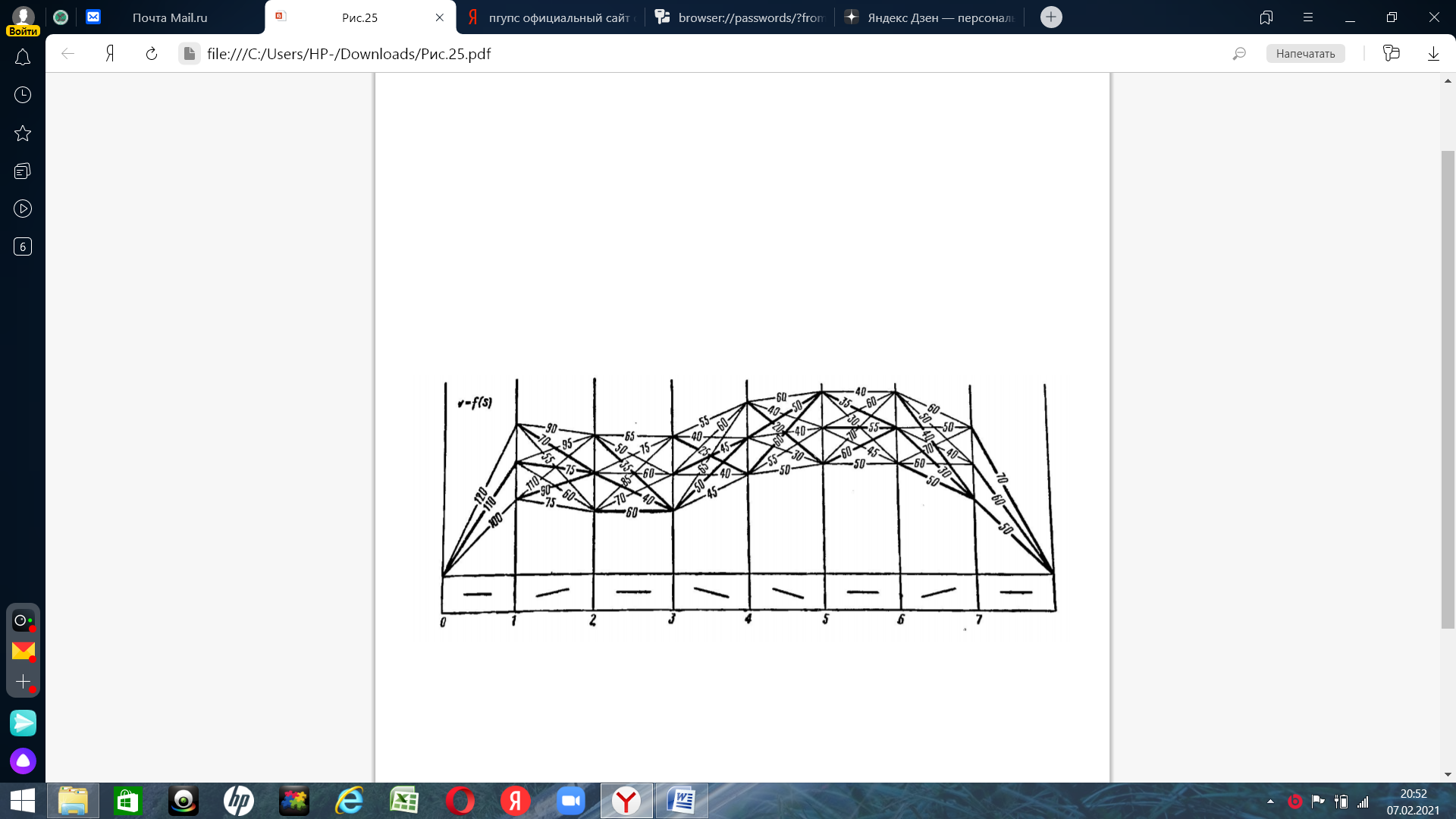


Рисунок 7.2 Определение наивыгоднейшего режима ведения поезда по участку

На предпоследний элемент поезд может вступить также с тремя значениями скорости (вертикаль 6). Однако из каждой точки можно попасть в конечную точку на вертикали 8 по трем вариантам (трем режимам). Поэтому находим условно-оптимальные стратегии на этом элементе (сравнивая между собой все варианты движения поезда) по минимуму общих расходов на обоих элементах пути





Аналогично определяем условно-оптимальные стратегии на следующем элементе по общему минимуму затрат на трех последних



 и т.д.

На предпоследнем этапе расчетов





Поскольку начальное состояние системы задано однозначно на первом этапе, находим лишь одно значение





Теперь, начиная с первой точки и проходя последовательного непрерывной ломаной жирной линии весь процесс до последней точки, определим оптимальную стратегию для всего процесса и на каждом элементе:



**Пример.** На схеме дорог, соединяющих 16 пунктов (рис. 7.3), требуется найти кратчайший путь из пункта 1 в пункт 16. Расстояния между пунктами указаны на соответствующих участках дорог. Понятно, что попасть из одного пункта 1 в другой 16 можно различными путями, т. е. задача многовариантная. Разобьем все пункты на насколько групп. В группу I включим только начальный пункт 1, в группу II – те пункты (2, 3 и 4), в которые можно попасть из пункта 1; к группе III отнесем пункты (5, 6 и 7), в которые можно попасть из пунктов группы II и т. д. К последней группе VII относится конечный пункт 16.

Кратчайший путь проходит через каждую группу пунктов, т. е. складывается из этапов I-II-III-IV-V-VI-VII. Алгоритм решения этой задачи состоит из последовательного определения кратчайших (условно-оптимальных) путей на каждом этапе, начиная с последнего. Для удобства расчетов обозначим через *lj* – кратчайший путь из пункта *j* в конечный; *lj* (*p*, *q*) – кратчайший путь из пункта *j* в конечный через пункты *p* и *q.*

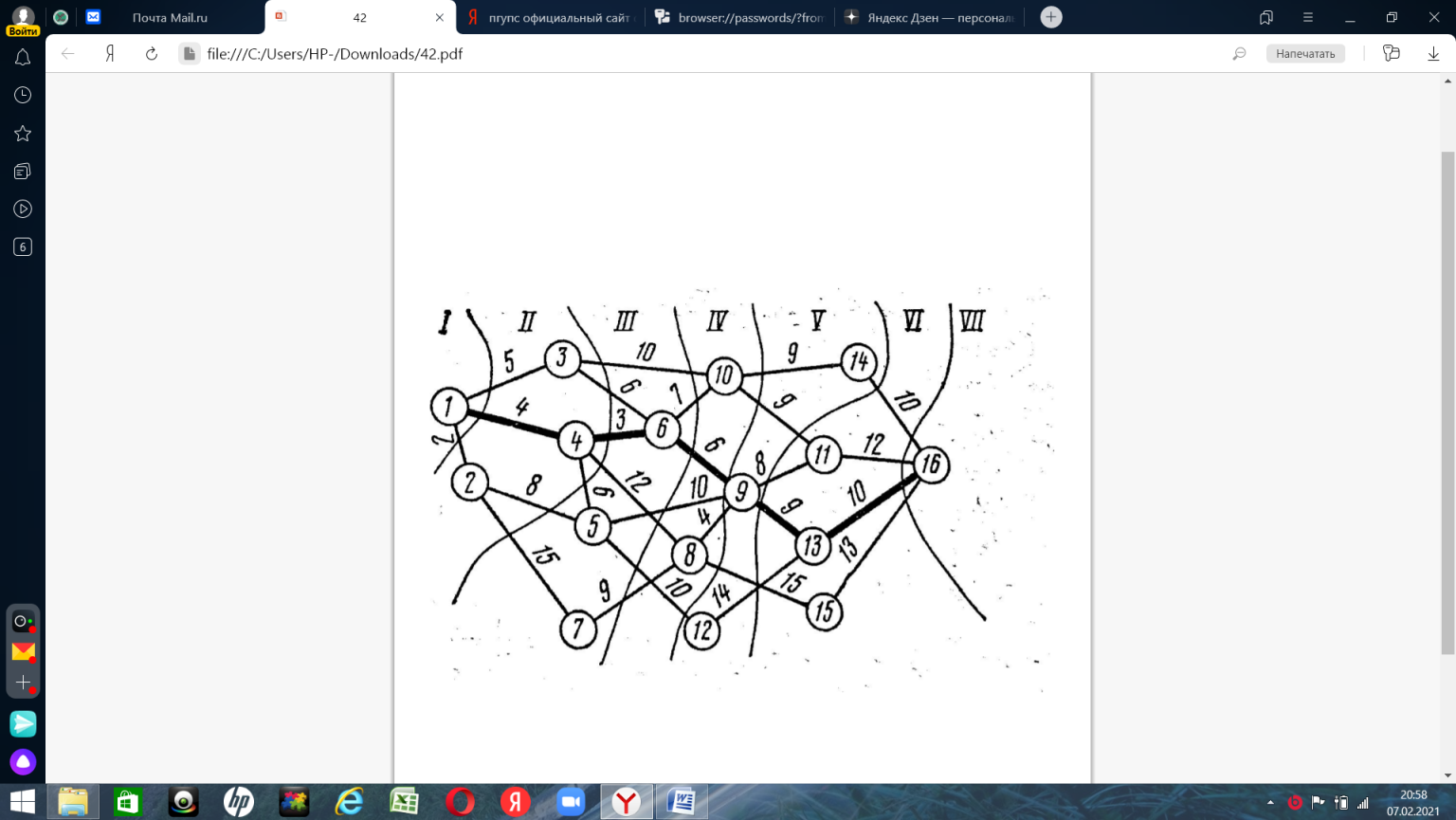


Рисунок 7.3 Кратчайший путь на сети дорог

По схеме находим кратчайшие пути:

1. Из каждого пункта группы VI в конечный (этап VI-VII) *l*15=13; *l*13=10; *l*11=12.

2. Из каждого пункта группы V в конечный с учетом результата первого шага (этап V-VI-VII):

*l*14=10; *l*12=14+*l*13=14+10=24; *l*9=min(8+*l*11; 9+*l*13)=min(8+12; 9+10)=19=*l*9(13).

3. Из каждого пункта группы IV в конечный с учетом результатов предыдущих шагов (этап IV-V-VI-VII):

*l*10=min(9+*l*14; 9+*l*11)=min(9+10; 9+12)=19=*l*10(14);

*l*8=min(4+*l*9; 15+*l*15)=min(4+19; 15+13)=23=*l8*(9, 13).

4. Из всех пунктов группы III до конечного (этап III-IV-V-VI-VII):

*l*7=9+*l*8=9+23=32 *l*7(8, 9, 13);

*l*6=min(7+*l*10; 6+*l*9)=min(7+19; 6+19)=25=*l6*(9, 13);

*l*5=min(10+*l*9; 10+*l*12)=min(10+19; 10+24)=29=*l5*(9, 13).

5. Из всех пунктов группы II (этап II-III-IV-V-VI-VII):

*l*4=min(3+*l6*; 12+*l*8; 6+*l*5)=min(3+25; 12+23; 6+29)=28=*l4*(6, 9, 13);

*l*3=min(10+*l10*; 6+*l*6)=min(10+19; 6+25)=29=*l3*(10, 14);

*l*2=min(8+*l5*; 15+*l*7)=min(8+29; 15+32)=37=*l2*(5, 9, 13).

6. Кратчайший путь из начального пункта в конечный (этап I-II-III-IV-V-VI-VII):

*l*1=min(5+*l3*; 4+*l*4; 2+*l*2)=min(5+29; 4+28; 2+37)=32=*l1*(4, 6, 9, 13).

Таким образом, кратчайший путь из пункта 1 в пункт 16 равен 32 и проходит через пункты 4, 6, 9 и 13 (на рис. 7.3 жирная линия).

**Пример.** Рассмотрим задачу об оптимизации развития пропускной способности линии.

На рис. 7.4 наклонной линией *Г*0+*гt* показан рост грузопотока в одном направлении или потребная пропускная способность в тоннах; *Г*0 – значение грузопотока в начале рассматриваемого периода; *г* – прирост грузопотока за год; *t* – число лет с начала периода; *Т* – период, в течение которого проводятся мероприятия по усилению пропускной способности. Будем считать, что на линии обращаются только грузовые поезда (пассажирским движением пренебрегаем). Уровень пропускной способности выражается горизонтальной линией. Начальный уровень наличной пропускной способности в тоннах *Г*н с ростом грузопотока будет исчерпан в год *t*1 эксплуатации линии.

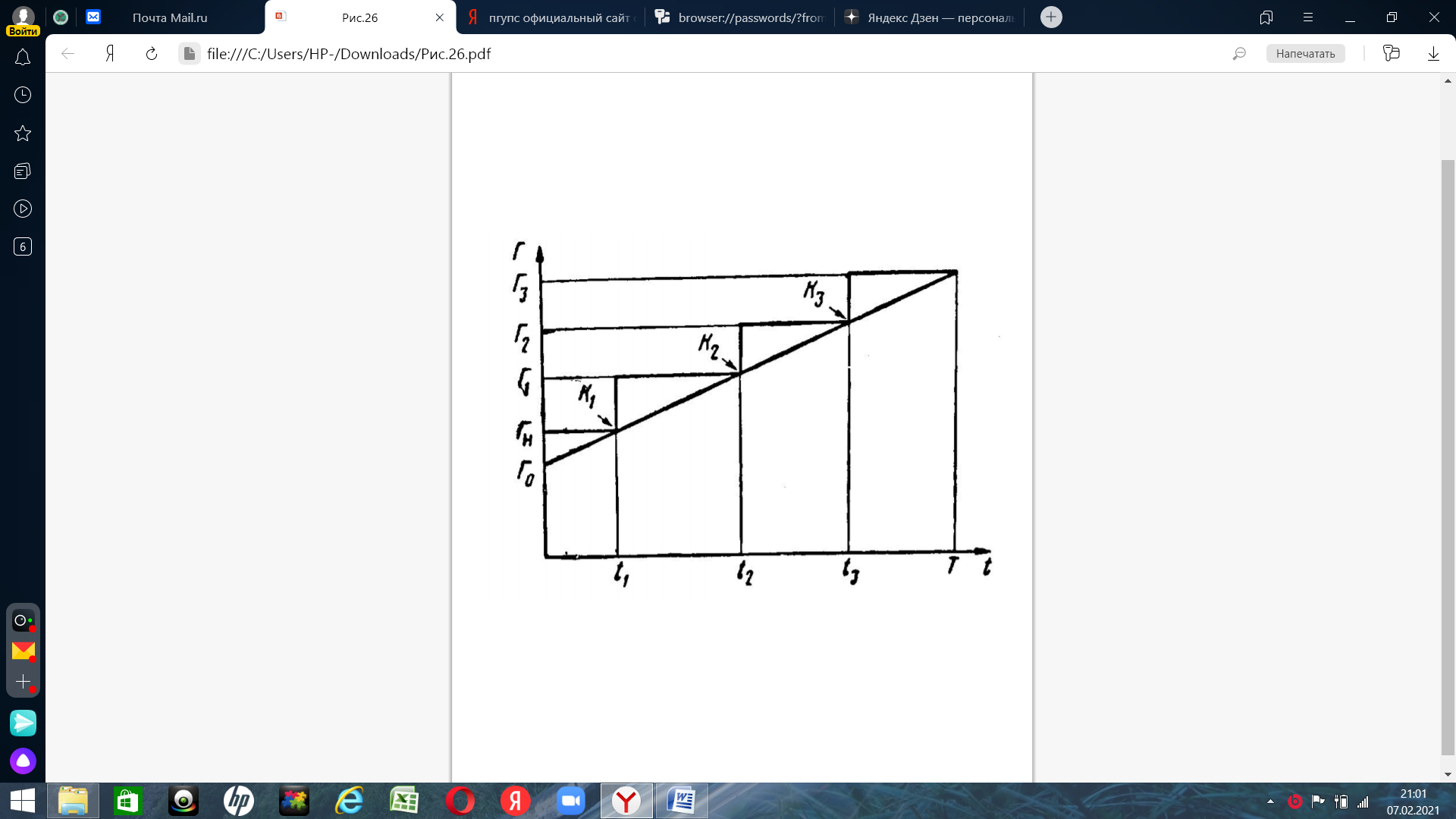


Рисунок 7.4 График роста грузопотока на линии

Усиление пропускной способности можно осуществить различными способами: устройством автоблокировки и диспетчерской централизации, удлинением станционных приемо-отправочных путей, укладкой двухпутных вставок для организации безостановочных скрещений поездов, вторых путей, использованием более мощных локомотивов и др. Каждый из этих способов требует определенных капитальных вложений. Затраты *К*1 какого-либо способа усиления пропускной способности позволяют поднять ее уровень до величины *Г*1. Затем необходимо изыскать другой способ, который потребует капитальных вложений *К*2 в срок *t*2 и т. д. Во многих случаях величину капитальных вложений определяет прирост грузопотока, который необходимо освоить. Например, мощность новых локомотивов должна быть тем больше, чем больше прирост грузопотока, который подлежит освоению ими, удлинение приемо-отправочных путей также зависит от величины дополнительного грузопотока, который необходимо освоить. Чем больше мощность локомотивов, или величина удлинения приемо-отправочных путей, тем больше и капитальные вложения на усиление пропускной способности тем или иным способом. С известной степенью условности можно считать, что величина капитальных вложений *К* пропорциональна приросту грузопотока Δ*Г*, а последний – времени эксплуатации линии в результате усиления пропускной способности данным способом *t*э Δ*Г=гt*э.

Таким образом, величину капитальных вложений можно выразить формулой

*К=αгt*э,

где α – некоторый коэффициент пропорциональности.

Предположим, что на рассматриваемой линии можно использовать три способа усиления пропускной способности в определенной последовательности на протяжении *Т* лет. Каждый из них позволяет освоить различные размеры дополнительного грузопотока в зависимости от величины капитальных затрат. Необходимо определить размеры и сроки вложения капитальных затрат, чтобы общие расходы на усиление пропускной способности всеми тремя способами были минимальными. При этом необходимо учесть экономический эффект от отдаления капитальных вложений, приводя их к начальному году периода *Т*. Чтобы не загромождать расчеты, будем учитывать лишь капитальные затраты в каждом способе усиления пропускной способности. Учет эксплуатационных расходов значительно усложняет решение задачи, однако не изменят принципиально методику расчетов. Приведенные капитальные затраты соответственно составят:







где Δ – нормативный коэффициент эффективности капитальных вложений, ;

*t*н – нормативный срок их окупаемости.

Таким образом, пропускная способность увеличивается в три этапа. требуется определить такие сроки капитальных вложений на каждом этапе, при которых общие приведенные затраты будут минимальными



При этом можно учитывать ограничения типа:

; ,

где Г', Г'' – максимальный уровень пропускной способности, которого можно добиться соответственно первым или вторым способом усиления.

Срок *t*1 определяется однозначно наличной пропускной способностью *Г*н на линии в начале рассматриваемого периода  или . Уровень пропускной способности в конце третьего этапа также определяется однозначно величиной потребной пропускной способности . Сроками же капитальных вложений *t*2 и *t*3 можно варьировать, изменяя тем самым и уровни пропускной способности Г1 и Г2 на первом и втором этапах. И так, необходимо определить сроки *t*2 и *t*3, которые минимизируют общие приведенные капитальные затраты на усиление пропускной способности на всех трех этапах. Задача решается методом динамического программирования в два этапа, поскольку необходимо найти значения двух неизвестных *t*2 и *t*3.

Первый этап оптимизации. Необходимо определить значение *t*3 для всех возможных *t*2, которые обращают общие приведенные капитальные затраты на двух последних этапах усиления пропускной способности *Е*2+*Е*3 в минимум. Исследуем функцию *Е*2+*Е*3 на минимум по аргументу *t*3. Для удобства дифференцирования преобразуем ее:



и определим первую производную по *t*3:



Если вторая производная по *t*3 положительна, то функция имеет минимум



так как *Т*˃*t*3. Приравняв первую производную нулю, попытаемся найти значение *t*3 в зависимости от *t*2, приводящее функцию расходов *Е*2+*Е*3 к минимуму. После соответствующих преобразований получим



Это уравнение типа  Решить его можно лишь графически, представив в виде  Решение этого уравнения – общая абсцисса точки пересечения графиков двух функций  и  (рис. 7.5). В нашем случае ; , где *а*=1+Δ;

;



Величина *а* определена однозначно, а *b* и *c* зависят от неизвестной *t*2. Предположим, что продолжительность периода *T*=*t*1+20 лет.

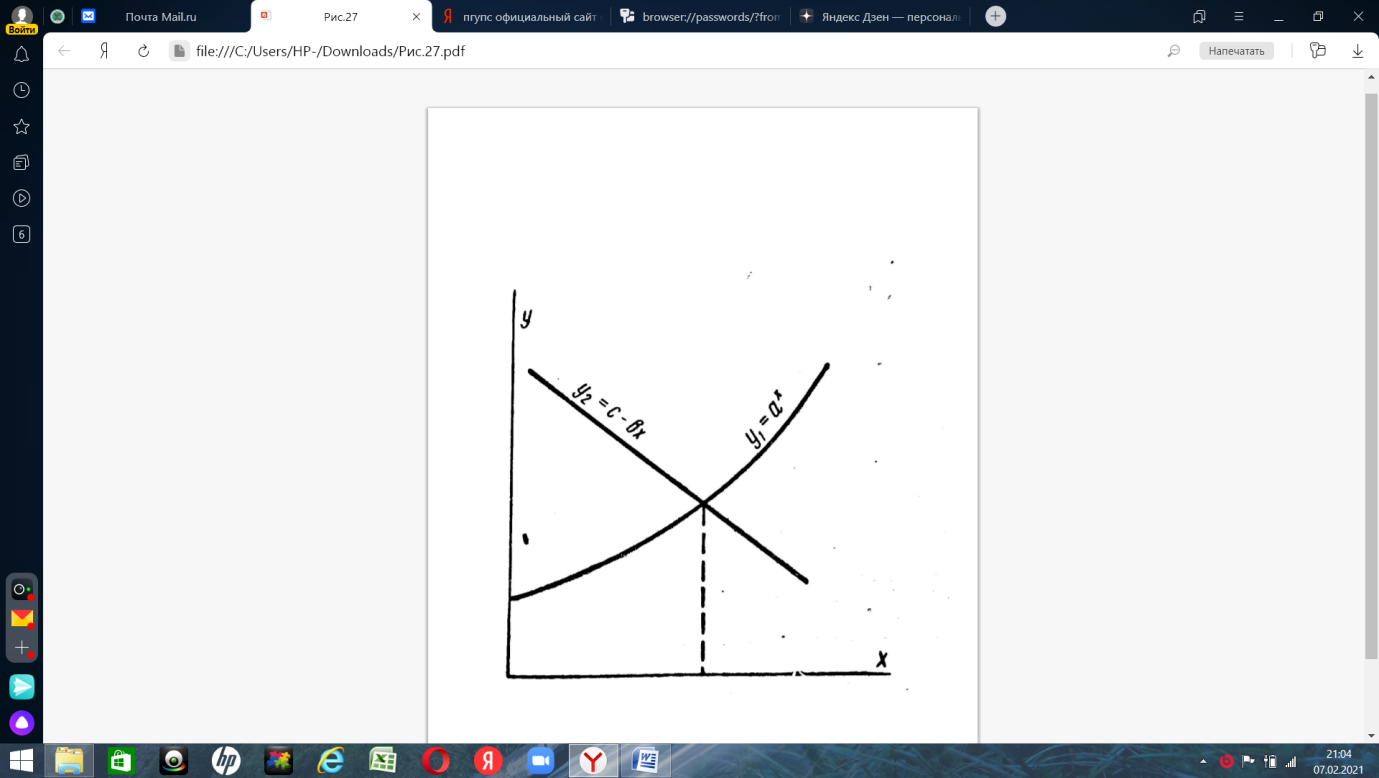
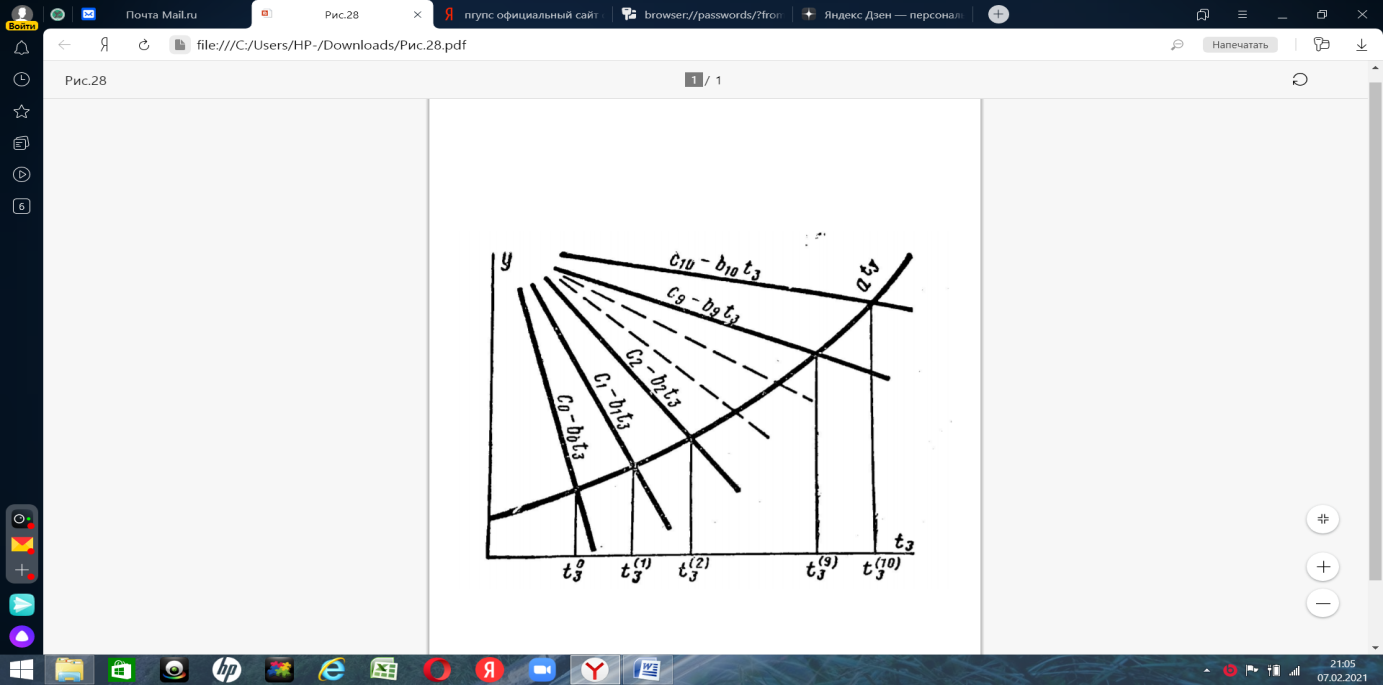
 

Рисунок 7.5 Графическое решение уравнения вида 

Рисунок 7.6 Графическое определение значений *t*3=f(*t*2)

Тогда *t*2 в общем случае может принимать целые значения от *t*2=*t*1 до *t*2=*t*1+20. Однако естественно предположить, что данным способом усиления пропускной способности нельзя освоить прирост грузопотока за весь период. Это может быть задано упомянутыми ранее ограничениями. С учетом их допустим, что величина *t*2 изменяется в пределах от *t*2=*t*1 до *t*2=*t*1+10, т. е. может принимать 11 целых значений. Для каждого из них можно подсчитать величины *b*=*b*0, *b*1, *b*2, ..., *b1*0 и *с*=*с*0, *с*1, *с*2, ..., *с1*0. Построив на графике соответствующие прямые (рис. 7.6) и определив точки пересечения их с кривой , получим соответствующие всем *t*2 значения *t*3:





которые минимизируют суммарные приведенные капитальные затраты на втором и третьем этапах усиления пропускной способности.

Второй этап оптимизации. Определим *t*2 в зависимости от значения *t*1 (известного), обращающее в минимум общие приведенные капитальные затраты на всех трех этапах,



Если бы в результате первого этапа была установлена функциональная зависимость *t*3=*f*(*t*1), то ее можно было бы подставить в данное выражение и исследовать последнее аналогично на минимум по *t*2 с помощью первой производной. Однако такой зависимости не выявлено, а установлены попарно 11 конкретных значений *t*3 и *t*2. Подставляя их в данное выражение, можно подсчитать 11 возможных значений суммарных приведенных капитальных затрат . Минимальному из них соответствуют наиболее рациональные сроки капитальных вложений *t*2 и *t*3, которые и определяют оптимальный вариант развития пропускной способности линии.