# PROBABILISTIC TRAVELING SALESMAN PROBLEMS

by

PATRICK JAILLET

Technical Report No. 185

OPERATIONS RESEARCH CENTER

**MASSACHUSETTS INSTITUTE**
**OF**
**TECHNOLOGY**

February 1985

# PROBABILISTIC TRAVELING SALESMAN PROBLEMS

by

Patrick Jaillet

Submitted in partial fulfillment of the requirements of Doctor of Philosophy

## ABSTRACT

In this thesis, we introduce and analyze a variation of the Traveling Salesman Problem: consider a problem of routing through a set of n points. On any given instance of the problem only a subset consisting of k out of the n points ($0 \leq k \leq n$) have to be visited, with the number k determined according to a known probability distribution (such as the binomial). We wish to find a priori a tour through all n points. On any given instance of the problem the k points present will then be visited in the same order as they appear in the a priori tour. The problem of finding such an a priori tour which is of minimum length in the expected value sense is defined as a "probabilistic traveling salesman problem" (PTSP). There are many possible variations of this generic version of the PTSP, each having potentially important applications.

First we present the derivations of closed-form expressions for computing efficiently the expected length of any given PTSP tour; we then derive some interesting, and occasionally counter-intuitive, properties of optimum PTSP tours; we also obtain relationships between the solutions to TSP and PTSP problems through n points and we develop useful bounds on the length of optimum PTSP tours.

This is followed by an analysis of the PTSP in the plane. We first present bounds on the expected length of the optimal PTSP tour for arbitrary sequences of n points lying in a square as well as for points uniformly and independently distributed over the square. We then analyze the asymptotic behavior of the expected value of the length of the tour obtained through the strategy of reoptimizing the optimal tour for each given instance of the problem. Finally we show that the expected length of the optimal PTSP tour through n points drawn from a uniform distribution in the unit square is almost surely (with probability 1) asymptotic to a constant times $\sqrt{n}$. We also discuss extensions of our results to sequences of points lying in a general subset of a d-dimensioned euclidean space.

Following this theoretical work, and based on it, we propose several different possible strategies to solve PTSPs. After formulating the problem as an integer nonlinear programming problem, we discuss a Branch-and-Bound procedure and also show the inadequacy of dynamic programming approaches for the PTSP. We then present an exposition of heuristic procedures, including: tour construction procedures; "hill-climbing" methods; and partitioning algorithms including a space-filling approach.

We conclude the thesis by suggesting briefly uses of this methodology in different areas of application such as preliminary planning of distribution systems and location of facilities.

Thesis Supervisor: Amedeo R. Odoni

Title: Professor of Aeronautics and Astronautics
and of Civil Engineering

# ACKNOWLEDGEMENTS

I wish to express my warmest feelings of gratitude to my thesis supervisor, Professor Amedeo Odoni for being a constant source of support and inspiration during the course of this dissertation and throughout my stay at MIT.

I would like to thank the other members of the thesis committee, Professors George Kocur, Thomas Magnanti and Harilaos Psaraftis for their interest in this work and for providing valuable comments on previous drafts.

This research was partially funded under a fellowship provided by the Center for Transportation Studies (MIT) through support from the UPS Foundation. Both the Center and the Foundation are acknowledged.

Thanks to Terri Demeris for her outstanding job of typing and organizing the final draft of this thesis.

Finally, and most of all, I want to express my deepest appreciation to my parents and friends for their understanding.

TABLE OF CONTENTS

Page

TITLE .................................................................. 1

ABSTRACT ............................................................... 2

ACKNOWLEDGEMENTS ....................................................... 4

TABLE OF CONTENTS ...................................................... 5

LIST OF FIGURES ........................................................ 8

CHAPTER 1: THE CONSIDERATION OF UNCERTAINTY IN ROUTING PROBLEMS ... 9

1.1 Introduction; Motivation ........................................... 9

1.2 Brief Literature Review ............................................ 14

1.2.1 Nondeterministic Distances Between Points ....................... 14

1.2.2 Nondeterministic Demand at Each Delivery (Pick-up) Location ..... 17

1.3 Outline of the Thesis .............................................. 18

CHAPTER 2: THE EXPECTED LENGTH OF TOURS UNDER VARIOUS CONDITIONS .. 24

2.1 Introduction ....................................................... 24

2.1.1 The Expected Length "In the PTSP Sense"; Outline of Chapter 2 ... 24

2.1.2 Background Information; Notation ................................ 26

2.2 The Expected Length of a Tour t in a Graph G with no Black Node .... 28

2.3 Extensions ......................................................... 35

2.3.1 The Expected Length of a Tour t in a Graph G with One Black Node . 35

2.3.2 The Expected Length of a Tour t in a Graph G with m Black Nodes .. 38

2.3.3 The Probabilistic Hamiltonian Path Problem (PHPP) ............... 40

2.3.4 General Probabilistic Assumptions ............................... 45

2.4 Conclusion ......................................................... 48

CHAPTER 3: GENERAL COMBINATORIAL PROPERTIES OF THE PTSP ............ 50

3.1 Introduction; Preliminaries ........................................ 50

Page

3.1.1 Contents of Chapter 3 .................................. 50
3.1.2 Preliminaries .................................. 52

3.2 Analysis of the Closed Form Expressions of Chapter 2 .......... 56

3.2.1 Properties of the $L_{m,t}^{(r)}$ .................................. 56
3.2.2 Properties of the Weights .................................. 64
3.2.3 Bounds on $E[L_t]$ .................................. 67

3.3 Perturbation of the Graph and Exploitable Properties of the Problem Structure .................................. 69

3.4 Relation Between the TSP and PTSP .................................. 76

3.5 The PHPP .................................. 90

3.6 Conclusion .................................. 91

CHAPTER 4: BOUNDS AND ASYMPTOTIC ANALYSIS FOR THE PTSP IN THE PLANE 92

4.1 Introduction; Notation .................................. 92

4.2 Bounds for Sequence of Points in $[o,r]^2$ .................................. 97

4.2.1 Case of an Arbitrary Sequence x .................................. 97
4.2.2 Case of a Uniform Sequence X .................................. 102

4.3 Asymptotic Analysis .................................. 109

4.3.1 Asymptotic Behavior of the "Strategy of Reoptimizing"... 109
4.3.2 Expected Length of the Optimal PTSP Tour When $n \to \infty$ ...... 113
4.3.3 Bounds on the Constant c(p) .................................. 119
4.3.4 Generalizations .................................. 123

4.4 Conclusion .................................. 130

CHAPTER 5: ALGORITHMIC INVESTIGATION .................................. 131

5.1 Introduction .................................. 131

5.1.1 Focus of Chapter 5 .................................. 131
5.1.2 Contents of Chapter 5 .................................. 132

5.2 Exact Optimization Methods .................................. 134

5.2.1 Integer Nonlinear Programming Formulation of the PTSP .. 137
5.2.2 Linearizations .................................. 145

Page

5.2.3 Discussion of Exact Methods .......... 153

5.3 Heuristic Procedures .......... 162

5.3.1 Theoretical Preliminaries .......... 163
5.3.2 Tour Construction Procedures .......... 175
5.3.3 Tour Improvement Procedures .......... 186
5.3.4 The PTSP in the Plane .......... 188

5.4 Conclusion .......... 194

CHAPTER 6: FINAL REMARKS AND CONCLUSIONS .......... 196

6.1 Summary of Results .......... 196

6.1.1 Methodological (Theoretical) Contributions .......... 196
6.1.2 Discussion .......... 197

6.2 Further Research .......... 202

6.2.1 Algorithmic Implementation .......... 202
6.2.2 Theoretical Investigation .......... 203
6.2.3 Applications of the PTSP Methodology .......... 204
6.2.4 Further Research on the Integration of Uncertainty into Combinatorial Optimization Problems .......... 204

REFERENCES .......... 206

APPENDIX A: Graphical Illustrations of the $L_{m,t}^{(r)}$ .......... 212

APPENDIX B: Case of a Non-Complete Graph G .......... 216

APPENDIX C: The Optimal Euclidean PTSP Tour May Interest Itself: An Example .......... 218

APPENDIX D: Exact Calculations for Lemma 3.9 .......... 224

APPENDIX E: Calculus Associated with the Asymptotic Behavior of The Star-Shaped Construction .......... 228

APPENDIX F: The Three Mathematical Programming Formulations of the PTSP .......... 233

APPENDIX G: A General Branch-and-Bound Scheme for the PTSP .......... 237

APPENDIX H: The Inadequacy of a Dynamic Programming Approach for the PTSP: A Simple Example .......... 239

# LIST OF FIGURES


| | | Page |
|---|---|---|
| 1.1 | Simple Graphical Example of a PTSP | 12 |
| 2.1 | A 24 Nodes Graph | 33 |
| 2.2 | Two Tours of the 24 Nodes Graph | 33 |
| 3.1 | Lemma 3.7 : The Set of Five Points | 72 |
| 3.2 | Lemma 3.7 : The Optimal PTSP Tours | 72 |
| 3.3 | Lemma 3.9 : Illustrations for m=1 n=4 | 79 |
| 3.4 | Lemma 3.9 : Illustrations for m=0 n=6 | 79 |
| 3.5 | Lemma 3.10 : The Graph $G_1$ | 85 |
| 3.6 | Lemma 3.10 : The Graph $G_2$ | 85 |
| 4.1 | Lemma 4.1 : Construction of a Tour | 98 |
| 4.2 | Lemma 4.2 : Construction of a Tour | 103 |
| 4.3 | Theorem 4.2 : The Subadditivity of $\phi_p$ | 116 |
| 5.1 | An Illustration of a D.P. Approach for the PTSP | 160 |
| 5.2 | Lemma 5.7 : Merging of Subtours | 173 |
| 5.3 | Tour Aggregation | 177 |
| 5.4 | "Non-Desirable" Type of Tour | 195 |

# CHAPTER 1

# THE CONSIDERATION OF UNCERTAINTY IN ROUTING PROBLEMS

## 1.1 Introduction; Motivation:

Vehicle routing problems involve finding a set of pick up and/or delivery routes from one or several central depots to various demand points (e.g. customers), in order to minimize some objective function (minimization of routing costs, or of the sum of fixed and variable costs, or of the number of vehicles required, etc.) Vehicles may have capacity and, possibly, maximum-route-time constraints. For example, the problem that arises when there is a single domicile (depot), a single vehicle of unlimited capacity, unit demands, only routing costs, and an objective function which minimizes total distance traveled, is the famous Traveling Salesman Problem (TSP). With several vehicles of common capacity, a single depot, known demands, and the same objective function as the TSP, we have a standard Vehicle Routing Problem (VRP). An excellent review of various "node-covering" routing problems can be found in Larson and Odoni [1981] as well as in Golden and Magnanti [1980]; a useful taxonomy for vehicle routing and scheduling problems is contained in Bodin and Golden [1981], one of many interesting papers in a special issue of Networks devoted to the Proceedings of a National Science Foundation Workshop on "Current and Future Directions in the Routing and Scheduling of Vehicles and Crews"; an extensive review of the state of the art in Routing and Scheduling of vehicles and crews is also given in Bodin et al. [1983] (with a bibliography containing around 700 papers!). In fact the scholarly literature devoted to the TSP is by itself quite impressive and one has simply to consult review papers such as Bellmore and Nemhauser [1968] and Parker and Rardin [1983] to convince oneself that the TSP is perhaps the most intensively investigated of all discrete optimization problems. The effort spent on the TSP problem is partially a reflection of the fact that this problem is an essential component of most other vehicle routing problems and that it also has numerous, and sometimes surprising, other applications (see, for example, Lenstra and Rinnooy Kan [1975]).

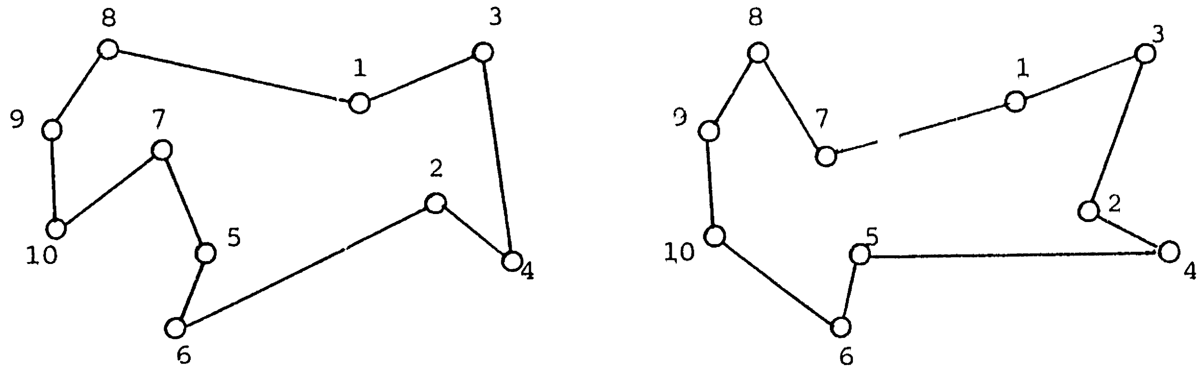
The applications of routing problems in general cover such diverse activities as retail distribution, mail and newspaper delivery, municipal waste collection, fuel oil delivery, etc.; indeed the number of instances in which the methodologies and algorithms thus developed have been used successfully in practical applications has been growing encouragingly over the last several years. We ought to point out, as well, applications of routing problems are not restricted to transportation planning since other settings can give rise to problems with the same mathematical structure (for example Job-Shop scheduling).

This brief discussion shows that literature on vehicle routing problems in a <u>deterministic context</u> is extensive. By a "deterministic context", we mean situations in which the number of "customers", their locations and the size of their demands are known with certainty before the routes are designed. One can identify, however, a practically endless variety of problems in which one or more of these parameters are random variables, i.e. subject to uncertainty in accordance with some probability distribution. In fact, these problems, specified as they are in a probabilistic context, are even more applicable than their deterministic counterparts. While much new ground has already been broken (see Section 1.2), the territory in this area is still virtually unexplored.

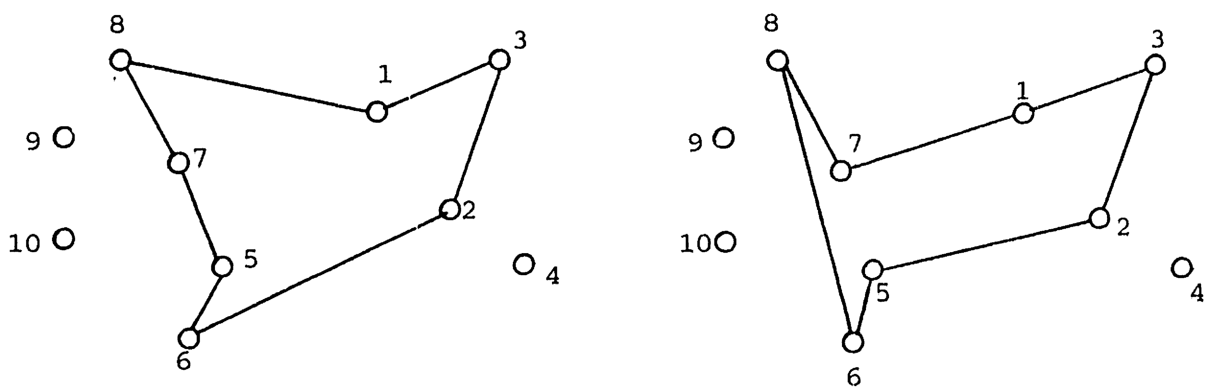
Let us concentrate on the specific case of the TSP and consider the following practical situation: assume a company wants to design a tour through n customers and desires to minimize the routing cost only; it is then legitimate to solve the corresponding TSP if all the customers must actually be visited every day. Assume however that this tour is to be used for a given prolonged period of time (more than one day) and that for this time-horizon, the set of customers to be visited on a daily basis varies; moreover, assume the company can not reoptimize or simply does not desire to reoptimize (see later for a discussion on this distinction) the route (on a daily basis); the vehicle will then follow a pre-designed tour everyday and, on any given day, will simply skip the missing customers from the original tour. The problem is not a TSP anymore, since not only must the tour be a "good" one (that is, have a small routing cost) when all customers are present, but it must also remain "well-behaved" when some customers are skipped from the original set. We have no guarantee that an optimal TSP tour through all the potential points has this desirable property.

This simple observation suggests the formulation and analysis of the following generic problem:

Consider a problem of routing through a set of n points. On any given instance of the problem only a subset consisting of k out of the n points ($0 \leq k \leq n$) must be visited, with the number k determined according to a known probability distribution. We wish to find a priori a tour through all n points. On any given instance of the problem, the k points present will then be visited in the same order as they appear in the a priori tour (see Figure 1.1 for an illustration). The problem of finding such an a priori tour which is of minimum length in the expected




1.1a: two <u>a-priori</u> tours through the same set of 10 points




1.1b: the two resulting tours when the points 4,9, and 10 need not be a visited

<u>Figure 1.1</u>: Simple Graphical Example of a PTSP

value sense is defined as a "probabilistic traveling salesman problem" (PTSP).

One can see that this corresponds to an idealized version of the previously described practical situation. There are at least two possible reasons why the company involved may not reoptimize its route on a daily basis:

1. The company cannot reoptimize its routes due to lack of advance information; that is, the information on who to visit on a particular day is not available at the beginning of the journey but becomes available only along the route itself.

2. The company does know who is to be visited on a particular day but chooses not to reoptimize the route either because it is too expensive to do so or because it prefers that the same driver visit the same stops every day (see Stewart and Golden [1983] for additional comments); this latter strategy obviously promotes regularity of service and covers situations in which a driver's knowledge of the route is an important factor affecting his efficiency and the level of service offered.

In fact the second reason corresponds, perhaps, to a broader range of applications; moreover, progress in communications and processing of information might render the first reason obsolete in the near future.

As an illustration of a PTSP, one can think of the actual case of a postman that delivers mail according to a fixed assigned route. On any particular day, upon delivery at a given location, he checks what address has to be visited next on his regular route and proceeds accordingly. This can be modeled by considering the simplest version of the general framework that we described earlier. By calling $p$ the probability that any particular address out of a set of $n$ addresses will require a visit on any given day (assuming independence between addresses and an equal p for all addresses), the number of addresses requiring a visit is a binomial random variable W with the following probability mass function:

$$Pr(W=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad 0 \leq k \leq n .$$

Motivated by examples such as this one, we will often call p the "coverage probability".

Having introduced the PTSP and before giving an outline of the thesis, let us briefly review the few papers that deal with routing problems having some probabilistic elements.

## 1.2 Brief Literature Review:

Very few researchers have, so far, introduced the element of uncertainty into routing problems, in contrast to the very impressive literature devoted to the deterministic version of these problems. (Of course, one of the reasons why researchers are only now turning to probabilistic types of problems is that deterministic-context routing problems have already proved to be more than sufficiently difficult).

One can identify two types of uncertainty that have been addressed in the literature to date:

1. the distances between points are nondeterministic
2. the demands at individual deliver (pick up) locations behave as random variables

We will now briefly look at the two cases separately.

### 1.2.1 Nondeterministic Distance Between Points

This subsection might itself be subdivided into two parts, depending on the underlying structure of the problem, i.e. on whether the points to be visited are on a graph/network or on a plane.

With an underlying network structure the arcs are themselves assumed to be of random length according to some probability distribution; very few papers treat this stochastic version of routing problems. One can note a paper by Kao [1978] in which a preference-order dynamic program is developed to treat the TSP under nondeterministic travel times. In it, Kao develops a solution procedure for finding a tour with maximum probability of completion by a specified time. In that respect, this paper resembles other attempts to define well-known problems on a probabilistic network (for example the Shortest Path problem, see Frank [1969]). In Leipala [1978] the expectation of the length of the optimal tour is estimated under various assumptions on the probability distribution of the random variables representing the internodal distances.

A different class of papers have been devoted to the study of routing problems in the plane. More precisely, the common approach is to assume that the set of points to be visited is drawn from a sequence of points in a 2-dimensional Euclidean space $R^2$ and that distances between points are given by the Euclidean metric. All these analyses introduce the probabilistic elements by assuming that the points are independently and uniformly distributed over a given area of the plane.

This probabilistic version of the TSP received considerable attention during the 1950's. Earlier papers treated the derivation of upper bounds on the value of the optimal TSP tour through n points lying in a square of side 1 (see Verblunsky [1951] and Few [1955]). It is, however, with Beardwood et al. [1959] that this probabilistic version of the TSP received a thorough and theoretically important treatment; this seminal paper showed that the value of the optimal tour through n points drawn from a uniform distribution in the unit square is almost certainly (i.e., with probability 1) equal to a constant $\beta$ times $\sqrt{n}$. In fact, this result was extended to an arbitrary Lebesgue measurable set of a d-dimensioned space and with an arbitrary probability distribution for the points. The paper shows that the constant of interest $\beta(d)$ depends only on the space dimensionality and not on the shape of the set considered ($\beta(2)$ has been estimated to be 0.765 (Stein [1977])). Although obtained in the 1950's, it was only recently that this result received some well-deserved attention. Karp [1977] was the first to use the theoretical result as a main argument in the probabilistic analysis of heuristics for the TSP. Following Karp's paper, results like that of Beardwood et al. suddenly gained considerable prominence, especially among computer scientists interested in algorithmic applications. Within the context of routing problems, several papers extended the basic results obtained in Beardwood et al.: Stein [1978] obtained a similar result for the single vehicle many-to-many Euclidean Dial-a-Ride problem (this problem can be modeled as a Traveling Salesman Problem with the feasibility constraint that each customer's origin must precede that customer's destination on the route); the first section of Haimovich and Rinnooy Kan [1983] obtained asymptotic expressions for the optimal solution value of the capacitated routing problem under various assumptions on the distribution of customers in the plane and on the capacity of the vehicle. These and other papers develop and analyze partitioning algorithms adopting the general idea contained in Karp [1977] to the specific problem at hand. It is worthwhile mentioning that results have also been obtained in contexts other than that of routing problems. To name a few, one can cite: minimal Matching of a set of random points by euclidean edges (Papadimitriou [1978]); optimal Triangulation of random points in the plane (Steele [1982]); Steinhaus's geometric Location problem for random points in the plane (Fisher and Hochbaum [1980]; Papadimitriou [1981]; Hochbaum and Steele [1982]; and Haimovich [1984]). A very powerful result has also been derived in Steele [1981a]: the author uses the theory of independent subadditive processes to obtain strong limit laws for a class of problems in geometrical probability that exhibit nonlinear growth (the TSP being one of these problems). To conclude, one should mention that this area of research is currently very active and is likely to lead to more exciting results.

### 1.2.2 Nondeterministic Demand at Each Delivery (Pick up) Location

Let us now turn to the other type of uncertainty that has recently been analyzed in routing problems. This subsection is concerned with routing problems for which we wish to design a minimum cost set of routes for a fleet of delivery (pick up) vehicles of fixed capacity; these problems are traditionally labeled VRPs. Except for an isolated analysis in the 1970's (Tillman [1969]), routing problems in which the demands at individual delivery (pick up) locations behave as random variables have received attention only very recently. A review of some results for stochastic vehicle routing is contained in Stewart and Golden [1983]. The approach undertaken is the following: the problems are formulated by using techniques from stochastic programming (i.e. chance-constrained optimization, or stochastic programming with recourse, see, for example, respectively Vajda [1972] and Hansotia [1979]) that allow one to transform these problems into deterministic VRPs.

By doing so, the problems can then be solved by using traditional heuristics developed for the deterministic version of the problem. One of the major drawbacks of these approaches is that it becomes necessary to introduce additional parameters (penalties) whose choice in terms of form and value is at the analysts' discretion and may only vaguely be related to routing costs.

### 1.3 Outline of the thesis:

In section 1.1 we presented the practical motivation behind the introduction of the PTSP. With respect to the brief literature review of section 1.2, one can mention an additional motivation for studying the PTSP, namely the author's desire to analyze a probabilistic version of a combinatorial problem while keeping its original flavor (i.e., without transforming the problem so that it is only indirectly related to its deterministic counterpart).

The central motivation behind our research is to bring new elements to bear in the analysis of routing models, thus making it possible to consider more realistic, hence more complex models. This thesis will then address the following two topics:

(i) Examination of the properties and characteristics of optimal solutions to PTSPs.

(ii) Development of alternative strategies to solve PTSPs.

The outline of the research is as follows: First in Chapter 2 we present efficient ways for computing the expected length of any given PTSP tour; after demonstrating (in section 2.1) the inefficiency of computing this expected length by explicit enumeration ($O(n2^n)$ steps), we present, in section 2.2, the derivation of a closed-form expression giving the expected length of any given tour through n points as a function of the "coverage" probability p and of n-1 fixed quantities intimately linked to the tour. This expression is obtained without any restrictions on the matrix of distances between points. We then show, using this expression and through a simple example, that the optimal TSP tour can be a very poor solution to the corresponding PTSP problem with a given probability p. In section 2.3 we extend our results in several directions. We first consider problems where some points are always present (one point, then m such points), the others being present only with probability p as previously assumed. We then define and analyze the Probabilistic Hamiltonian Path Problem (PHPP). Finally, in a fourth subsection we generalize all our previous results by considering more general probabilistic assumptions about the presence of points. When generalization of the closed-form expression is not possible we give an alternative way (recursive relationships) to compute efficiently the expected length of any given fixed tour.

In Chapter 3, after providing a general "weight-form" representation of every closed-form expression of Chapter 2 by means of a single formula, we first derive, in section 3.2, properties of the two classes of elements of which this expression consists. These properties are then used to derive lower and upper bounds on the expected length of a tour. In section 3.3 we try to obtain, using, in parts, results from section 3.2, some useful characterizations and/or exploitable properties of the problem structure: we first show that there are some manipulations of the distance matrix that leave the problem unchanged; we then show that, surprisingly, the optimal PTSP tour in the Euclidean plane can intersect itself; we finally consider the behavior of the expected length of a given tour with respect to perturbations of the graph such as adding or deleting a node. Section 3.4 is then concerned with finding some relationships between the optimal TSP tour and the optimal PTSP tour: after providing some results for problems of small size, we then derive, using results from section 3.2, worst-case bounds on the absolute difference between the expected length of the optimal TSP tour and the expected length of the optimal PTSP tour. We discuss the tightness of our bounds and we provide insights into some very peculiar aspects of the PTSP problem. We also point out that all our previous results are also valid for the PHPP problem and finally we present a useful result based on a variation of the PHPP.

Chapter 4 is concerned with methods of analysis considerably different from those of the previous two chapters. An asymptotic approach is taken in which set-theoretic concepts are used instead of graph-theoretic ones. We consider a set of points in 2-dimensional Euclidean space $R^2$, assuming the distance between points to be the ordinary Euclidean distance. In section 4.2 we present an upper bound on the expected length of the optimal PTSP tour for an arbitrary sequence of n points lying in a square of side r. Assuming the points are uniformly and independently distributed over the square, we obtain a second upper bound as well as a lower bound on the expected length of the optimal PTSP tour. In section 4.3 we turn our attention to asymptotic behavior (i.e., $n \to \infty$). In a first subsection we present the asymptotic behavior of the expected value of the length of the tour obtained through the strategy of reoptimizing the optimal tour for each realization of the random variables (i.e., for each subset of the points that will actually need a visit on a particular problem instance, we construct the optimal TSP tour and compute its length). This, of course, constitutes a lower bound on the expected length of the optimal PTSP tour. The second subsection contains the most important theoretical result of this chapter: we show that the expected length of the optimal PTSP tour (coverage probability p) through n points drawn from a uniform distribution in the unit square is almost surely (with probability 1) asymptotic to $c(p)\sqrt{n}$, where c(p) is a constant depending on the coverage probability p. The third subsection is then concerned with the derivation of bounds on c(p).

We then present, in section 4.4, generalizations of our results in several directions: first, we note that all our previous results extend to the case where one of the points is always present (a depot); we then present extensions for cases where more than one point is always present. We also discuss extensions of all our results to any bounded Lebesgue measurable set of a d-dimensioned Euclidean space and, after mentioning the use of other metrics, we conclude the chapter with a brief discussion of the practical implications of our results.

Chapter 5 is concerned with using the extensive theoretical investigation provided in Chapters 2,3,4 to develop solution procedures for solving the PTSP. Our emphasis is on the conceptualization of solution procedures and all the proposed algorithms are based on a theoretical foundation provided by the previous chapters. In section 5.2 we are concerned with exact optimization methods for solving the PTSP. We first show how one can use the integer linear programming formulation of the TSP to formulate the PTSP as an integer nonlinear programming problem; then, we successively transform this formulation first, to a mixed integer linear program, and finally, to a pure integer linear program. In a third subsection, we discuss the relative merits of those three formulations, followed by a proposed Branch-and-Bound procedure. We conclude section 5.2 by showing that the relationship between the PTSP and TSP is not as simple as one could imagine. Indeed, we will show that a seemingly natural extension of the dynamic programming formulation of the TSP does not solve the PTSP, and that, in fact, one cannot use dynamic programming approaches to provide an exact solution procedure for this problem.

The second main section (5.3) contains an exposition of heuristic procedures (that is, not guaranteed to obtain an optimal solution). After providing a brief discussion on the necessity of developing such procedures for a complex problem like the PTSP, we first present some theoretical preliminaries, on which the proposed procedures will be built. Based on those preliminaries, we present a host of procedures under the generic term of Tour construction procedures (a term originally used for the TSP). First, we discuss extensions of the Clarke and Wright savings approach (see Golden et al. [1980], for example) and label them Supersavings Algorithms. Then we also introduce the "Almost" Nearest Neighbor Algorithm and finally conclude this section on Tour construction procedures by listing several "insertion" procedures. In a second subsection 5.3.3 we briefly mention the use of "hill-climbing" methods (similar to the "2-opt" or "3-opt" heuristics proposed for the TSP). Finally, in a third subsection we consider the PTSP in the plane. Based mainly on results from Chapter 4, this section will analyze a recent heuristic for the TSP based on spacefilling curves (see Platzman and Bartholdi [1983]) and we also look at procedures based on partitioning approaches (see Karp [1977]). We conclude Chapter 5 with a review of the most interesting results and most promising approaches proposed to solve the PTSP.

Finally in Chapter 6, after a brief section on the applications of the methodology to preliminary planning and to location of facilities, we review the major results of the research and then discuss further research on the PTSP as well as the general idea of integrating the consideration of uncertainty into combinatorial optimizations problems.

# CHAPTER 2

# THE EXPECTED LENGTH OF TOURS UNDER VARIOUS CONDITIONS

## 2.1 Introduction

### 2.1.1 The Expected Length "In The PTSP Sense"; Outline of Chapter 2

Let us first briefly reintroduce the PTSP; as indicated in Chapter 1, we are concerned with a variation of the TSP in which it is no longer certain that each of the n points (i.e. customers) must be visited; rather each point is present only with a fixed probability p (independently of each other). A tour through all points has to be found before knowing which points actually have to be visited; once a tour (sequence of points) is given, missing points will simply be skipped. Hence, for each realization of the random variables (each corresponding to a specific subset of the points that will actually need a visit), the relative sequence of the remaining points along the tour will not be changed. A tour has, by definition, a random total length; the problem of interest (PTSP) is to find a predetermined tour of minimum expected total length.

It should be stressed that, for a given tour, the computation of its expected length (as mentioned previously) corresponds to the strategy of visiting the points (requesting a visit) according to the prescribed sequence in which they appear in the predetermined tour (the other points being skipped); we will sometimes refer to this strategy as "in the PTSP sense".

Consider now a given tour; a way to compute its expected length would be to enumerate all cases and report for each of them the resulting length of the tour under consideration; however, there are $2^n$ such cases (indeed, each point is either present or not, independently of the n-1 other nodes) and, if we assume that computing the length of a tour given by a sequence of k points requires k-1 steps (basic additions), this method will then require

$$\sum_{k=1}^{n} (k-1) \binom{n}{k} = (n-1)2^n \qquad \text{number of steps.}$$

This is clearly not satisfactorily and one of the goals of this chapter is to reduce substantially the number of computations required to get the expected length of a given tour.

After specifying the notation to be used in the statements and proofs of our results (subsection 2.1.2), we will analyze the expected length "in the PTSP sense" under various conditions. In section 2.2 we derive a closed-form expression giving the expected length of any given tour through n points as a function of the "coverage" probability p and of n-1 fixed quantities intimately linked to the tour. This expression is obtained without any restrictions on the distance between points. We then show, using this expression and through a simple example that the optimal TSP tour can be a very poor solution to the corresponding PTSP problem with a given probability p; this result provides an additional motivation for studying in detail the PTSP. In a third section we extend our results in several directions. We first consider problems in which some points are always present (one point, then m such points), the others being present only with probability p as previously assumed; we then define and analyze the Probabilistic Hamiltonian Path Problem (PHPP); finally, in the fourth subsection we generalize all our previous results by considering more general probabilistic assumptions about the presence of points. When generalization of the closed-form expression is not possible we give an alternative way (recursive relationships) to compute efficiently the expected length of any given fixed tour.

### 2.1.2 Background Information; Notation

Throughout this chapter $G = (N,A,D)$ denotes a complete, directed graph where

$N$ = the node set of cardinality $|N|$ (set of customers)

$A$ = the set of arcs joining the nodes of $N$

$D$ = distance (cost) matrix; the elements of the matrix $D$, $d(i,j)$, represent the distance (cost) from node $i$ to node $j$, i.e. the weight of arc $(i,j)$. Unless otherwise specified, we assume a general distance matrix (not necessarily respecting a metric).

$t = (i_1, i_2, \ldots, i_{|N|}, i_1)$ will represent a sequence of nodes forming a Hamiltonian circuit (tour) of the graph $G$. The length $L(t)$ of the tour $t$ is determined by

$$L(t) = \sum_{j=1}^{|N|} d(i_j, i_{j+1}) \text{ where } i_{|N|+1} \equiv i_1 \qquad (2.1)$$

The TSP asks for a tour $t$ which minimizes $L(t)$.

An alternative way of formulating the TSP is by considering the set of all cyclic permutations $\Pi$ on $|N|$ objects. A cyclic permutation $\Pi$ represents a tour $t$ if we interpret $\Pi(j)$ to be the node visited after node $j$, $j \varepsilon [1..|N|]$. Then the weight $w(\Pi)$ of the permutation $\Pi$ is given by

$$w(\Pi) = \sum_{j=1}^{|N|} d(j,\Pi(j)) \tag{2.2}$$

and of course is identical to the length L(t) of the corresponding tour t.

Notation and Assumptions:

In this chapter each node of G is colored either white or black.

A black node of G is a node that will always require a visit during each execution of a tour t. $N_1$ will be the set of black nodes with $|N_1| = m$.

A white node of G, on the other hand, is a node that will not always require a visit. Unless otherwise specified (see section 2.3.4), each white node is present with a fixed probability p, independently of each other - (in a general setting, P will represent the set of rules defining the probabilistic assumptions about the white nodes) - $N_2$ will be the set of white nodes ($N_1 \cup N_2 = N$, $N_1 \cap N_2 = \phi$) and $|N_2| = n$.

A tour $t = (i_1, i_2, \ldots, i_{|N|}, i_1)$ is cyclic by definition; we will then adopt the following notational simplification on indices: $i_{j+r+1}$ (with j running from 1 to $|N|$) will stand for $i_{j+r \bmod(|N|)+1}$ (i.e. $i_{|N|+1} \equiv i_1$ etc.)

We can now formulate the general PTSP as follows:

> Given G = (N,A,D,P) find the fixed tour
>
> $t = (i_1, \ldots, i_{|N|}, i_1)$ of minimum expected length.

Before presenting our results, let us indicate that graphical illustrations of some of the quantities involved in the theorems are provided in Appendix A in order to give some intuitions behind their (otherwise) formal introduction.

## 2.2 The Expected Length of a Tour t in a Graph G With No Black Node

In this case we have $|N_1| = 0$, $|N| = |N_2| = n$.

Let $t = (i_1, i_2, \ldots, i_n, i_1)$ be a given tour of G.

Let us introduce n-1 quantities obtained from t:

$$L_t^{(r)} = \sum_{j=1}^{n} d(i_j, i_{j+1+r}) \qquad \forall\, r \in [0..n-2] \qquad (2.3)$$

Note:

- for r=0 $L_t^{(o)}$ is simply the length L(t) of the tour t (as given in (2.1) with $|N| = n$)

- for r in general, $L_t^{(r)}$ is the sum of n elements, each representing the distance from the node $i_j$ to its $(r+1)^{th}$ successor with respect to the tour t (that is, from each $i_j$ we "skip" the first r nodes appearing on the tour t). With respect to the formulation given in (2.2) is is easily seen that $L_t^{(r)}$ can be written

$$L_t^{(r)} = \sum_{j=1}^{n} d\left(j, \Pi^{r+1}(j)\right) \qquad (2.4)$$

where $\Pi^{r+1}(j)$ is defined recursively as $\Pi^{r+1}(j) \equiv \Pi\left(\Pi^r(j)\right)$. We now have all the elements to state the following lemma:

Lemma 2.1:

Given a graph G with only white nodes, and a coverage probability p, the conditional expected length of a tour t, given k white nodes are missing from the original tour, is:

(i) $E[L_t | k \text{ missing nodes}] = 0$ if $k=n-1$ or $k=n$

(ii) $E[L_t | k \text{ missing nodes}] = \left(1/\binom{n}{k}\right) \left[ \sum_{r=0}^{k} \binom{n-2-r}{k-r} L_t^{(r)} \right]$

if $k \in [0..n-2]$

Proof: (i) is obvious. We will prove (ii) by using a combinatorial argument consisting of two observations:

1) there are $\binom{n}{k}$ different and equiprobable possible ways of having k missing nodes out of n (or equivalently of having n-k remaining nodes). If then we call $D_q$ the resulting length of the tour t for the $q^{th}$ way, $q \in \left[1..\binom{n}{k}\right]$, of having k missing customers, then, simply, $E[L_t | k \text{ missing nodes}] = 1/\binom{n}{k} \sum_{q=1}^{\binom{n}{k}} D_q$. It remains now to evaluate $\sum_{q=1}^{\binom{n}{k}} D_q$.

2) When k nodes are missing, the tour t consists only of n-k nodes and its resulting length is thus the sum of n-k elements of the distance matrix D. So $\sum_{q=1}^{\binom{n}{k}} D_q$ consists of the sum of $\binom{n}{k}$ (n-k) elements of D; as some elements will appear more than once (in different $D_q$'s) we simply have to regroup them differently (according to the number of times they appear in the computation of the $D_q$'s).

Let us take $d(i_j, i_{j+r+1})$ (that is, a generic element of $L_t^{(r)}$) for a given $r \in [0..n-2]$; for this element to be part of the computation of the $D_q$'s

- nodes $i_j$ and $i_{j+r+1}$ <u>have</u> to be present, <u>and</u>

- nodes $i_{j+1}, \ldots, i_{j+r}$, i.e., r consecutive nodes, <u>have</u> to be missing.

Since we have <u>only</u> k missing nodes:

- if $r > k$, $d(i_j, i_{j+r+1})$ will <u>never</u> appear and then $L_t^{(r)}$ for $r > k$ will not be part of the computation of $E[L_t | k \text{ missing nodes}]$.

- if $r \leq k$, then we have the freedom to choose k-r nodes to be missing (r having been determined already) among n-2-r, i.e. a total of $\binom{n-2-r}{k-r}$ ways. This is true $\forall\, j \in [1..n]$ and hence $\binom{n-2-r}{k-r}$ is the number of times $L_t^{(r)}$ will appear in $\sum_q D_q$.

We can check that, regrouping elements this way, we did not miss any by simply noting that each $L_t^{(r)}$ is composed of the sum of n elements. Then $\sum_{r=0}^{k} \binom{n-2-r}{k-r} L_t^{(r)}$ will be composed of

$$n \sum_{r=0}^{k} \binom{n-2-r}{k-r} = n\binom{n-1}{k} = (n-k) \binom{n}{k} \text{ elements}$$

which is the number of elements involved in $\sum_q D_q$

Q.E.D.

We can now state our basic result giving the unconditional expected length of a tour t:

<u>Theorem 2.1</u>:

Given a graph G with only white nodes and a coverage probability p, the expected length of a tour t is:

$$E[L_t] = p^2 \left[\sum_{r=0}^{n-2} (1-p)^r L_t^{(r)}\right]$$

Proof: we have:

$$E[L_t] = \sum_{k=0}^{n} E[L_t | k \text{ missing nodes}] \Pr\{k \text{ missing nodes}\}.$$

$$\Pr\{k \text{ missing nodes}\} = \binom{n}{k} p^{n-k}(1-p)^k$$

We obtain, using Lemma 2.1,

$$E[L_t] = \sum_{k=0}^{n-2} \left[(1/\binom{n}{k}) \sum_{r=0}^{k} \binom{n-2-r}{k-r} L_t^{(r)} \right]\left[\binom{n}{k} p^{n-k}(1-p)^k\right] \tag{2.5}$$

$$= \sum_{r=0}^{n-2} L_t^{(r)} \left[\sum_{k=r}^{n-2} \binom{n-2-r}{k-r} p^{n-k}(1-p)^k\right].$$

Setting $u \equiv k-r$ and $s \equiv n-2-r$ we have

$$\sum_{k=r}^{n-2} \binom{n-2-r}{k-r} p^{n-k}(1-p)^k = p^2(1-p)^r\left(\sum_{u=0}^{s} \binom{s}{u} p^{s-u}(1-p)^u\right)$$

$$= p^2(1-p)^r[p+(1-p)]^s = p^2(1-p)^r$$

Q.E.D.

As can be seen from (2.3) the computation of each $L_t^{(r)}$ (given a tour t of a graph G) requires n-1 additions, $r \in [0..n-2]$. Hence, using Theorem 2.1, we can determine the expected length of any tour in $O(n^2)$ steps. Moreover, since we have a closed form expression, once the $L_t^{(r)}$ are determined, we can, in $O(n)$ steps, compute the expected length of the tour t for different values of the coverage probability p.

Let us present a numerical example illustrating these points; the example will also provide additional motivation for studying the PTSP. Very often in practice, companies design delivery or pick-up routes with reference to their full set of customers and do not reoptimize routes to take account of daily variations in demand; our example will show that directly taking such variations into account in designing the initial route might lead to substantial routing cost savings compared to the traditional application of the TSP; in other words we show that the optimal TSP tour can be a very poor solution by comparison to the optimal PTSP tour if the set of customers to be visited is not fixed and if reoptimization for each instance of the problem is not considered - (for some problems reoptimization may even be infeasible due to lack of sufficient advanced information).

In this example, the graph G contains 24 white nodes that are positioned at the vertices of two concentric 12-gons as shown in Figure 2.1. In Figure 2.2 two tours have been designed through this set of nodes; tour a is the optimal TSP tour, tour b is an alternative tour. Although L(b) is greater than L(a), for p=0.5 the expected length of tour a is 30% larger than the expected length of tour b. More precisely, for this example $G = (N_1 \cup N_2, A, D)$ is such that:

$|N_1| = 0$ $|N_2| = 24$ and D is given by the Euclidean distance between the nodes. The nodes correspond to the vertices of two concentric regular 12-gons; the inside 12-gon is such that each of its vertices is positioned "between" two successive vertices of the outside 12-gon:

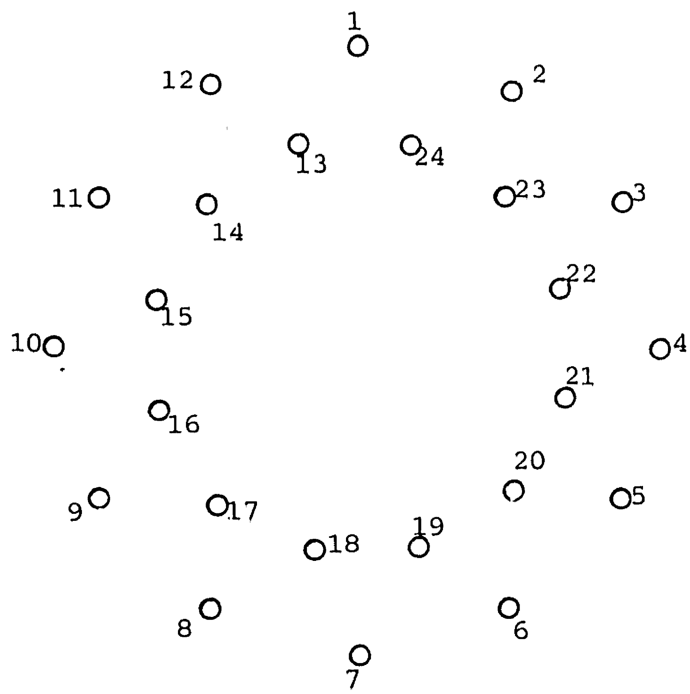



Figure 2.1: a 24 Nodes Graph

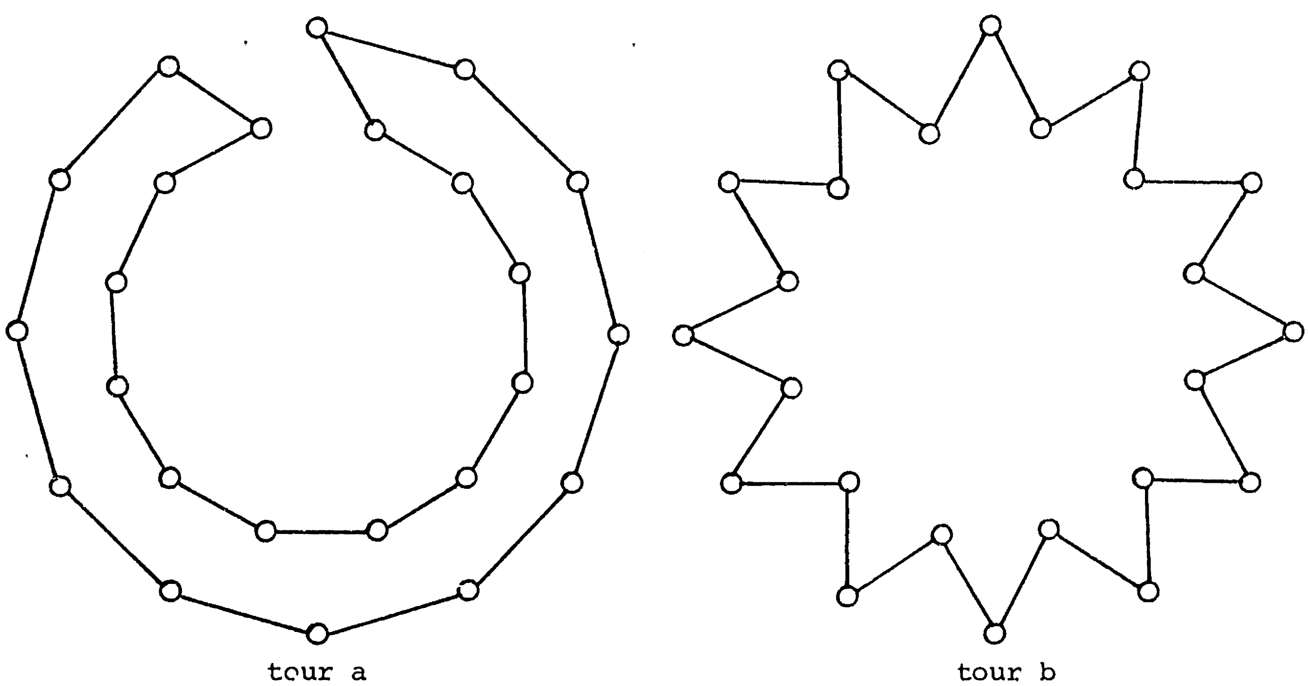



Figure 2.2: Two Tours of the 24 Nodes Graph

the distance between two successive vertices of the outside 12-gon is 1

the distance between two successive vertices of the inside 12-gon is 0.629

With this information, we can compute all distances between pairs of nodes and obtain D.

Let us give the $L^{(r)}$'s for tour a and tour b of figure 3:

| tour a | tour b |
|---|---|
| $L_a^{(0)} = L_a^{(22)} = 19.561$ | $L_b^{(0)} = L_b^{(22)} = 19.704$ |
| $L_a^{(1)} = L_a^{(21)} = 35.86$ | $L_b^{(1)} = L_b^{(21)} = 19.548$ |
| $L_a^{(2)} = L_a^{(20)} = 48.445$ | $L_b^{(2)} = L_b^{(20)} = 32.976$ |
| $L_a^{(3)} = L_a^{(19)} = 57.064$ | $L_b^{(3)} = L_b^{(19)} = 37.764$ |
| $L_a^{(4)} = L_a^{(18)} = 61.856$ | $L_b^{(4)} = L_b^{(18)} = 47.952$ |
| $L_a^{(5)} = L_a^{(17)} = 63.304$ | $L_b^{(5)} = L_b^{(17)} = 53.412$ |
| $L_a^{(6)} = L_a^{(16)} = 62.184$ | $L_b^{(6)} = L_b^{(16)} = 60.840$ |
| $L_a^{(7)} = L_a^{(15)} = 59.428$ | $L_b^{(7)} = L_b^{(15)} = 65.412$ |
| $L_a^{(8)} = L_a^{(14)} = 56.047$ | $L_b^{(8)} = L_b^{(14)} = 70.080$ |
| $L_a^{(9)} = L_a^{(13)} = 52.984$ | $L_b^{(9)} = L_b^{(13)} = 72.960$ |
| $L_a^{(10)} = L_a^{(12)} = 51.067$ | $L_b^{(10)} = L_b^{(12)} = 74.928$ |
| $L_a^{(11)} = 51.08$ | $L_b^{(11)} = 75.528$ |

tour a is the optimal TSP tour of length 19.561

tour b is an alternative tour of length 19.704

However, for p=0.9, we have from Theorem 2.1

$$E[L_a] \cong 19.193 > E[L_b] \cong 17.846,$$

i.e. $E[L_a]$ is 7% greater

and for p = 0.5

$$E[L_a] \cong 16.110 > E[L_b] \cong 12.286,$$

i.e. $E[L_a]$ is 31% greater

2.3 Extensions:

In section 2.2 all nodes of G were white. In practical routing applications, however, one of the nodes generally represents the depot where the tour begins and ends; this depot is fixed and is always present. It is therefore important to extend out results to graphs G where $N_1 \neq \phi$; we will do so by first considering $|N_1| \equiv m = 1$ (one depot) and then a general m.

Then, after introducing the Probabilistic Hamiltonian Path Problem and giving similar expressions for it, we will finally consider more general probabilistic assumptions for all previous problems.

2.3.1 The Expected Length of a Tour t in a Graph G With One Black Node

We have $|N_1| = 1$, $|N_2| = n$, hence $|N| = n+1$. Let $i_1$ be the black node and let $t = (i_1, \ldots, i_{n+1}, i_1)$ be a given tour of G.

Here again we shall introduce quantities (n, this time) obtained from t; as they differ slightly from the $L_t^{(r)}$ defined in section 2.2 (recall that now we have a black node which always requires a visit) we will use

$L^{(r)}_{1,t}$ for this subsection and

$L^{(r)}_{m,t}$ for the next subsection where G

contains m black nodes.

$L^{(r)}_{1,t}$ is defined as follows:

- $L^{(o)}_{1,t} = \sum_{j=1}^{n+1} d(i_j, i_{j+1})$ (hence $L^{(o)}_{1,t}$ still represents the

length of the tour t) and

- $L^{(r)}_{1,t} = \sum_{j=1}^{n+1} d_{1,t}(i_j, i_{j+1+r})$ $\forall$ r $\varepsilon$ [1..n-1]

where $d_{1,t}(.,.)$ is obtained from d(.,.) according to the following rules:

$\forall$ r $\varepsilon$ [1..n-1]

(a) $d_{1,t}(i_j, i_{j+1+r}) = d(i_j, i_{j+1+r})$ if $1 \leq j \leq n-r+1$

(2.6)

(b) $d_{1,t}(i_j, i_{j+1+r}) = d(i_j, i_1) + d(i_1, i_{j-n+r})$ if $n-r+1 < j \leq n+1$

Note that in case (b) above, node $i_1$ is one of the first r nodes following $i_j$ on the tour t and that it cannot be "skipped". In other words, for $n-r+1 < j \leq n+1$, to go from $i_j$ to its $(r+1)^{th}$ successor we have to "visit" node $i_1$ first.

Lemma 2.2:

Given a graph G with n white nodes, one black node, and a coverage probability p, the conditional expected length of a tour t, given k white nodes are missing from the original tour is:

(i) $E[L_t \mid n \text{ missing nodes}] = 0$ if $k = n$

(ii) $E[L_t \mid n-1 \text{ missing nodes}] = (1/n)L_{1,t}^{(n-1)}$ if $k=n-1$

(iii) $E[L_t \mid k \text{ missing nodes}] = \left(1/\binom{n}{k}\right)\left[\sum_{r=0}^{k} \binom{n-2-r}{k-r} L_{1,t}^{(r)}\right]$ if $k \in [0..n-2]$

Proof:

(i) is obvious. (ii) is also obvious from the definition of $L_{1,t}^{(n-1)}$ and the fact that each white node of G has an equal probability of being the only one present (indeed $L_{1,t}^{(n-1)}$ can be equivalently written from its definition as follows: $L_{1,t}^{(n-1)} = \sum_{j=2}^{n+1} \left(d(i_j,i_1) + d(i_1,i_j)\right)$. The proof of (iii) follows with some minor changes the one for Lemma 2.1. The first observation in the proof of Lemma 2.1 remains valid; for the second observation the proof goes as follows: the elements $d(i_j,i_{j+r+1})$ $\forall$ $r \in [0..k]$ still appear $\binom{n-2-r}{k-r}$ times, except if $i_j \equiv i_1$ or $i_{j+r+1} \equiv i_1$; in those cases the elements $d(i_j,i_{j+r+1})$ appear $\binom{n-1-r}{k-r}$ times (since now, out of the n white nodes, only <u>either</u> $i_j$ <u>or</u> $i_{j+r+1}$ has to be present, and not both as previously). This complication does not create a problem since terms like $d(i_1,i_{j+r+1})$ or $d(i_j,i_1)$ do not appear only in $L_{1,t}^{(r)}$ but also in $L_{1,t}^{(r+1)}$, $L_{1,t}^{(r+2)}$, ..., $L_{1,t}^{(k)}$ (according to our definition of those quantities). Hence keeping the same weight (as in Lemma 2.1) for each $L_{1,t}^{(r)}$ (i.e. $\binom{n-2-r}{k-r}$) is still valid; indeed for terms unique to $L_{1,t}^{(r)}$ this weight represents the true number of times this element remains on the tour t when k nodes are missing, and for the elements belonging to $L_{1,t}^{(r)}$, ... $L_{1,t}^{(k)}$ these weights will lead to $\sum_{\ell=0}^{k-r} \binom{n-2-r-\ell}{k-r-\ell} = \binom{n-1-r}{k-r}$, which is the quantity we derived previously. Q.E.D.

Theorem 2.2 can now be stated; its proof is not given since it follows exactly the lines of the proof of Theorem 1 (using Lemma 2.2 instead of Lemma 2.1).

Theorem 2.2:

Given a graph G with n white nodes, one black node, and a coverage probability p, the expected length of a tour t is:

$$E[L_t] = p^2\left[\sum_{r=0}^{n-2} (1-p)^r L_{1,t}^{(r)}\right] + p(1-p)^{n-1} L_{1,t}^{(n-1)}$$

2.3.2 The Expected Length of a Tour t in a Graph G With m Black Nodes:

We have $|N_1| = m$, $|N_2| = n$, $|N| = n+m$. Let $t = (i_1, i_2, \ldots, i_{n+m}, i_1)$ be a given tour of G. Let us again define quantities as follows:

- $L_{m,t}^{(o)} = \sum_{j=1}^{n+m} d\left(i_j, i_{j+1}\right)$ (as before it represents the length of the tour t through the n+m nodes of G)

- $L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}\left(i_j, i_{j+1+r}\right) \quad \forall\, r \in [1..n-1]$

where $d_{m,t}(.,.)$ is obtained from $d(.,.)$ according to the following rules:

$\forall\ r \in [1..n-1]$

(a) $d_{m,t}(i_j, i_{j+r+1}) = d(i_j, i_{j+r+1})$ (2.7)

whenever the nodes $i_{j+1}, \ldots, i_{j+r}$ are all <u>white</u> nodes

(b) $d_{m,t}(i_j, i_{j+r+1}) = \sum_{e=0}^{s} d(k_e, k_{e+1})$

where $k_o \equiv i_j$, $k_{s+1} \equiv i_{j+r+1}$, and where $(k_1, k_2, \ldots, k_s)$ is the sequence of black nodes drawn from $(i_{j+1}, \ldots, i_{j+r})$ $(s \in [1..\min(m,r)])$

- finally we define $L^{(n)}_{m,t}$ to be the length of the tour t through the m black nodes (i.e. when no white nodes are present).

Given these definitions we can now state two general results. We will do so without giving proofs since, although cumbersome, these proofs are straightforward extensions of the ones previously given in subsection 2.3.1 and section 2.2.

<u>Lemma 2.3</u>:

Given a graph G with n white nodes, m black nodes, and a coverage probability p, the conditional expected length of a tour t, <u>given</u> k white nodes are missing from the original tour, is:

(i) $E[L_t \mid n \text{ missing nodes}] = L^{(n)}_{m,t}$ if k=n

(ii) $E[L_t \mid n-1 \text{ missing nodes}] = (1/n)\ L^{(n-1)}_{m,t}$ if k=n-1

(iii) $E[L_t \mid k \text{ missing nodes}] = \left(1/\binom{n}{k}\right)\left[\sum_{r=0}^{k} \binom{n-2-r}{k-r} L^{(r)}_{m,t}\right]$

if $k \in [0..n-2]$

Theorem 2.3

Given a graph G with n white nodes, m black nodes, and a coverage probability p, the expected length of a tour t is:

$$E[L_t] = p^2\left[\sum_{r=0}^{n-2} (1-p)^r L_{m,t}^{(r)}\right] + p(1-p)^{n-1} L_{m,t}^{(n-1)} + (1-p)^n L_{m,t}^{(n)}$$

We now present similar results for a related problem.

### 2.3.3 The Probabilistic Hamiltonian Path Problem: (PHPP)

This is essentially the same problem as the PTSP except that in addition to the graph G and the coverage probability p, we need to specify two fixed nodes between which a Hamiltonian path is to be constructed.

We use the same notation as before for describing the underlying graph G (note that now we require $m \geq 2$ since two black nodes have to be specified). The starting node and the finishing node of the Hamiltonian path will always be labeled $i_1$ and $i_{|N|}$.

Let $h = (i_1, i_2, \ldots, i_{|N|})$ represent a Hamiltonian path from $i_1$ to $i_{|N|}$ of the graph G with length $L(h) = \sum_{j=1}^{|N|} d(i_j, i_{j+1})$

The formulation of the general PHPP is then:

Given $G = (N,A,D,P)$, and two black nodes $i_1$, $i_{|N|}$, find the Hamiltonian path $h = (i_1, \ldots, i_{|N|})$ between node $i_1$ and node $i_{|N|}$ of minimum expected length.

In this subsection, we do not consider tours anymore, so that the assumptions for indices are as follows: for $j \in [1..|N|-1]$, $i_{j+r+1}$ will stand for $i_{(j+r-1 \bmod (|N|-1)+1)}$ (i.e., $i_{|N|+1} \equiv i_2$ and so on).

For the problem of finding the expected length of a Hamiltonian path h, we will present the results for the general case where $|N_1| = m \geq 2$, $|N_2| = n$ and the coverage probability is p.

Let us define the $L^{(r)}_{m,h}$'s:

- $L^{(o)}_{m,h} = \sum_{j=1}^{n+m-1} d(i_j, i_{j+1})$ (as in all previous sections, this

represents the length L(h) of the path h)

- $\forall\, r \in [1..n-1]$

$$L^{(r)}_{m,h} = \sum_{j=1}^{n+m-1} d_{m,h}(i_j, i_{j+r+1})$$

where $d_{m,h}(.,.)$ is obtained from $d(.,.)$ in the same fashion as $d_{m,t}(.,.)$ (see (2.7)); in fact for the purpose of the construction of $d_{m,h}(.,.)$ it is very convenient to think of $i_1$ and $i_{|N|}$ to be the <u>same</u> node and to apply the rules (2.7). Hence, for example,

$$d_{m,h}(i_{n+m-1}, i_2) \equiv d(i_{n+m-1}, i_{n+m}) + d(i_1, i_2)$$

and <u>not</u> $d(i_{n+m-1}, i_{n+m}) + d(i_{n+m}, i_1) + d(i_1, i_2)$

- finally let $L^{(n)}_{m,h}$ be the length of the resulting path when all the n white nodes are missing.

We can state the following two results; again the proofs are not given since they are straightforward extensions of those in section 2.2 and subsection 2.3.2.

Lemma 2.4:

Given a graph G with n white nodes, m black nodes - two of them being $i_1$ and $i_{|N|}$ - and a coverage probability p, the conditional expected length of a Hamiltonian path h, given k white nodes are missing from the original path, is:

(i) $E[L_h | n \text{ missing nodes}] = L^{(n)}_{m,h}$ if k=n

(ii) $E[L_h | n-1 \text{ missing nodes}] = (1/n)\ L^{(n-1)}_{m,h}$ if k=n-1

(iii) $E[L_h | k \text{ missing nodes}] = \left(1/\binom{n}{k}\right) \left[ \sum_{r=0}^{k} \binom{n-2-r}{k-r} L^{(r)}_{m,h} \right]$

if $k \in [0..n-2]$

This lemma leads to the next theorem:

Theorem 2.4:

Given a graph G with n white nodes, m black nodes - two of them being $i_1$ and $i_{|N|}$ - and a coverage probability p, the expected length of a Hamiltonian path h from $i_1$ to $i_{|N|}$ is:

$$E[L_h] = p^2 \left[ \sum_{r=0}^{n-2} (1-p)^r L^{(r)}_{m,h} \right] + p(1-p)^{n-1} L^{(n-1)}_{m,h} + (1-p)^n L^{(n)}_{m,h}$$

One can note that an alternative way of obtaining $E[L_h]$ is to proceed as follows: first transform the Hamiltonian path $h = (i_1, \ldots, i_{|N|})$ into a tour $t = (i_1, \ldots, i_{|N|}, i_1)$ by adding the arc $(|N|, 1)$; then compute $E[L_t]$ by using Theorem 2.3; finally $E[L_h]$ is obtained from $E[L_t]$ by subtracting $d(i_{|N|}, i_1)$, i.e.

$$E[L_h] = E[L_t] - d(i_{|N|}, i_1)$$

This approach corresponds to an alternative (but equivalent) definition of $L_{m,h}^{(r)}$ (in function of $L_{m,t}^{(r)}$ as defined by the rules (2.7)):

$$L_{m,h}^{(r)} = L_{m,t}^{(r)} - (r+1)\ d(i_{|N|}, i_1) \text{ for } r \in [0..n-1]$$

$$L_{m,h}^{(n)} = L_{m,t}^{(n)} - d(i_{|N|}, i_1)$$

Let us briefly mention a slightly modified version of the Probabilistic Hamiltonian Path Problem for which no previous results can be adequately used so that a special treatment is needed. This version corresponds to the case in which all nodes (including $i_1$ and $i_{|N|}$) are whites. Let $h = (i_1, \ldots, i_n)$ be a hamiltonian path through n white nodes; here again by defining n-1 quantities $L_h^{(r)}$ ($r \in [0..n-2]$) one can easily show that $E[L_h]$ can be obtained by using a closed-form expression identical to the one given in Theorem 2.1 (replacing $L_t^{(r)}$ by $L_h^{(r)}$); for this problem the $L_h^{(r)}$ are defined as follows:

$$L_h^{(r)} = \sum_{j=1}^{n-1-r} d(i_j, i_{j+r+1}) \qquad \forall\ r \in [0..n-2]$$

Note:

- $L_h^{(o)}$ is still the length of h
- $L_h^{(r)}$ is composed of only n-1-r elements

- In the case of n white nodes and one black node ($i_1$), the Hamiltonian path $h=(i_1, \ldots, i_{n+1})$ has an expected length given by Theorem 2.2 (replacing $L_t^{(r)}$ by $L_h^{(r)}$); the $L_h^{(r)}$ are defined as follows:

$$L_h^{(r)} = \sum_{j=1}^{n-r} d(i_j, i_{j+r+1}) + \sum_{j=0}^{r} d(i_1, i_{j+1}) \quad \forall\ r \in [0..n-1]$$

Before turning to the next section in which more general probabilistic assumptions will be made, let us make a brief digression on the efficiency of the computation of the expected length of a tour t or a Hamiltonian path h (in a graph G with n white nodes and m black nodes) using our previous results.

Let us take first the case of a Hamiltonian path h of G; although Theorem 2.4 is very important for obtaining a closed-form expression for the most general case, it might not be the only way of computing $E[L_h]$ numerically. Indeed it might be more efficient to decompose the original Hamiltonian path into several non-overlapping sections, each of them being a path with a starting black node, a finishing black node, and only white nodes in the middle; then one can simply apply Theorem 2.4 (by taking m=2) to each of these paths and then add their respective expected lengths to obtain the desired answer. For example let us consider $h=(i_1,i_2,...,i_{10})$ and assume that besides $i_1$ and $i_{10}$ nodes $i_3$, $i_6$, and $i_7$ are black; then either we compute $E[L_h]$ by using Theorem 2.4 once with n=5 m=5 or we can compute $E[L_h]$ by splitting h into four paths: $h_1=(i_1, i_2,i_3)$ $h_2=(i_3,...,i_6)$, $h_3=(i_6,i_7)$, and $h_4=(i_7,..,i_{10})$ and then applying Theorem 2.4 to each of them with m=2 giving:

$$E[L_h] = \sum_{i=1}^{4} E[L_{h_i}]$$

This method may be more efficient if m is relatively large, i.e. when the proportion of black nodes is not negligible. Formally the first method requires an order $O(n^2m)$ steps and the second requires $O(n^2+m)$ steps.

For the case of a tour t in G with $|N_1| = m$, $|N_2| = n$ the same idea applies: either we use Theorem 2.3 once, or Theorem 2.4 m times.

2.3.4 General Probabilistic Assumptions:

The combinatorial properties of the problems that led to the various lemmas in sections 2.2 and 2.3 are still valid if one chooses a general (instead of a binomial) pmf for W, the number of present nodes, provided that these nodes are indistinguishable, in the sense that, given W takes on the value k, the k present nodes are chosen at random from the set of white nodes. Given this condition, Lemmas 2.1 to 2.4 remain valid for the corresponding cases and Theorems 2.1 to 2.4 can be readily modified by simply leaving Pr(W=k) in general form. For example, Theorem 2.1 would result in the following:

$$E[L_t] = \sum_{k=0}^{n-2} \left[\left(1/\binom{n}{k}\right) \sum_{r=0}^{k} \binom{n-2-r}{k-r} L_t^{(r)}\right]\left[\Pr(W=n-k)\right]$$

$$= \sum_{r=0}^{n-2} L_t^{(r)} \left[\sum_{k=r}^{n-2} \left(\binom{n-2-r}{k-r}/\binom{n}{k}\right) \Pr(W=n-k)\right]$$

This is still a closed-form expression giving the expected length of a tour t as a weighted sum of the $L_t^{(r)}$'s; the weights are now more complicated but can be calculated very rapidly once we have Pr(W=k). Theorems 2.2 to 2.4 lead to the same generalization.

The general pmf for W (together with the condition of indistinguishable nodes with respect to W) naturally includes cases where there are dependences among nodes; for example if we specify Pr(W=0) = 1/2, Pr(W=n) = 1/2 then if one white node has to be visited, this will be true for the n-1 others. Hence the generalization of Theorem 2.4 along this line provides a wide range of possibilities for modeling purposes.

On the other hand this does not include cases where general node-specific probabilities or correlations exist; one notable exception can be exhibited from the following observation: k nodes, each with a coverage probability p (independently of each other) and positioned at the exact same location are in fact equivalent to a single node with a coverage probability $p'=1-(1-p)^k$. Hence for special cases of node-specific probabilities (that is, each node j has to have a coverage probability $p_j$ of the form $1-(1-p)^{k_j}$ for a common p) one can use (by duplicating nodes) the previous theorems. For modeling purposes, the importance of the limitations imposed by these special node-specific probability cases is reduced by the fact that, very often, we have only rough estimates of the coverage probabilities. Most of the time we know simply the relative frequency of visiting these nodes (i.e., node i is visited more often than node j); we can then assume a coverage probability $p_i = 1-(1-p)^{k_i}$ for each node i of the graph (with $k_i > k_j$ if node i is visited more often than node j).

For the <u>general</u> node-specific probability cases, one can use an alternative method for efficiently computing the expected length of a given tour which is based on recursive relationships; this method, however, does not provide closed-form expressions for the expected length.

For the PHPP, with n white nodes and two black nodes, let (1, 2, 3, ..., n+2) be the Hamiltonian path h under consideration, and let nodes (1) and (n+2) be the black nodes between which the path is defined (assume the white nodes are indexed according to their order of appearance along the path h). Let E[k] be the expected length of the path (k, k+1, ..., n+2) for $k \varepsilon [1..n+1]$. Then the following recursive relationship holds:

- $E[n+1] = d(n+1,\ n+2)$

  (initial condition)

- $$E[k] = \sum_{j=1}^{n+1-k} p_{k+j}\Big[\prod_{i=1}^{j-1}\big(1-p_{k+i}\big)\Big]\big(d(k,k+j) + E[k+j]\big) + \Big[\prod_{i=1}^{n+1-k}\big(1-p_{k+i}\big)\Big]d(k,n+2) \qquad \forall\ k \in [1..n] \tag{2.8}$$

The reasoning for (2.8) is simple: starting from node k, the first node to be reached is either k+j for $j \in [1..n-1]$ (with probability $p_{k+j}(1-p_{k+1}) \ldots (1-p_{k+j-1})$) or node n+2; in the first case, after reaching node k+j, the remaining expected distance is simply E[k+j].

For a given k, E[k] can thus be obtained numerically in $O\big((n+1-k)^2\big)$ steps. Hence E[1] (the expected length of the path) is obtained in $O(n^2)$ steps.

Following the discussion of section 2.3.3, the method presented here (for the special case of a PHPP with n white nodes and two black nodes) is sufficient to compute the expected length of any Hamiltonian path for a graph with m > 2; it is also sufficient to compute the expected length of any tour provided the graph contains at least one black node. For graphs with no black node, one can use the previous method by proceeding as follows:

let t = (1,2,...n,1) be the tour under consideration, then $E[L_t]$ can be obtained by considering the following steps:

step 1: k=0

step 2: k=k+1 (2.9)

step 3:
- consider $t_k = (k,k+1,k+2,\ldots,n,k)$
- assume k is a black node (it is a black node with probability $p_k$)
- compute $E[L_{t_k}]$ using (2.8)

step 4: if $k = n-1$ stop

otherwise go the step 2.

Then
$$E[L_t] = p_1\, E\left[L_{t_1}\right] + p_2(1-p_2)\, E\left[L_{t_2}\right] + \ldots + p_{n-1}(1-p_1)\ldots(1-p_{n-2})\, E\left[L_{t_{n-1}}\right]$$

Since the computation of $E\left[L_{t_k}\right]$ requires $O\left((n-k)^2\right)$ steps, $E[L_t]$ requires $O(n^3)$ steps.

It should be noted that the recursive approaches ((2.8) and (2.9)) could have been applied to the cases discussed in sections 2.3.1 - 2.3.3 and 2.2 (where $p_j = p \;\forall\; j$); but the closed form expressions developed by using specific combinatorial features of those special cases are equally efficient computationally and are certainly much more powerful for analytical purposes.

## 2.4 Conclusion

In this chapter we presented expressions giving the expected length of any given tour in an efficient manner; for example Theorem 2.1 leads to a reduction from $O\left(n2^n\right)$ (derived in the introduction by considering complete enumeration) to $O(n^2)$ number of steps to calculate this expected length (in fact it is worthwhile mentioning that for a general distance matrix D, we cannot find a more efficient method since, to compute the expected length of a tour, one must consider every elements of D and there are $O(n^2)$ such elements). Not only our results reduce substantially the number of computations required to get the expected length of a given tour, they do so through the derivation of closed-form expressions that will prove useful, later on, for the development of combinatorial properties and algorithmic procedures for the PTSP. Using such expressions we demonstrated, through an example, that the optimal TSP tour can be a very poor solution for the corresponding PTSP (we will present a generalization of this example in Chapter 3 during the investigation of combinatorial properties of the PTSP).

# CHAPTER 3

# GENERAL COMBINATORIAL PROPERTIES OF THE PTSP

## 3.1 Introduction; Preliminaries

### 3.1.1 Content of Chapter 3

We have seen in Chapter 2 how to express efficiently the expected length (in the PTSP sense) of a given tour of a complete graph G containing n white nodes (i.e., nodes which are present only probabilistically) and m black nodes (i.e., nodes that are always present). The most general result obtained assumed a general probability mass function (p.m.f.) for W - the number of white nodes present - provided that white nodes are "treated equally" in the probabilistic sense (that is, given W takes on the value k, those k present white nodes are chosen at random among the original n white nodes). Chapter 3 can be considered as a companion to Chapter 2: based on the results of the previous chapter, we will derive general combinatorial properties of the PTSP and we will provide insights into the peculiarities and, sometimes, counterintuitive behavior of this problem.

Let us first give a general outline of the contents of Chapter 3 as well as of the basic underlying ideas contained in our results. First, in the rest of this introductory section, after discussing the case of a non-complete (but connected) graph, we present a "weight-form" representation of every closed-form expression of Chapter 2 by means of a single formula; this unified representation allows us to present all our results in the most general way possible (assuming a general p.m.f. for W); we conclude section 3.1 by giving some additional notation and conventions which are subsequently used throughout the statements and proofs of our results.

In section 3.2 we take a close and systematic look at the expressions derived in Chapter 2; to do so we derive properties of the two classes of elements forming $E[L_t]$, first properties of the $L_{m,t}^{(r)}$ under various conditions for the distance matrix D, then properties of the weights; those properties are then used to derive lower and upper bounds on the expected length of a tour $E[L_t]$.

In section 3.3 we come back to a more global view of the problem, and using, in part, results from section 3.2, we try to obtain some useful characterizations and/or exploitable properties of the problem structure; we first show that there are some manipulations of D that leave the problem unchanged (paralleling similar characterizations of the TSP); we then show that surprisingly the optimal PTSP tour in the Euclidean plane can intersect itself; we finally consider the behavior of the expected length of a given tour with respect to perturbations of the graph such as adding or deleting a node, switching the color of a node, etc.

Section 3.4 is then concerned in finding some relationships between the optimal TSP tour (through all the demand points) and optimal PTSP tour (the existence of relations between those tours can be expected, since the TSP is a special case of the PTSP in which we have only black nodes); we first show that for problems of small size the optimal TSP tour can be the optimal PTSP tour for any probability assumptions for W; we then derive, using results from section 3.2, worst-case bounds on the absolute difference between the expected length of the optimal TSP tour and the expected length of the optimal PTSP tour; we discuss the tightness of our bounds and finally provide an analysis of the limiting behavior (as $n \to \infty$) of our "star-shaped" construction (introduced in Chapter 2 for n=24); this section contains also a discussion on some very peculiar aspects of the PTSP problem.

In section 3.5 we merely point out that all our previous results are also valid for the PHPP problem as originally defined in Chapter 2 (a Hamiltonian path between two black nodes); we then present a useful result based on the variation of the PHPP in which we do not require the nodes between which the Hamiltonian path is to be found to be black. The concluding section reviews the main results of this chapter and indicates where in the subsequent material they play an important role.

It should be pointed out that section 3.2 is extremely "technical" and that, despite its crucial importance for the development of the subsequent sections, might not contain (depending on the reader's taste) results as interesting as those obtained in section 3.3 and 3.4.

### 3.1.2 Preliminaries:

A. Case of a non-complete graph G:

The results presented in Chapter 2 all assumed that the given graph G was complete; in fact, by definition of the quantities $L_{m,t}^{(r)}$ for a given tour t any distance d(i,j) is used at least once in the computation of $E[L_t]$ - the expected length (in the PTSP sense) of a given tour t - and thus it is necessary to require D to have a finite value in each of its elements (except maybe the diagonal elements d(i,i) that do not need to be specified by definition of the problem); indeed if any of the elements of D is + ∞ then $E[L_t]$ will automatically be + ∞.

Let us then assume that we are given a non-complete graph G, together with a tour and are asked to compute its expected length (see Appendix B for more details on the discussion of non-complete graphs). A natural way of doing this would be exactly as before by simple assuming that if the arc $(i_j, i_{j+r+1})$ does not exist we simply go from $i_j$ to $i_{j+r+1}$

using intermediary nodes that exist along the tour t (for example taking the shortest path along the tour if several possibilities are offered). One immediate inconvenience of this approach is that it is tour-dependent; indeed, for each tour t we would need to find $d(i_j, i_{j+r+1})$; moreover, it introduces asymmetry even if the original graph was symmetric since there is no reason why $d(i_j, i_{j+r+1}) = d(i_{j+r+1}, i_j)$.

All these problems lead to the following alternative approach that we will assume from now on: if one is given a non-complete graph G, i.e., a matrix D with some empty entries, then for each non-existent arc (i,j), we create an artificial arc (i,j) whose length will be the shortest distance (using arcs from the original graph G) from i to j; in other words, we transform the matrix D into a matrix $D'$ which will correspond to a complete graph $G'$; we then define the expected length of the tour t on the graph G to be the expected length of t in $G'$ as it was previously defined for a complete graph. Note that the transformation $D \to D'$ is tour-independent and preserves some properties of D (i.e., if D is symmetric, then $D'$ is too).

Keeping this transformation in mind, we can now concentrate solely on complete graphs; let us now present a unified approach for the presentation of our results in subsequent sections.

B. Weight form notation:

Given a graph $G = (N_1 \cup N_2, A, D)$ with n white nodes (the set $N_2$) and m black node (the set $N_1$), given a general p.m.f. for W - the number of white nodes present -, and given a tour t of G, we have expressed the expected length (in the PTSP sense) of t, $E[L_t]$, as:

$$\text{for } m=0, \ E[L_t] = \sum_{r=0}^{n-2} \alpha_r \, L_{0,t}^{(r)}$$

$$\text{for } m=1,\ E[L_t] = \sum_{r=0}^{n-2} \alpha_r\, L_{1,t}^{(r)} + \frac{1}{n}\Pr(W=1)L_{1,t}^{(n-1)}$$

$$\text{for } m\geq 2,\ E[L_t] = \sum_{r=0}^{n-2} \alpha_r\, L_{m,t}^{(r)} + \frac{1}{n}\Pr(W=1)L_{m,t}^{(n-1)} + \Pr(W=0)L_{m,t}^{(n)}$$

where:

$$\alpha_r = \sum_{k=r}^{n-2} \left(\binom{n-2-r}{k-r}\Big/\binom{n}{k}\right) \Pr(W=n-k) \ \forall\ r \in [0..n-2]$$

(see section 2.3.4)

$L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}(i_j, i_{j+1+r})$ (where the $d_{m,t}(.,.)$ are obtained from $d(.,.)$ according to rules given in Chapter 2; the reader is reminded that graphical illustrations of the $L_{m,t}^{(r)}$'s for the different cases are provided in Appendix A).

We will express $E[L_t]$ by means of a single formula valid for every case (i.e., m=0,1 or m$\geq$2) as follows:

$$E[L_t] = \sum_{r=0}^{n} \alpha_r\, L_{m,t}^{(r)}$$

with the following specific rules:

- $L_{0,t}^{(n-1)} = L_{0,t}^{(n)} = 0$ ; $L_{1,t}^{(n)} = 0$ .
- $\alpha_{n-1} = \frac{1}{n}\Pr(W=1)$ ; $\alpha_n = \Pr(W=0)$ .

Note:

(1) When W has the originally used binomial p.m.f. (corresponding to the cases for which each white point is present with a probability p, independently of each other) $\alpha_r$ takes the following more familiar form:

$$\alpha_r = p^2(1-p)^r \quad \forall\ r \in [0..n-2]$$

$$\alpha_{n-1} = p(1-p)^{n-1}$$

$$\alpha_n = (1-p)^n \quad .$$

(2) Given a general p.m.f. for W, it is easy to see that one can recover individual probabilities as follows:

$$\text{Prob}\{j \text{ \underline{specific} nodes } i_1, i_2, .., i_j \text{ are present}\} = \sum_{k=j}^{n} \left(\binom{k}{j} / \binom{n}{k}\right) \Pr(W=k) \tag{3.0}$$

C. Specific convention and usual definitions:

We end these preliminaries by defining a terminology that will be adopted throughout this chapter (everything else is assumed to be as in Chapter 2).

(1) "Binomial case" corresponds to the case where $\Pr(W=k)=\binom{n}{k}p^k(1-p)^{n-k}$ (that is, each white node is present with a probability p, independent of the other nodes).

(2) A metric is used in the traditional sense, that is, a function d, mapping NxN into R such that:

$d(i,j) = 0$ if and only if $i=j$

$d(i,j) = d(j,i) \geq 0 \;\forall\; i,j$ (symmetry)

$d(i,j) \leq d(i,k) + d(k,j) \;\forall\; i,j,k$ (triangular inequality).

(3) We will always assume that $n \geq 1$; $n+m \geq 3$.

Note that for n=0 the PTSP reduces to the traditional TSP and for $m+n \leq 2$ we have a unique tour; both cases are obviously of little interest for our purpose.

We are now ready to present our results.

In all cases a complete graph G is assumed and we are given a tour t of G; two specific tours will play an important role in Section 3.4.

$t_1$ : optimal TSP tour

$t_p$ : optimal PTSP tour

## 3.2 Analysis of the Closed Form Expressions of Chapter 2

In this section we will be mainly concerned with analysing in detail the component elements of $E[L_t]$ (for a given tour t of a given graph G

$E[L_t] = \sum_{r=0}^{n} \alpha_r L_{m,t}^{(r)}$ ) namely $L_{m,t}^{(r)}$ and $\alpha_r$.

### 3.2.1 Properties of the $L_{m,t}^{(r)}$

We will proceed from properties valid under a general distance matrix D to properties valid under some restrictions on D.

A. The distance matrix D is general:

Fact 3.1: (Number of elements of D in $L_{m,t}^{(r)}$ )

Given a complete graph G with n white nodes, m black nodes, and given a tour t, then:

(i) $L_{m,t}^{(r)}$ is composed of the addition of n+(r+1)m elements of the matrix D and this is true:

- $\forall\ r \in [0..n-2]$ if $m = 0$
- $\forall\ r \in [0..n-1]$ if $m \geq 1$

(ii) $L_{m,t}^{(n)}$ is composed of the sum of m elements $\forall\ m \geq 2$

(iii) for m = 0,1 every element of $L_{m,t}^{(r)}$ is distinct

Proof: by definition of the $L_{m,t}^{(r)}$; for example let us give a formal argument for (i) and $m \geq 1$ (the other cases are even simmpler to prove); we have

$$L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}\,(i_j, i_{j+r+1})$$

where $d_{m,t}(i_j,i_{j+r+1})$ is by definition the sum of $k_j+1$ elements $k_j$ being the number of black nodes among $\{i_{j+1},\ldots,i_{j+r}\}$ hence the total number of elements involved in the computation of $L_{m,t}^{(r)}$ is

$$\sum_{j=1}^{n+m} (k_j+1) = n+m + \sum_{j=1}^{n+m} k_j$$

and since each node of the tour t appears in r different sets $\{i_{j+1},\ldots,i_{j+r}\}$ and as $k_j$ counts only the black nodes we have

$$\sum_{j=1}^{n+m} k_j = mr \qquad \text{Q.E.D.}$$

Fact 3.2: (Summation $\sum_r L_{m,t}^{(r)}$ )

Under the conditions of Fact 3.1 we have:

(i) $\sum_{r=0}^{n} L_{m,t}^{(r)}$ = constant independent of t if m = 0.

(ii) $\sum_{r=0}^{n} L_{m,t}^{(r)}$ is in general tour-dependent if $m \geq 1$, but $L_{1,t}^{(n-1)}$ and $L_{2,t}^{(n)}$ are tour-independent

Proof: Again this is obvious by definition of the $L_{m,t}^{(r)}$:

for (i): $\sum_{r=0}^{n} L_{0,t}^{(r)} = \sum_{r=0}^{n-2} \sum_{j=1}^{n} d(i_j,i_{j+r+1}) = \sum_{k\neq\ell}\sum d(k,\ell)$

for (ii) one can argue as follows: take for simplicity m=1 and assume that $i_1$ is the black node; then, by definition of the $L_{1,t}^{(r)}$'s, $d(i_1,i_2)$ will appear in all of them, $d(i_1,i_3)$ in all but $L_{1,t}^{(o)}$, etc. so that $\sum_r L_{m,t}^{(r)}$ will be dependent on the order of the sequence of nodes (in other words on the tour t). The cases of $m \geq 2$ can be proved identically. Finally to show that $L_{1,t}^{(n-1)}$ and $L_{2,t}^{(n)}$ are tour-independent it suffices to remark that:

$$L_{1,t}^{(n-1)} = \sum_{j=1}^{n} d_{1,t}(i_j, i_{j+n}) = \sum_{k\neq 1} \big(d(i_1,i_k) + d(i_k,i_1)\big)$$

where $i_1$ is the black node, and that $L_{2,t}^{(n)}$ is a tour through 2 points and is then unique.

Q.E.D.

Fact 3.3: (Network characterization of the $L_{m,t}^{(r)}$ )

(i) for m = 0, $L_{0,t}^{(r)}$ consists of $g_r$ distinct (i.e., without common nodes) subtours, each of them containing $\frac{n}{g_r}$ points, where $g_r$ = G.C.D. (n,r+1).

(ii) for m = 1, $L_{1,t}^{(r)}$ consists of r+1 subtours having the black node as a common node.

(iii) in general for any $m \geqslant 0$, $L_{m,t}^{(r)}$ "visits" each white node once and each black node (r+1) times.

Proof: The validity of (i) is best seen by writing $L_{0,t}^{(r)}$ as a function of the cyclic permutation $\Pi$ (determined uniquely by the tour t).

$$L_{0,t}^{(r)} = \sum_{j=1}^{n} d(i, \Pi^{r+1}(i)).$$

Then it is a classical result (see, for example Berge [1971]) that if $\Pi$ is a cyclic permutation on n objects (that is, with one cycle of length n) then $\Pi^{r+1}$ is a permutation with $g_r$ cycles of length $\frac{n}{g_r}$ where $g_r$ = G.C.D.(n,r+1).

The proof of (ii) is obtained directly from the definition of $L_{1,t}^{(r)}$ and an argument similar to the one used in Fact 3.1.

(iii) is also straightforward to check and in the case of the black nodes the argument parallels again the one used in Fact. 3.1.

Q.E.D.

This fact will be of use now to obtain a lower bound on $L_{m,t}^{(r)}$.

<u>Lemma 3.1</u>: (Lower bound on $L_{m,t}^{(r)}$ )

Given a graph G, an optimal TSP tour $t_1$ of G, and any other tour t of G, we have:

(i) <u>for m=0</u> $L_{0,t}^{(r)} \geq L_{0,t_1}^{(o)}$ $\forall$ r such that G.C.D.(n,r+1)=1

$$L_{0,t}^{(r)} \geq \frac{1}{2} \sum_{j=1}^{n} \left(d_{in}^{(1)}(j)+d_{out}^{(1)}(j)\right) \equiv B_{SL} \text{ otherwise}$$

(ii) <u>for m≥1</u>

$$L_{m,t}^{(r)} \geq \frac{1}{2} \sum_{j \text{ white}} \left(d_{in}^{(1)}(j)+d_{out}^{(1)}(j)\right) + \frac{1}{2} \sum_{j \text{ black}} \left(\sum_{k=1}^{r+1} \left(d_{in}^{(k)}(j) + d_{out}^{(k)}(j)\right)\right)$$

for r $\varepsilon$ [0..n-1]

where $d_{in}^{(k)}(j)$ $\left(d_{out}^{(k)}(j)\right)$ is the length of the $k^{th}$ shortest arc coming into (out of ) node j.

<u>Proof</u>: this lemma is a direct consequence of Fact 3.3. Indeed for (i) when G.C.D.(n,r+1)=1, then $L_{0,t}^{(r)}$ is the length of a single tour and thus cannot be smaller than the length of the optimal TSP tour. For all the other cases the bound is a direct consequence of part (iii) in Fact 3.3 (which states that there is a single arc going in and coming out of any white node and (r+1) arcs in and out of black node). One can note that for case (i) if n is prime then $L_{0,t}^{(r)} \geq L_{0,t_1}^{(o)}$ for all r.

Q.E.D.

Facts 3.1 - 3.3 and Lemma 3.1 are properties of $L_{m,t}^{(r)}$ without any assumptions on the distance matrix D. We now turn to additional properties that can be obtained by restricting D to some specific cases: we consider first D to be symmetric (i.e. each element of D is such that $d(i,j) = d(j,i)$) and then assume that the triangular inequality holds.

B. <u>The matrix D is symmetric</u>:

By making this additional assumption, we obtain the following fact:

<u>Fact 3.4</u>: (Symmetric D)

Given a graph G with a symmetric distance matrix D and a tour t:

(i) $L_{0,t}^{(n-2-r)} = L_{0,t}^{(r)} \ \forall\, r \in [0..n-2]$ and this is (in general) not true for $L_{m,t}^{(r)}$ when $m > 1$.

(ii) $\sum_r L_{1,t}^{(r)}$, $L_{2,t}^{(n-1)}$, and $L_{3,t}^{(n)}$ all become tour-independent.

<u>Proof</u>:

- The validity of (i) can be seen by using the permutation definition of $L_{0,t}^{(r)}$; indeed we have

$$L_{0,t}^{(n-2-r)} = \sum_{i=1}^{n} d\left(i, \Pi^{n-2-r}(i)\right) \equiv \sum_{i=1}^{n} d\left(\Pi^{r+1}(i), \Pi^{r+1}\left(\Pi^{n-2-r}(i)\right)\right)$$

$$= \sum_{i=1}^{n} d\left(\Pi^{r+1}(i), \Pi^{n}(i)\right) = \sum_{i=1}^{n} d\left(\Pi^{r+1}(i), i\right) .$$

The fact that this is not true for $m > 1$ is obvious from the network characterization of $L_{m,t}^{(r)}$ given in Fact 3.3.

- Let us now consider part (ii). When D is symmetric there is a single tour through 3 points and this implies that $L_{3,t}^{(n)}$ is tour-independent; this also implies the same for $L_{2,t}^{(n-1)}$ once we note that $L_{2,t}^{(n-1)}$ can be written as:

$$L_{2,t}^{(n-1)} = \sum_{k \neq 1,2} \left( d(i_k, i_1) + d(i_1, i_2) + d(i_2, i_k) \right)$$

where $i_1$ and $i_2$ are assumed to be the two black nodes. To show that $\sum_r L_{1,t}^{(r)}$ becomes tour-independent under the assumption of a symmetric distance matrix one simply has to note that now (assuming $i_1$ to still be the black node) elements of the form $d(i_1, i_{k+2})$ will appear n+1 times in total summation: indeed they will appear once in each of $L_{1,t}^{(r)}$ (for $r \in [k..n-1]$) under the form $d(i_1, i_{k+2})$ and once in each of $L_{1,t}^{(r)}$ ($r \in [0..k-1]$ and $r = n-1$) under the form $d(i_{k+2}, i_1)$; this is true $\forall\, k \in [0..n-1]$ which implies that

$$\sum_{r=0}^{n-1} L_{1,t}^{(r)} = (n+1) \sum_{k \neq 1} d(i_1, i_k) + \frac{1}{2} \sum_{\substack{k \neq \ell \\ k \neq i_1}} \sum_{\ell \neq i_1} d(k, \ell)$$

Q.E.D.

C. <u>The matrix D satisfies the triangular inequality</u>:

Under this additional restriction, we can state the following result:

<u>Lemma 3.2</u>: (Another lower bound on $L_{m,t}^{(r)}$ for $m \geq 1$)

Given a graph G with a distance matrix D that satisfies the triangular inequality and given a tour t of G, we have:

$$L_{m,t}^{(r)} \geq L_{m,t_1}^{(o)} \quad \forall\, m \geq 1 \quad \forall\, r \in [0..n-1]$$

where $t_1$ is an optimal TSP tour of G)

Proof: Let us consider first the case where m=1.

From Fact 3.3 we know that $L_{1,t}^{(r)}$ corresponds to the length of r+1 subtours having the black node in common; one can transform this set of r+1 subtours into a single tour whose length will not be greater than $L_{1,t}^{(r)}$ under the triangular inequality assumption; the length of this tour is in turn not smaller than the length of the optimal PTSP tour; this establishes Lemma 3.2 for m=1.

For m $\geqslant$ 2 one can note that by the triangular inequality

$L_{m,t}^{(r)} \geqslant L_{1,t}^{(r)}$ (by keeping only one black node and turning the others into white); this is true simply because $d_{m,t}(i_j,i_{j+r+1}) \geqslant d_{1,t}(i_j,i_{j+r+1})$ $\forall$ j. We can then apply Lemma 3.2 on $L_{1,t}^{(r)}$ to obtain:

$$L_{m,t}^{(r)} \geqslant L_{1,t}^{(r)} \geqslant L_{1,t_1}^{(o)}$$

and since $L_{m,t_1}^{(o)} \equiv L_{1,t_1}^{(o)}$ Lemma 3.2 is proved.

Q.E.D.

In addition to these lower bounds, the triangular inequality assumption allows us to derive an upper bound on $L_{m,t}^{(r)}$:

Lemma 3.3: (Upper bound on $L_{m,t}^{(r)}$ )

Given a graph G with n white nodes and m black nodes and given any tour t of G the following is true:

$$L_{m,t}^{(r)} \leqslant L_{m,t}^{(r_1)} + L_{m,t}^{(r_2)} \qquad \forall\, r \in [1..n-1]$$

$$\forall\, r_1, r_2 : r_1 + r_2 = r-1$$

As a consequence $L_{m,t}^{(r)} \leqslant (r+1)\, L_{m,t}^{(o)} \quad \forall\, r \in [1..n-1]$

Proof: We have $L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}(i_j,i_{j+r+1})$; now consider $d_{m,t}(i_j,i_{j+r+1})$

and $i_{j+r_1+1}$ with $0 \leq r_1 \leq r-1$:

If $i_{j+r_1+1}$ is a black node then

$$d_{m,t}(i_j, i_{j+r+1}) = d_{m,t}(i_j, i_{j+r_1+1}) + d_{m,t}(i_{j+r_1+1}, i_{j+r+1})$$

If $i_{j+r_1+1}$ is a white node, let $k_1$ and $k_2$ be the black nodes immediately preceeding and following $i_{j+r_1+1}$ along the sequence $(i_j, \ldots, i_{j+r+1})$.

By the triangular inequality we have

$$d(k_1, k_2) \leq d(k_1, i_{j+r_1+1}) + d(i_{j+r_1+1}, k_2)$$

which implies in turn that:

$$d_{m,t}(i_j, i_{j+r+1}) \leq d_{m,t}(i_j, i_{j+r_1+1}) + d_{m,t}(i_{j+r_1+1}, i_{j+r+1}) \quad .$$

With every case considered we thus obtain:

$$L_{m,t}^{(r)} = \sum_{j=1}^{n+m} d_{m,t}(i_j, i_{j+r+1}) \leq \sum_{j=1}^{n+m} d_{m,t}(i_j, i_{j+r_1+1}) + $$

$$+ \sum_{j=1}^{n+m} d_{m,t}(i_{j+r_1+1}, i_{j+r+1})$$

but since

$$\sum_{j=1}^{n+m} d_{m,t}(i_{j+r_1+1}, i_{j+r+1}) = \sum_{j=1}^{n+m} d_{m,t}(i_j, i_{j+r+1-r_1-1})$$

we finally obtain

$$L_{m,t}^{(r)} \leq L_{m,t}^{(r_1)} + L_{m,t}^{(r-r_1-1)}$$

Q.E.D.

Note: In the remainder of this chapter (and Thesis) we will always assume that D satisfies the triangular inequality so that all our previous results hold except Fact 3.4 for which D is required to be symmetric. The triangular inequality is most often satisfied in practice and one of the important cases where this assumption is automatically true is when d(i,j) corresponds to the length of the shortest path from i to j in the original given graph G.

Let us now turn our attention to properties of $\alpha_r$, the other element involved in the closed-form expressions of Chapter 2.

### 3.2.2 Properties of the Weights

As indicated in section 3.1 we will present our results assuming a general probability mass function for W - the random variable representing the total number of present white nodes. Most of the time we will also indicate the form taken by our general results for the binomial case. (These conventions are valid for the rest of this chapter).

The first property is concerned with the relative values of the weights:

Fact 3.5: (Relative values of the weights)

Given any choice for the probability mass function of W we have:

$$\alpha_r > \alpha_{r'} \qquad \text{for } 0 \leq r < r' \leq n-2$$

Proof: Before presenting the general proof, one can easily see that Fact 3.5 holds for the binomial case for which $\alpha_r = p^2(1-p)^r$ since then $\alpha_{r'} = (1-p)^{r'-r}\,\alpha_r$ and $1-p<1$.

In general, $\alpha_r = \sum_{k=r}^{n-2} \left(\binom{n-2-r}{k-r}\Big/\binom{n}{k}\right) \Pr(W=n-k)$ so that:

$$\alpha_r - \alpha_{r+1} = \Pr(W=n-r)/\binom{n}{r} + \sum_{k=r+1}^{n-2} \frac{\binom{n-2-r}{k-r} - \binom{n-3-r}{k-r-1}}{\binom{n}{k}} \Pr(W=n-k) \qquad (3.1)$$

and since $\binom{n-2-r}{k-r} - \binom{n-3-r}{k-r-1} = \binom{n-3-r}{k-r}$ , (3.1) becomes:

$$\alpha_r - \alpha_{r+1} = \sum_{k=r}^{n-3} \left(\binom{n-3-r}{k-r}/\binom{n}{k}\right) \Pr(W=n-k) \geq 0 \quad .$$

This in turn implies Fact 3.5.

Q.E.D.

One can also note that by applying the same logical step to $\alpha_r - \alpha_{r+1} \equiv \beta_r$ we obtain:

$$\beta_r - \beta_{r+1} = \sum_{k=r}^{n-4} \left(\binom{n-4-r}{k-r}/\binom{n}{k}\right) \Pr(W=n-k) \geq 0 \quad .$$

(one can then apply that step to $\gamma_r = \beta_r - \beta_{r+1}$ and so on..)

Our second result is concerned with summations of the $\alpha_r$'s.

<u>Fact 3.6</u>: (Summation of the $\alpha_r$'s)

Given any choice for the probability mass function of W we have:

$\forall\, n \geq 1$

(i) $\displaystyle\sum_{r=0}^{n-1} \alpha_r = \frac{E[W]}{n}$

(ii) $\displaystyle\sum_{r=0}^{n-1} (r+1)\alpha_r = 1 - \Pr(W=0)$

Proof: First note that for n=1, Fact 3.6 is obvious; indeed (i) merely reflects that $\alpha_o = Pr(W=1) = E[W]$ and (ii) the trivial fact that $Pr(W=1) = 1 - Pr(W=0)$. Let us then concentrate on the non-trivial case of $n \geq 2$:

1. to show (i) note that by definition of $\alpha_r$

$$\sum_{r=0}^{n-2} \alpha_r = \sum_{r=0}^{n-2} \left[ \sum_{k=r}^{n-2} \left( \binom{n-2-r}{k-r} / \binom{n}{k} \right) Pr(W=n-k) \right]$$

which becomes, after reversing the summation,

$$\sum_{r=0}^{n-2} \alpha_r = \sum_{k=0}^{n-2} \left( \left( Pr(W=n-k) / \binom{n}{k} \right) \sum_{r=0}^{k} \binom{n-2-r}{k-r} \right) \tag{3.2}$$

Also, $$\sum_{r=0}^{k} \binom{n-2-r}{k-r} = \binom{n-1}{k} \tag{3.3}$$

Hence (3.2) and (3.3) lead to:

$$\sum_{r=0}^{n-2} \alpha_r = \sum_{k=0}^{n-2} \binom{n-1}{k} / \binom{n}{k} \, Pr(W=n-k) \quad . \tag{3.4}$$

Now, by replacing n-k by u in (3.4) we obtain:

$$\sum_{r=0}^{n-2} \alpha_r = \sum_{u=2}^{n} \frac{u}{n} Pr(W=u) = \frac{1}{n} \left[ E[W] - Pr(W=1) \right]$$

and since $\alpha_{n-1} = \dfrac{Pr(W=1)}{n}$ (i) of Fact 3.6 is proved.

2. It remains to show (ii):

We have

$$\sum_{r=0}^{n-2} (r+1)\alpha_r = \sum_{r=0}^{n-2} \sum_{k=r}^{n-2} \left(\binom{n-2-r}{k-r} \Big/ \binom{n}{k}\right) (r+1) \Pr(W=n-k)$$

$$= \sum_{k=0}^{n-2} \left(\Pr(W=n-k) \Big/ \binom{n}{k}\right) \sum_{r=0}^{k} (r+1) \binom{n-2-r}{k-r} \quad . \tag{3.5}$$

Again it is easy to show that

$$\sum_{r=0}^{n-2} (r+1) \binom{n-2-r}{k-r} = \binom{n}{k} \tag{3.6}$$

so that (3.5) and (3.6) give the following:

$$\sum_{r=0}^{n-2} (r+1)\alpha_r = \sum_{k=0}^{n-2} \Pr(W=n-k) = 1-\Pr(W=0) - \Pr(W=1) \tag{3.7}$$

Now, since $\alpha_{n-1} = \Pr(W=1)/n$, (ii) is proved.

Q.E.D.

We now have all the facts we need to prove the following result which gives lower and upper bounds for the expected length of any given tour t of a graph G.

### 3.2.3 Bounds on $E[L_t]$

Lemma 3.4: (Upper and Lower Bounds on $E[L_t]$)

For any tour t through n white nodes and m black nodes, we have (with $t_1$ being an optimal TSP tour of the graph G):

(i) for m = 0

$$E[L_t] \leq L_{0,t}^{(o)} \left[1 - \Pr(W=0) - \Pr(W=1)\right]$$

$$E[L_t] \geq L_{0,t_1}^{(o)} \left[\frac{E[W]-\Pr(W=1)}{n}\right] \text{ if n prime}$$

$$E[L_t] \geq \alpha_o\, L_{0,t_1}^{(o)} \left[\frac{E[W]-\Pr(W=1)}{n} - \alpha_o\right] B_{SL} \text{ otherwise}$$

(where $B_{SL} = \frac{1}{2} \sum_{j=1}^{n} \left(d_{in}^{(1)}(j) + d_{out}^{(1)}(j)\right)$ as defined in Lemma 3.1).

(ii) for m = 1

$$E[L_t] \leq L_{1,t}^{(o)} \left[1 - Pr(W=0)\right]$$

$$E[L_t] \geq L_{1,t_1}^{(o)} \left[\frac{E[W]}{n}\right] .$$

(iii) for $m \geq 2$

$$E[L_t] \leq L_{m,t}^{(o)}$$

$$E[L_t] \geq L_{m,t_1}^{(o)} \left[\frac{E[W]}{n}\right] .$$

Proof: (i) follows from Lemma 3.1, Lemma 3.3, and Fact 3.6.

(ii) and (iii) follow from Lemma 3.2, Lemma 3.3 and Fact 3.6.

Q.E.D.

Before concluding section 3.2, one can note that the lower bound given in Lemma 3.4 for the case m=0, n not prime can be improved by considering a lower bound valid for $\sum_r \alpha_r L_{0,t}^{(r)}$ as a whole (and not for $L_{0,t}^{(r)}$ individually). This is made possible by Fact 3.5.

Lemma 3.5: (Improved lower bound on $E[L_t]$ for m=0 n not prime)

Define $B_{SL}^{(r)} = \frac{1}{2} \sum_{j=1}^{n} \left(d_{in}^{(r)}(j) + d_{out}^{(r)}(j)\right)$

where $d_{in}^{(r)}(j)$ , $d_{out}^{(r)}(j)$ have the same definition as in Lemma 3.1.

Then we have:

$$E[L_t] \geq \alpha_o L^{(o)}_{1,t_1} + \sum_{j=1}^{n-2} \alpha_r B^{(r)}_{SL}.$$

Proof: A consequence of Fact 3.1, Fact 3.3, and Fact 3.5.

Q.E.D.

This concludes this somewhat overly technical section; it should, however, be stressed that the results developed in this section will be often used in later sections of this chapter as well as later on as, for example, in Chapter 5.

Let us now take a slightly more "macroscopic" view in order to obtain some useful characteristic of the combinatorial problems under investigation.

## 3.3 Perturbation of the Graph and Exploitable Properties of the Problem Structure

In this section we attempt to obtain some useful characterization of the PTSP problem, such as specific properties of the problem structure; in doing so, we sometimes compare the PTSP to the well-understood TSP. Our first result is concerned with some "allowable" manipulations of the distance matrix D.

Lemma 3.6:

Given a graph G and given a tour t of expected length $E[L_t]$, if we subtract a constant b from every element of the row i of the distance matrix D, the tour to has a new expected length $E'[L_t]$ which is obtained from $E[L_t]$ as follows:

(i) $m = 0$ $\quad E'[L_t] = E[L_t] - \frac{b}{n}\left[E[W]-Pr(W=1)\right]$

(ii) $m = 1$

$$E'[L_t] = E[L_t] - \frac{b}{n} E[W] \quad \text{if row } i \text{ corresponds to a white node}$$

$$E'[L_t] = E[L_t] - b\left[1 - Pr(W=0)\right] \quad \text{otherwise.}$$

(iii) $m \geq 2$

$$E'[L_t] = E[L_t] - \frac{b}{n} E[W] \quad \text{for a white node}$$

$$E'[L_t] = E[L_t] - b \quad \text{otherwise.}$$

Proof: There are more than one ways of proving this lemma; the simplest one for a general p.m.f. for W is to use results from the previous section. Let us first consider case (i) with no black node:

$$\text{If } m = 0 \text{ then } E[L_t] = \sum_{r=0}^{n-2} \alpha_r L_{0,t}^{(r)} .$$

From Fact 3.3 each $L_{0,t}^{(r)}$ $r \in [0..n-2]$ "visits" node i exactly once so that there is a single arc going out of node i that is included in each $L_{0,t}^{(r)}$ ; since we subtract b from every such arc we have

$$E'[L_t] = E[L_t] - \sum_{r=0}^{n-2} \alpha_r b$$

and this together with Fact 3.6 prove the desired result.

For cases (ii) and (iii) the arguments parallel the one given above.

Q.E.D.

Note: (1) By subtracting a constant b from a row i of a symmetric matrix D, we end up with an asymmetric matrix D'; to avoid this possibly annoying fact, one simply has to subtract the same constant b from the column i of D and then Lemma 3.6 holds by replacing b by 2b. On the other hand, one should be aware that the transformations of Lemma 3.6 do not necessarily preserve the triangular inequality.

(2) An important consequence of Lemma 3.6 is that, by subtracting constants from any row or column of D (for any number of times) we change the expected length of any tour t by the same value (independent of the tour t) and this implies in turn that solving the PTSP in the original graph G is equivalent to solving it in the transformed graph G' (corresponding to the transformed matrix $D'$). This has the advantage of being able to work with matrices with zero entries, a sometimes "computationally" useful factor.

The following result will give an interesting and surprising characterization of the PTSP tour.

Lemma 3.7:

The optimal PTSP tour may intersect itself in the plane with the Euclidean metric.

Proof: We are going to present a counterexample; the full set of calculations is provided in Appendix C. This example corresponds to five points on the plane, two of them being black (i.e., always present), the other three being present only with a fixed probability p, independently of each other; their relative position is shown in Figure 3.1; in Figure 3.2 we present the optimal PTSP tours for $p > 0.75$ in (a) and for $p < 0.25$ in (b) (the proof that they are the optimal PTSP tours is given in Appendix C). This example has been constructed for m=2; to see that it also proves Lemma 3.7 for m=1 (or m=0) it suffices to replace one (or two) of the black nodes by a sufficiently large number of superimposed white nodes so that the probability of not visiting anyone of them is arbitrarily close to zero.

Q.E.D.

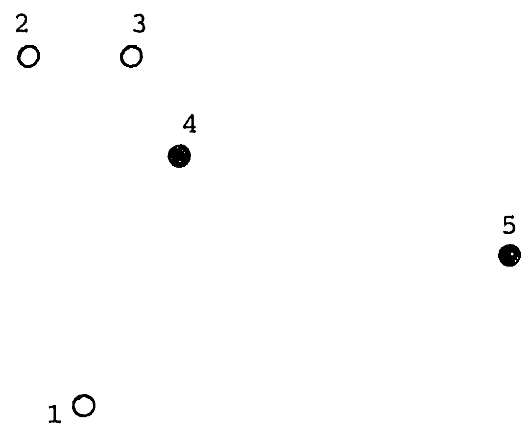
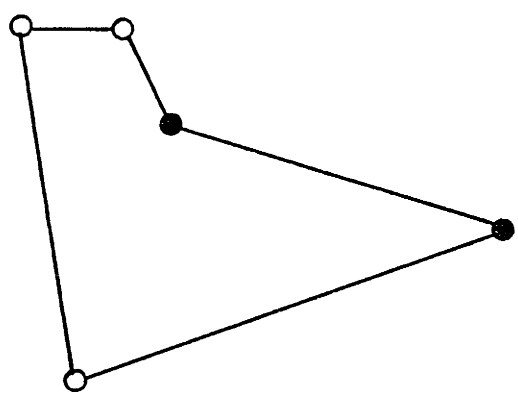
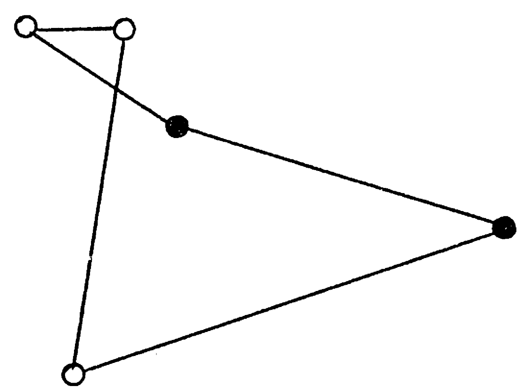



Figure 3.1: the set of five points



3.2a: Optimal tour (tour 1)
when p > 0.75



3.2b: Optimal tour (tour 2)
when p < 0.25

Figure 3.2: the optimal PTSP tours

Note: This result, of course, does not have as positive a consequence as the previous lemma (Lemma 3.6); indeed, for the TSP in the plane it is easy to show that the optimal TSP tour does not intersect itself and this property has often been exploited for developing and improving heuristic algorithms to solve the TSP in the plane (see Larson and Odoni [1981] for a discussion on this subject). For our problem this property of the TSP is not true anymore and we cannot discard the tours that intersect themselves from the pool of potential optimal PTSP tours.

We shall conclude this section with a somewhat different concern. Our interest will be to explore what happens to the expected length of a given tour t of a graph G after performing small perturbations of this graph, such as adding or deleting a node, and switching the color of a node.

Our main findings are contained in the following lemma:

Lemma 3.8:

Let $G = (N_1 \cup N_2, A, D)$ be a given graph with m black nodes and n white nodes and let t be a given tour of expected length $E[L_t]$. Assume W is given by the "Binomial case". Then the following results are obtained:

(a) if one node i is deleted from G, the resulting tour t', obtained from t by simply removing i from the sequence, has an expected length $E[L_{t'}]$ which satisfies:

$$E[L_t] - 2d_i^* \leq E[L_{t'}] \leq E[L_t] \quad \text{if i is black}$$

$$E[L_t] - 2d_i^* p \leq E[L_{t'}] \leq E[L_t] \quad \text{if i is white}$$

(b) if one white node i of G becomes black then the new expected length of t, $E'[L_t]$, is such that:

$$E[L_t] \leq E'[L_t] \leq E[L_t] + 2d_i^*(1-p)$$

where $d_i^* = \max_{j \neq i} \{d(i,j);\ d(j,i)\}$

Proof: To prove (a) or (b) we have several possibilities. One can use the method of proof used in Lemma 3.6, but it seems easier to proceed differently (and this approach also has the advantage of presenting a method of proof very useful for problems containing probabilistic elements).

The idea is the following:

(1) Assume i to be a white node; then:

$$E[L_t] = E[L_t | i \text{ present}] \Pr\{i \text{ is present}\} + E[L_t | i \text{ absent}] \Pr\{i \text{ is absent}\} \quad . \qquad (3.8)$$

The main observation here is that:

$E[L_t | i \text{ present}]$ is nothing else than $E'[L_t]$ of part (b)

$E[L_t | i \text{ absent}]$ is nothing else than $E[L_{t'}]$ of part (a) .

Now it is easy to verify that:

$$E[L_t | i \text{ present}] \geq E[L_t | i \text{ is absent}] \qquad (3.9)$$

$$E[L_t | i \text{ present}] \leq E[L_t | i \text{ is absent}] + 2d_i^* \quad . \qquad (3.10)$$

(To see the validity of (3.10) it suffices to realize that $E[L_t | i \text{ present}]$ corresponds to the expected length of a tour through n-1 white nodes and m+1 black nodes and that $E[L_t | i \text{ absent}]$ corresponds to the same tour but with only n-1 white nodes and m black nodes; hence for each subsets of the n-1 white nodes the two tours will differ at most by $2d_i^*$, which implies (3.10)).

Now: (3.8) and (3.9) imply:

$$E[L_t] > E[L_t | i \text{ is absent}] \equiv E[L_{t'}]$$

$$E[L_t] < E[L_t | i \text{ is present}] \equiv E'[L_t] \quad .$$

Also, (3.8) and (3.10) imply:

$$E[L_t] < E[L_t | i \text{ is absent}] + 2d_i^*p \equiv E[L_{t'}] + 2d_i^*p$$

$$E[L_t] > E[L_t | i \text{ is present}] + 2d_i^*(1-p) \equiv E'[L_t] + 2d_i^*(1-p) \quad .$$

This takes care of part (a) (i white) and part (b).

(2) To verify the validity of (a) when i is black it suffices to note that under this condition

$$E[L_t | i \text{ is present}] \equiv E[L_t]$$

$$E[L_t | i \text{ is absent}] \equiv E'[L_t] \quad .$$

so that (3.9) and (3.10) give the desired result.

Q.E.D.

<u>Note</u>:

(1) If one node i is <u>added</u> to G, Lemma 3.8(a) remains valid if one exchanges t' and t in the inequalities.

(2) If one black node i of G becomes white, Lemma 3.8(b) remains valid if one exchanges E' and E in the inequalities.

(3) The results can be extended to the case of a general p.m.f. for W.

## 3.4 Relation Between the TSP and PTSP

As pointed out in the introduction, the TSP is a special case of a PTSP in which all nodes are black; it is then natural to investigate the possible links between the two problems. We will be concerned with two different (but related) issues in this section: first we would like to determine conditions (if any) under which a PTSP problem is solved by an optimal tour of the corresponding TSP problem; then following this "qualitative" search we will analyze quantitatively the question of how far the optimal TSP tour can be from optimality for the PTSP problem.

Lemma 3.9, which, in fact, will be mainly limited to very small size problems, will provide elements of the answer to the first issue; Theorem 3.1 will then provide an upper bound on the absolute "distance from optimality" of a TSP tour as a function of $\frac{E[W]}{n}$, the "average percentage" of present nodes among the set of white nodes.

We will then provide, through Lemmas 3.10 and 3.11, some indications concerning the sharpness of this bound and will also offer some insights into additional peculiarities of this problem.

Lemma 3.9:

Let $G = (N_1 \cup N_2, A, D)$ be a given graph with m black nodes ($N_1$) and n white nodes ($N_2$); then we have the following:

(i) Provided that the distance matrix D is symmetric then for any G of size up to 4 ($n+m=4$) we have $t_1 \equiv t_p$ for any p.m.f. for W; in case $m=0$ then this is also true for $n=5$.

(ii) If D is not symmetric, or if $n+m>5$, or if $m=0$, $n>6$, then it is possible to construct problem instances where $t_1 \neq t_p$ for some p.m.f. for W.

(iii) Let Convex($N_1 \cup N_2$) be the set of points of $N_1 \cup N_2$ that belong to the convex hull of $N_1 \cup N_2$ (Convex($N_1 \cup N_2$) $\subset N_1 \cup N_2$); if D corresponds to the Euclidean metric and if Convex($N_1 \cup N_2$) = $N_1 \cup N_2$ then $t_1 \equiv t_p \; \forall \; n \; \forall \; m$ and for any p.m.f. for W.

<u>Proof</u>: Let us consider the three cases separately:

(i): if $m+n \leq 3$ and D symmetric then we have a single tour, hence (i) is trivially true.

if $m+n=4$: then we have either (discarding $n=0$)

$$m=0 \Rightarrow E[L_t] = \alpha_0 L_{0,t}^{(o)} + \alpha_1 L_{0,t}^{(1)} + \alpha_2 L_{0,t}^{(2)}$$

$$m=1 \Rightarrow E[L_t] = \alpha_0 L_{1,t}^{(o)} + \alpha_1 L_{1,t}^{(1)} + \alpha_2 L_{1,t}^{(2)}$$

$$m=2 \Rightarrow E[L_t] = \alpha_0 L_{2,t}^{(o)} + \alpha_1 L_{2,t}^{(1)} + \alpha_2 L_{2,t}^{(2)}$$

$$m=3 \Rightarrow E[L_t] = \alpha_0 L_{3,t}^{(o)} + \alpha_1 L_{3,t}^{(1)} \quad .$$

But from Facts 3.2 and 3.4 we have that:

for $m=0$, $\sum_{r=0}^{2} L_{0,t}^{(r)}$ = constant and $L_{0,t}^{(2)} = L_{0,t}^{(o)}$;

for $m=1$, $L_{1,t}^{(2)}$ = constant and $\sum_{r=0}^{2} L_{1,t}^{(r)}$ = constant;

for $m=2$, $L_{2,t}^{(1)}$ = constant $L_{2,t}^{(2)}$ = constant;

for $m=3$, $L_{3,t}^{(1)}$ = constant .

Hence, for every case, $E[L_t] = K_1 L_{m,t}^{(o)} + K_2$

where: for $m=0$, $K_1 = \alpha_0+\alpha_2-2\alpha_1=(\alpha_0-\alpha_1)-(\alpha_1-\alpha_2)$ which is positive from Fact 3.5;

for m=1, $K_1 = \alpha_0 - \alpha_1$ also positive from Fact 3.5;

for m=2 and m=3, $K_1 = \alpha_0 > 0$.

So for all cases $\min_t E[L_t] \equiv \min_t L_{m,t}^{(o)}$ and this establishes (i) when m+n=4.

Let us now look at the case m=0 n=5. We have:

$$E[L_t] = \alpha_0 L_{0,t}^{(o)} + \alpha_1 L_{0,t}^{(1)} + \alpha_2 L_{0,t}^{(2)} + \alpha_3 L_{0,t}^{(3)}$$

and now from Facts 3.2 and 3.4 we have

$$\sum_{r=0}^{3} L_{0,t}^{(r)} = \text{constant and } L_{0,t}^{(3)} = L_{0,t}^{(o)},\ L_{0,t}^{(2)} = L_{0,t}^{(1)} .$$

This implies again that $E[L_t] = K_1 L_{m,t}^{(o)} + K_2$ and this time $K_1 = (\alpha_0+\alpha_3)-(\alpha_1+\alpha_2) = (\alpha_0-\alpha_1)-(\alpha_2-\alpha_3)$ and since $\alpha_0-\alpha_1 > \alpha_1-\alpha_2 > \alpha_2-\alpha_3$ from Fact 3.5 we proved the last case of (i)

(ii) to prove (ii) we will provide a counterexample for each of the two following cases:

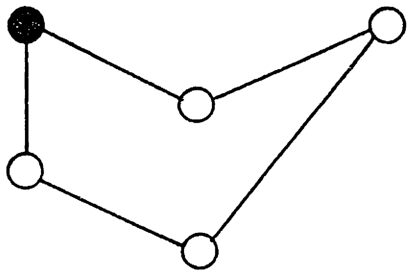
(1) m=0, n=6, D euclidean, W binomial with p=0.8

(2) m=1, n=4, D euclidean, W binomial with p=0.8

(One can also find an example when D is asymmetric and m=0, n=4.)

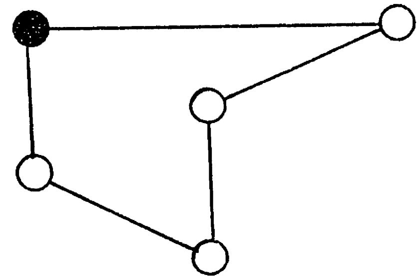
Figures 3.3, 3.4 provide graphical illustrations of the examples; the exact calculations are provided in Appendix D.

(iii) To show this part we use the following well-known property for the Euclidean TSP: the order in which the points on the convex hull appear in the optimal TSP tour must be the same as the order in which



Optimal TSP tour

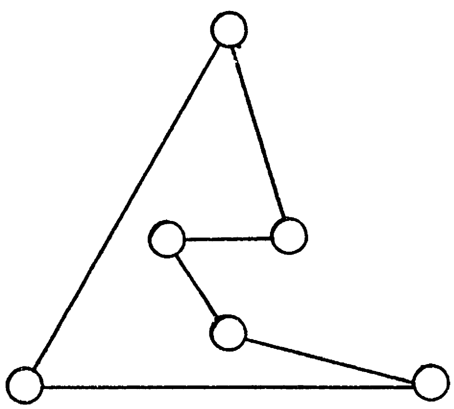
tour 1



Optimal PTSP tour for $p < 0.66$

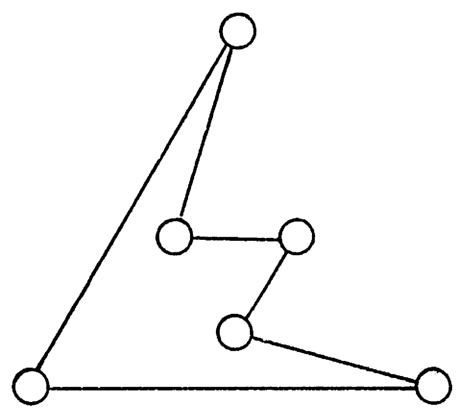
tour 2

Figure 3.3: Illustrations for m=1 n=4



Optimal TSP tour:

tour 3



Optimal PTSP tour for any p

tour 4

Figure 3.4: Illustrations for m=0 n=6

these points appear on the convex hull; this property follows directly from the fact that the optimal TSP tour does not intersect itself when D corresponds to the Euclidean metric.

Now, when $\text{Convex}(N_1 \cup N_2) = N_1 \cup N_2$, the TSP tour through $N_1 \cup N_2$ is simply the sequence of points appearing along $\text{Convex}(N_1 \cup N_2)$.

If we drop any set of points from $N_2$ (white nodes), say A, then it is easy to see that

$$\text{Convex}(N_1 \cup N_2) = N_1 \cup N_2 \Rightarrow \text{Convex}(N_1 \cup (N_2 - A)) = N_1 \cup (N_2 - A)$$

which implies that by simply skipping those points from the original TSP tour, we still end up with the optimal TSP tour through the restricted set of points (without reoptimizing) hence (iii) is proved.

Q.E.D.

<u>Note</u>: It is interesting to note that (iii) constitutes a case for which not only $t_p' \equiv t_1$ but also for which the expected length in the PTSP sense is identical to the expected length assuming reoptimization on every instance of the problem.

Let us now turn our attention to the general case for which $t_1 \neq t_p$ and give an upper bound on $\frac{E[L_{t_1}] - E[L_{t_p}]}{E[L_{t_p}]}$ that is a worst-case ratio for the TSP tour.

<u>Theorem 3.1</u>:

Let G be a given graph with n white nodes and m black nodes; then we have:

$$\text{(i)} \qquad \frac{E[L_{t_1}] - E[L_{t_p}]}{E[L_{t_p}]} < \frac{n - E[W]}{E[W]} = \frac{1 - E[W]/n}{E[W]/n}$$

(ii) $\dfrac{L^{(o)}_{t_p} - L^{(o)}_{t_1}}{L^{(o)}_{t_1}} < \dfrac{1-E[W]/n}{(E[W^2]-E[W])/n(n-1)}$

$\forall\ m \geq 1 \quad \forall\ n$

and $m=0 \quad \forall\ n$ prime .

<u>Proof</u>: (i) From Lemma 3.4 we have:

for $m \geq 1 \quad E[L_{t_1}] \leq L^{(o)}_{m,t_1}$

$$E[L_{t_p}] \geq L^{(o)}_{m,t_1} \ \frac{E[W]}{n}$$

which proves (i) for $m \geq 1$.

for $m=0$ and $n$ prime

$$E[L_{t_1}] \leq L^{(o)}_{0,t_1} \ \left[1-Pr(W=0)-Pr(W=1)\right]$$

$$E[L_{t_p}] \geq L^{(o)}_{0,t_1} \ \left[\frac{E[W]-Pr(W=1)}{n}\right] .$$

Hence $\dfrac{E[L_{t_1}] - E[L_{t_p}]}{E[L_{t_p}]} < \left(\dfrac{1-E[W]/n}{E[W]/n}\right)\left(\dfrac{1-(nPr(W=0)+(n-1)Pr(W=1))/(n-E[W])}{1-Prob(W=1)/E[W]}\right)$

and, since $Pr(W=1) \leq E[W][Pr(W=0)+Pr(W=1)]$,

this implies $\quad 1 - \dfrac{nPr(W=0)+(n-1)Pr(W=1)}{n-E[W]} \leq 1 - \dfrac{Pr(W=1)}{E[W]}$

and hence proves (i) for $m=0$ and $n$ prime. (We conjecture that (i) is still valid for $n$ not prime, but we haven't been able to prove it, since now Lemma 3.4 is of no use anymore.)

(ii) We will prove this part for m≥2. The other cases can be proved by exactly the same method. We have

$$E[L_{t_p}] = \sum_{r=0}^{n} \alpha_r L_{m,t_p}^{(r)} = \alpha_0 L_{m,t_p}^{(o)} + \sum_{r=1}^{n} \alpha_r L_{m,t_p}^{(r)} \quad . \tag{3.11}$$

Now from Lemma 3.2 together with Fact 3.6, we have

$$E[L_{t_p}] \geq \alpha_0 L_{m,t_p}^{(o)} + L_{m,t_1}^{(o)} \left[\frac{E[W]}{n} - \alpha_0\right] \tag{3.12}$$

From Lemma 3.4 we still have

$$E[L_{t_1}] \leq L_{m,t_1}^{(o)} \tag{3.13}$$

and by definition of the optimal PTSP tour $t_p$ we have:

$$E[L_{t_1}] - E[L_{t_p}] \geq 0 \tag{3.14}$$

Hence (3.12), (3.13), and (3.14) lead to:

$$L_{m,t_1}^{(o)} \left[1 + \alpha_0 - \frac{E[W]}{n}\right] - \alpha_0 L_{m,t_p}^{(o)} \geq 0$$

which in turn implies:

$$\frac{L_{m,t_p}^{(o)} - L_{m,t_1}^{(o)}}{L_{m,t_1}^{(o)}} \leq \frac{1 - E[W]/n}{\alpha_0} \quad . \tag{3.15}$$

It remains to evaluate $\alpha_0$. We have by definition:

$$\alpha_0 = \sum_{k=0}^{n-2} \left[\binom{n-2}{k} \Big/ \binom{n}{k}\right] \Pr(W=n-k) \quad .$$

which becomes, by setting u=n-k and noticing $\binom{n-2}{n-u} \Big/ \binom{n}{n-u} \equiv \binom{u}{2} \Big/ \binom{n}{2}$,

$$\alpha_0 = \sum_{u=2}^{n} \left(\binom{u}{2}/\binom{n}{2}\right) Pr(W=u) = \frac{1}{n(n-1)} \sum_{u=1}^{n} u(u-1)Pr(W=u) \quad .$$

Hence $\alpha_0 = \frac{1}{n(n-1)} \left(E[W^2] - E[W]\right)$

Q.E.D.

Comments:

1. When W is a binomial random variable with parameter p, i.e.,

$Pr(W=k)=\binom{n}{k} p^k(1-p)^{n-k}$ , Theorem 3.1 gives:

$$\frac{E[L_{t_1}] - E[L_{t_p}]}{E[L_{t_p}]} < \frac{1-p}{p} \text{ and } \frac{L_{t_p}^{(o)} - L_{t_1}^{(o)}}{L_{t_1}^{(o)}} < \frac{1-p}{p^2}$$

2. For the general case $\frac{E[W]}{n}$ represents the "average percentage of present white nodes" to be expected.

3. When $\frac{E[W]}{n}$ approaches zero the upper bounds of Theorem 3.1 approach $+ \infty$ .

An interesting question concerns the sharpness of the bounds provided by Theorem 3.1; it is reasonable to assume that when $\frac{E[W]}{n}$ is close to 1 the bounds of Theorem 3.1 are quite good; what can we say when $\frac{E[W]}{n}$ is close to zero?

Theorem 3.1(i) has been established by bounding $E[L_{t_1}]$ from above and $E[L_{t_p}]$ from below using results from Lemma 3.4. For example, for

$m>1$, $\qquad E[L_{t_1}] < L_{m,t_1}^{(o)}$

$$E[L_{t_p}] > L_{m,t_1}^{(o)} \quad \frac{E[W]}{n}$$

Our next lemma (3.10) shows that those bounds are the best possible in the sense that for each of them we can construct problem instances for which $E[L_{t_1}] = L^{(o)}_{m,t_1}$ or $E[L_{t_p}] = L^{(o)}_{m,t_1} \frac{E[W]}{n}$ (but not both at the same time!).

On the other hand, another lemma (3.11) will give more insight into why it is highly improbable (we conjecture impossible) that we can construct a problem instance where the TSP tour is "arbitrarily bad" for the PTSP (i.e., for which $\frac{1-E[W]/n}{E[W]/n}$ is met with $\frac{E[W]}{n}$ arbitrarily small). We will conclude the discussion by pointing out some additional peculiarities of the PTSP.

Lemma 3.10:

The lower and upper bounds given in Lemma 3.4 are the best possible in the sense that:

(i) there exists a problem instance $G_1(n,m)$ such that $E[L_{t_1}]$ equals the upper bounds given in Lemma 3.4.

(ii) there exists a problem instance $G_2(n,m)$ such that $E[L_{t_p}]$ equals the lower bounds given in Lemma 3.4.

Proof: The constructions are given in Figure 3.5 and Figure 3.6.

(i) for this part we will assume for simplicity that W is a binomial random variable with parameter p and that m=0.

In this example, the graph $G_1$ contains n white nodes that are positioned at the vertices of a n-gon and D is given by the Euclidean distance between the nodes; we choose the distance between two successive vertices of the n-gon to be $\frac{1}{n}$. It is then easy to see that the distance between one vertex and its $r^{th}$ subsequent vertex (clockwise or not) is given by:

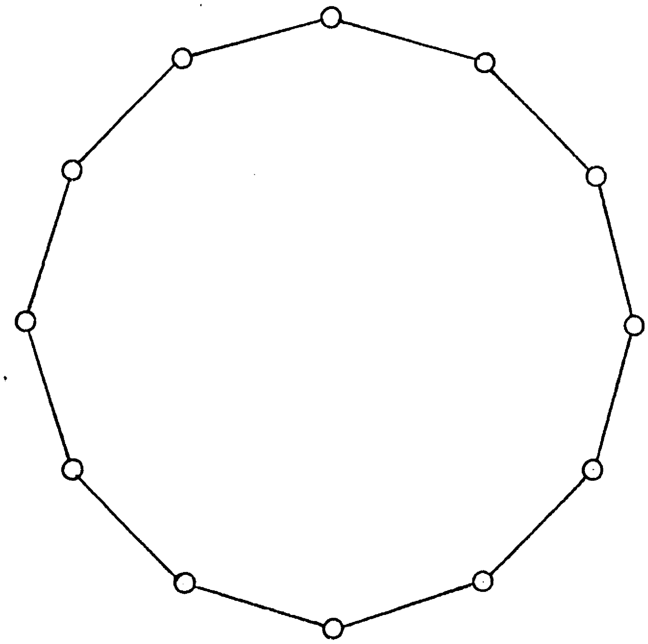


Figure 3.5: The Graph $G_1$

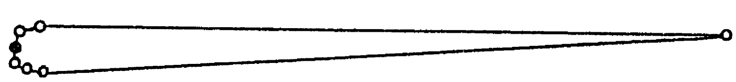


Figure 3.6: The Graph $G_2$

$$\frac{\sin(r\Pi/n)}{n\sin(\Pi/n)} \qquad \text{for } r \in [1..\lfloor\tfrac{n}{2}\rfloor]$$

Hence, by definition of the $L^{(r)}_{0,t_1}$, we have:

$$L^{(r)}_{0,t_1} = \frac{\sin\left((r+1)\Pi/n\right)}{\sin(\Pi/n)} \quad \text{for} \quad r \in [0..\lfloor\tfrac{n}{2}\rfloor-1]$$

(remember that $L^{(n-2-r)}_{0,t_1} \equiv L^{(r)}_{0,t_1}$) which implies that:

$L^{(o)}_{0,t_1} = 1$ for any n,

and $\lim_{n\to\infty} L^{(r)}_{0,t_1} = (r+1)$ (for any finite r).

Together with the fact that $\lim_{r\to\infty} (1-p)^r(r+1) = 0$ for any $0<p<1$, we finally have (the exact derivations are similar to the ones given in Appendix E for our next lemma):

$$\forall\ \varepsilon > 0 \quad \forall\ 0 < p < 1 \quad \exists\ N_{p,\varepsilon}: \ \forall\ n > N_p.$$

$$E[L_{t_1}] > L^{(o)}_{0,t_1} - \varepsilon \qquad .$$

Q.E.D.

(ii) For this case the construction is much simpler. For any n white nodes and m black nodes, construct a graph $G_2$ such that the m black nodes and n-1 white nodes are at almost the same position and the remaining white node is positioned at a distance x from this agglomeration of n+m-1 nodes.

We obviously have $L^{(o)}_{0,t_1} = 2x$ and $L^{(r)}_{0,t_1} = 2x\ \forall\ r$. So

$$E[L_{t_p}] \equiv E[L_{t_1}] = L^{(o)}_{0,t_1} \frac{E[W]}{n} \quad .$$

Q.E.D.

We have thus been able to construct problem instances for which $E[L_{t_1}] \sim L^{(o)}_{m,t_1}$ in one case and $E[L_{t_p}] \sim L^{(o)}_{m,t_1} \frac{E[W]}{n}$ in the other case. However, this does not imply that we can find one problem instance for which both bounds are met at the same time, hence providing a case where the optimal TSP tour can be made arbitrarily bad for the PTSP problem. In fact the conjecture here is that this cannot be done and that one of the worst cases against the TSP is given by the following lemma (based on the limiting behavior, i.e. $n \to \infty$, os the "star-shaped construction" introduced in Chapter 2 for n=24):

Lemma 3.11:

For any $0<p<1$ and $\varepsilon>0$ there exists a natural number $N_{p,\varepsilon}$ such that for any $n \geq N_{p,\varepsilon}$, one can construct a graph $G_3$ with n white nodes for which:

$$E[L_{t_1}]\left(\frac{1}{2-p} - \frac{p(1-p)^2}{2-p}\right) - \varepsilon \leq E[L_{t_p}] \leq E[L_{t_1}] \ \frac{1}{2-p} + \varepsilon$$

Proof: First, when p=1 Lemma 3.11 is obvious. Let us then concentrate on $0<p<1$.

Consider the following extension of the construction given in Chapter 2 in which a graph $G_3$ contains 2n white nodes positioned at the vertices of two concentric regular n-gons (see Figure 2.2). Consider as before tour a and tour b; assume that the distance between successive vertices of the outside n-gons is 1/n and then choose the distance between successive vertices of the inside n-gons $d_2$ such that (we want a to be the optimal TSP tour):

$$L_a^{(o)} = (n-1)\left[\frac{1}{n} + d_2\right] + 2\alpha \leq 2n\alpha = L_b^{(o)} \tag{3.16}$$

$$\alpha = \frac{1}{2 \sin \pi/n}\left[\sqrt{(1/n)^2 + d_2^2 - (2/n)\, d_2 \cos \pi/n}\right. \tag{3.17}$$

Using calculus one can show that

$$d_2 = \frac{1}{n}\left[K_n - \sqrt{K_n^2 - 1}\right] \text{ gives } L_a^{(o)} = L_b^{(o)} \quad \forall\, n$$

$$\text{where } K_n = \frac{1 + \cos\frac{\pi}{n} - \cos^2\frac{\pi}{n}}{\cos^2\frac{\pi}{n}} \quad .$$

It is then a matter of somewhat lengthly calculus (see Appendix E for details) to show that:

$$\lim_{n\to\infty} L_a^{(r)} = 2(r+1) \qquad \text{for any finite } r$$

$$\lim_{n\to\infty} L_b^{(r)} = 2\sqrt{r^2-r+1} \qquad \text{for any finite } r \tag{3.18}$$

and then to deduce that:

$$\forall\, \varepsilon_1 > 0 \;\; \forall\, \varepsilon_2 > 0 \; \forall\, 0 < p < 1 \text{ (hence } \forall\, \varepsilon = \varepsilon_{1/(2-p)}+\varepsilon_2)$$

$$\exists\; N_{p,\varepsilon} : \forall\, n > N_{p,\varepsilon} \;\; 2-\varepsilon_1 \leq E[L_a] \leq 2 + \varepsilon_1 \tag{3.19}$$

$$2\left[\frac{1}{2-p} - \frac{p(1-p)^2}{2-p}\right] - \varepsilon_2 \leq E[L_b] \leq \frac{2}{2-p} + \varepsilon_2 \; .$$

Q.E.D

Note: Lemma 3.11 implies that by choosing p very small one can construct examples where the optimal TSP tour has an expected length (in the sense of the PTSP) twice as big as the expected length of the optimal PTSP tour (that is, a worst case behavior of 100%).

In the process of proving Lemmas 3.10 and 3.11 we obtained the following interesting fact:

For graphs with white nodes and with W a binomial random variable with parameter p, no matter how small p is ($p=\epsilon>0$) one can constuct examples where:

(i) $E[L_t] = L_t^{(o)}$ (Lemma 3.10)

(ii) $E[L_{t_1}] \sim 2\, E[L_{t_p}]$ (Lemma 3.11) .

On the other hand, when p=0, then the expected length of any tour t of any arbitrary graph G is zero.

In other words:

for p = 0, the PTSP for any graph G is trivially solved.

for p > 0, no matter how small p is, there exists graphs G where the difference in expected length between tours is not negligible, implying that even for p very small different tours might not be (almost) equivalent with respect to the PTSP.

This can be explained very easily by noticing that given p is very small (but strictly positive), one can always choose n so that

$E[W] = np$ is arbitrarily large (however for p=0 $E[W] \equiv 0 \; \forall n$).

## 3.5 The PHPP.

All the results of the previous sections can be directly applied to the Probabilistic Hamiltonian Path Problem as originally defined in Chapter 2 (that is, Hamiltonian paths between two black nodes); this is true since, as mentioned in Chapter 2, one can always transform a path $h = (i_1, \ldots, i_{|N|})$ into a tour $t = (i_1, \ldots, i_{|N|})$ with expected lengths (under the assumption that $i_1$ and $i_{|N|}$ are black nodes) such that:

$$E[L_h] \equiv E[L_t] - d(i_{|N|}, i_1)$$

For the variation of the PHPP in which $i_1$ and $i_{|N|}$ are not necessarily black one can conduct an analysis similar to the one in the previous sections. We give now an additional lemma for this case of graphs without black nodes.

Lemma 3.12:

Let $h = (i_1, \ldots, i_n)$ be a path through n white nodes; let $h \equiv h_1 \oplus h_2$ where

$$h_1 = (i_1, \ldots, i_k),\ h_2 = (i_k, \ldots, i_n) \qquad \text{for } 1<k<n;$$

then $$E[L_{h_1 \oplus h_2}] \geq E[L_{h_1}] + E[L_{h_2}]$$

Proof:

$$E[L_{h_1 \oplus h_2}] = E[L_{h_1 \oplus h_2} | i_k \text{ is present}]\ \frac{E[W]}{n} + E[L_{h_1 \oplus h_2} | i_k \text{ is absent}]\ \left(1 - \frac{E[W]}{n}\right) \quad . \tag{3.20}$$

- $E[L_{h_1 \oplus h_2} | i_k$ is present] corresonds to $i_k$ being black and as we have seen in Chapter 2 we then have

$$E[L_{h_1 \oplus h_2} | i_k \text{ is present}] = E[L_{h_1} | i_k \text{ is present}] + E[L_{h_2} | i_k \text{ is present}]$$

- on the other hand

$$E[L_{h_1 \oplus h_2} | i_k \text{ is absent}] > E[L_{h_1} | i_k \text{ is absent}] + E[L_{h_2} | i_k \text{ is absent}]$$

Since the left-hand-side term is equal to the right hand side term plus the expected length of the segment joining the last node of $h_1$ to the first node of $h_2$ (which is $> 0$).

Q.E.D.

## 3.6 Conclusion

More than any other chapter of this thesis, this one presents a host of "technical" results.

In summary the important points are:

1. The development of upper and lower bounds on $E[L_t]$ (Lemma 3.4) that can be shown to be the best possible (Lemma 3.10).

2. The entire section 3.3 in which properties of the optimal PTSP tour have been established (Lemmas 3.6, 3.7, and 3.8).

3. The entire section 3.4 in which we compare the optimal PTSP tour with the optimal TSP tour (Lemmas 3.9, 3.11, Theorem 3.1).

4. The lemma concerning the "non-additivity" of $E[L]$ (Lemma 3.12).

All these results will play key role in Chapter 5, the chapter dealing with the development of solution procedures for the PTSP.

# CHAPTER 4

# BOUNDS AND ASYMPTOTIC ANALYSIS FOR THE PTSP IN THE PLANE

## 4.1 Introduction; Notation

The methods of analysis considered in this chapter are considerably different from those of the previous two chapters. An asymptotic analysis is performed in which set-theoretic concepts are used instead of graph-theoretic ones. The main objective of this chapter is to obtain strong (i.e., concerned with convergence with probability 1) limit laws for the PTSP similar to the celebrated result for the TSP obtained by Beardwood et al. [1959]. Informally stated, this result states that the value of the optimal TSP through n points drawn from a uniform distribution in the unit square is almost surely (with probability 1) asymptotic to $\beta\sqrt{n}$. (The constant $\beta$ has been estimated to be approximately 0.765; see Stein [1977]). In fact, this result can be extended to an arbitrary Lebesgue measureable set of a d-dimensioned Euclidean space and with an arbitrary probability distribution for the location of the points; the constant of interest $\beta(d)$ depends only on the space dimensionality and not on the shape of the set considered.

This theoretical result has become widely recognized to be at the heart of the probabilistic evaluation of the performance of heuristic algorithms for vehicle routing problems. In fact it was used as the main argument in the probabilistic analysis of heuristics for the TSP (Karp [1977]). Because of those algorithmic applications, results like that of Beardwood et al. have gained considerable practical interest. Steele [1981a] uses the theory of independent subadditive processes to obtain strong limit laws for a class of problems in geometrical probability that exhibit nonlinear growth (Steele's theorem will be essential in the development of some of our results).

Before presenting the notation to be used in the statements and proofs of our results, let us briefly sketch the outline of this chapter. First we consider a set of points in 2-dimensional Euclidean space $R^2$, assuming the distance between points to be the ordinary Euclidean distance. In section 4.2 we present an upper bound on the expected length of the optimal PTSP tour for an arbitrary sequence of n points lying in a square of side r. Assuming the points are uniformly and independently distributed over the square, we obtain another upper bound as well as a lower bound on the expected length of the optimal PTSP tour. These results apply to problems with a finite set of n points; in section 4.3 we turn our attention to asymptotic behavior (i.e., $n \to \infty$). In a first subsection we present the asymptotic behavior of the expected value given by the strategy of reoptimizing the optimal tour for each realization of the random variables (i.e., for each subset of the points that will actually need a visit on a particular instance, we construct the optimal TSP tour and compute its length); this, of course, constitutes a lower bound on the expected length of the optimal PTSP tour. The second subsection contains the most important theoretical result of this chapter: we show that the expected length (in the PTSP sense) of the optimal PTSP tour (coverage probability p) through n points drawn from a uniform distribution in the unit square is almost surely (with probability 1) asymptotic to $c(p)\sqrt{n}$, where c(p) is a constant depending on the coverage probability p; the third subsection of section 4.3 is then concerned with the derivation of bounds (upper and lower) on c(p). We then present, in section 4.4, generalizations of our results in several directions: first, we notice that all our previous results extend to the case where one of the points is always present (a depot); we then present extensions for cases where more than one point is always present (we also briefly mention similar results for the PHPP). Then we discuss extensions of our results to any bounded Lebesgue measurable set of a d-dimensioned Euclidean space. Finally, we conclude the chapter with a brief discussion of the practical implications of our results.

Notation:

- $R^2$ denotes the 2-dimensional Euclidean space.
- $[o,r]^2$ is the square of side r (r being a positive real number)
- $x = \{x_1, x_2, \ldots\}$ represents a (countably infinite) sequence of points on $R^2$; $x^{(n)}$ indicates the first n points of x, i.e., $x^{(n)} = \{x_1, x_2, \ldots, x_n\}$. If the positions of the points are random, the sequence will be denoted by upper-case letters, i.e., $X = \{X_1, X_2, \ldots\}$.
- given a sequence x, $t(x^{(n)})$ represents a tour through the points of the sequence $x^{(n)}$; $L_t(x^{(n)})$ is the length of the tour $t(x^{(n)})$; $EL_t(x^{(n)}, p)$ is the expected length (in the PTSP sense) of the tour $t(x^{(n)})$ when each point of the sequence $x^{(n)}$ requires a visit only with probability p, independently of each other. Two specific tours will play a key role in our derivations:

$t_1(x^{(n)})$ is the optimal TSP tour through $x^{(n)}$;

(i.e., the tour $t_1$ that solves $\min_t (L_t(x^{(n)}))$)

$t_p(x^{(n)})$ is the optimal PTSP tour through $x^{(n)}$

when the coverage probability is p;

(i.e., the tour $t_p$ that solves $\min_t (EL_t(x^{(n)}, p))$).

- W is the random variable representing the total number of points to be visited ($0 \leq W \leq n$); S is the random variable representing the specific subset of the points of $x^{(n)}$ that actually need a visit (there are $2^n$ possible subsets $s_j$); hence we have for any subset $s_j$ of cardinality $|s_j|=k$

$$\Pr(S=s_j) = 1/\binom{n}{k} \Pr(W=k) \tag{4.1}$$

- if for each realization of S, say $s_j$, we construct the optimal TSP tour $t_1(s_j)$, the expected value is:

$$E[L_{t_1}(S)] = \sum_{j=1}^{2^n} L_{t_1}(s_j) \Pr(S=s_j) \tag{4.2}$$

together with (4.1), this expression becomes:

$$E[L_{t_1}(S)] = \sum_{k=0}^{n} \left(1/\binom{n}{k}\right) \left[ \sum_{s_j, |s_j|=k} L_{t_1}(s_j) \right] \Pr(W=k) \tag{4.3}$$

[Note the difference between $EL_{t_1}(x^{(n)},p)$ and $E[L_{t_1}(S)]$. The former is the expected length (in the PTSP sense) of the optimal TSP tour through $x^{(n)}$ when points are present only with probability p; the latter quantity represents the expected length computed under the strategy of reoptimizing the tour for each subset $s_j$ (and we have: $E[L_{t_1}(S)] \leq EL_{t_p}(x^{(n)},p) \leq EL_{t_1}(x^{(n)},p)$).]

In fact this notation is not specific to $t_1$; for any algorithm A which, given a set of points on the plane, produces a feasible tour $t_A$, we write:

$EL_{t_A}(x^{(n)},p)$ for the expected length (in the PTSP sense) of $t_A(x^{(n)})$, and

$E[L_{t_A}(S)]$ for the expected length obtained under the strategy of applying the algorithm A for each realization $S=s_j$ (to get $t_A(s_j)$).

[Note that in general $E[L_{t_A}(S)] \neq EL_{t_A}(x^{(n)},p)$].

- It is useful to note that our notation implies some very specific rules:

$L_t(x^{(n)})$ is a real number

$L_t(X^{(n)})$ is a random variable (according to the position of the points)

also $EL_t(x^{(n)},p)$ is a real number but

$EL_t(X^{(n)},p)$ is a random variable.

So when we take the expectation of our random variables (•) with respect to the position of the points we will use the Eu[(•)] to avoid confusion with the other type of expectation (with respect to which point has to be visited).

- Finally:

$x^{(n)} - s_j$ denotes the set of points of $x^{(n)}$ that do not belong to $s_j$ ,

$g(n) = o(f(n))$ means $\lim_{n \to \infty} \frac{g(n)}{f(n)} = 0$ ,

$h(n) = O(f(n))$ means that there exists a constant $\alpha > o$ such that $h(n) < \alpha f(n)$ for all $n \geq 1$.

## 4.2 Bounds for Sequence of Points in $[o,r]^2$

### 4.2.1 Case of an Arbitrary Sequence x

Lemma 4.1

Let x be an arbitrary sequence of points in $[o,r]^2$ and p be the coverage probability for each point, then the expected length of the optimal PTSP tour satisfies:

$$\text{(i)} \quad EL_{t_p}\left(x^{(n)},p\right) \leq \left(\sqrt{2(np-2)} + \frac{13}{4}\right) r \qquad \forall\ np \geq 2.5$$

$$\text{(ii)} \quad EL_{t_p}\left(x^{(n)},p\right) \leq \left(\frac{np}{2} + 3\right) r \qquad \forall\ np \leq 2.5$$

Proof: If p=0 then $EL_{t_0}\left(x^{(n)},0\right) \equiv 0$ and the lemma is trivially verified; so let us suppose $0 < p \leq 1$. The demonstration of this lemma involves the construction of two tours $t_{c1}\left(x^{(n)}\right)$ and $t_{c2}\left(x^{(n)}\right)$ (a similar construction was used by Few [1955] to improve an upper bound for the TSP developed by Verblunsky [1951]). The construction goes as follows: (see Figure 4.1 for a graphical description).

Divide the initial square $[o,r]^2$ into 2h rows of equal width (h is a positive integer to be chosen later); there are then 2h+1 horizontal lines and 2 vertical lines on the square. Starting from the top horizontal line and moving downward discard every other horizontal line; connect each point of $x^{(n)}$ to the nearest of the remaining h+1 horizontal lines by a double vertical link, then connect each of the h+1 horizontal

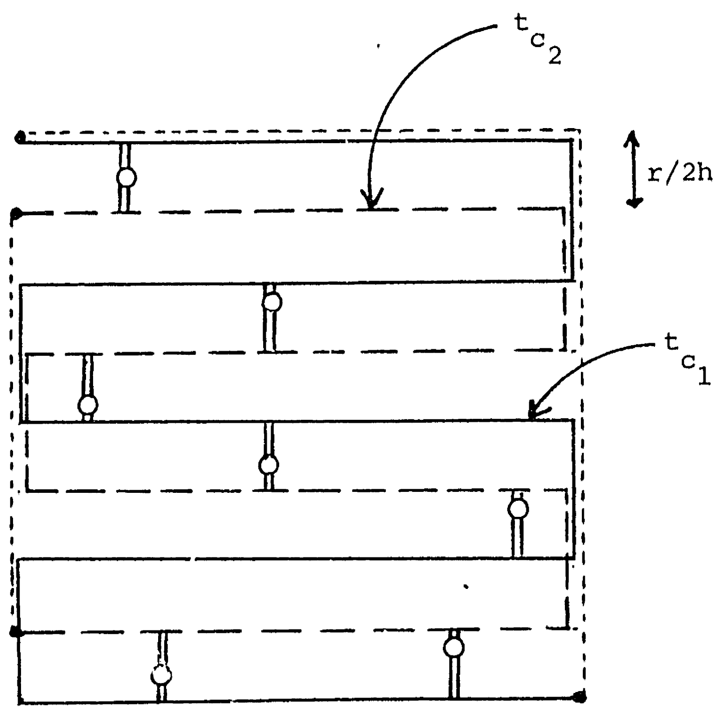



Figure 4.1: Construction of a Tour

lines to a line immediately above and/or below it by using some portion of the two vertical sides of the square to finally obtain a hamiltonian path through $x^{(n)}$. Construct a similar Hamiltonian path using the other h horizontal lines (previously discarded). Then extend each of these two hamiltonian paths to a tour by adding appropriate horizontal and vertical segments (giving respectively $t_{c1}$ and $t_{c2}$). Those two tours have the following properties:

(a) given any subset $s_j$ of $x^{(n)}$ ($j \in [1..2^n]$), the construction of $t_{ci}(s_j)$ gives the <u>same</u> tour as the resulting process of skipping the points of $(x^{(n)} - s_j)$ from $t_{ci}(x^{(n)})$ (in the PTSP sense). This is because the position of each point along the tour $t_{ci}(x^{(n)})$ depends only on its location within the square and not on the locations or presence of the other points. This property implies that

$$EL_{t_{ci}}(x^{(n)},p) \equiv E[L_{t_{ci}}(S)]$$

(b) given the <u>number</u> of points present is k, then $L_{t_{c1}}(s_j) + L_{t_{c2}}(s_j)$ is <u>constant</u> for all subsets $s_j$ of $x^{(n)}$ containing k points ($\binom{n}{k}$ such subsets) that is, it does not depend on <u>which</u> points are present (this is so by construction; indeed taking any point within the square $[o,r]^2$, the sum of the double vertical links that connect it to $t_{c1}$ and the double vertical link that connect it to $t_{c2}$ is exactly $\frac{r}{h}$, hence independent of its exact location). So if we call this constant $\ell_k$ when the number of points present is k, then this property implies (using (4.3))

$$E[L_{t_{c1}}(S) + L_{t_{c2}}(S)] = \sum_{k=o}^{n} \ell_k \Pr(W=k) \ . \qquad (4.4)$$

These two properties and their consequences imply that

$$\sum_{i=1}^{2} EL_{t_{ci}}\left(x^{(n)},p\right) = \sum_{k=0}^{n} \ell_k \Pr(W=k) \tag{4.5}$$

and since $t_p\left(x^{(n)}\right)$ is the optimal PTSP tour

$$EL_{t_p}\left(x^{(n)},p\right) \leq EL_{t_{ci}}\left(x^{(n)},p\right) \qquad \text{for } i = 1,2$$

hence

$$EL_{t_p}\left(x^{(n)},p\right) \leq \frac{1}{2} \sum_{k=0}^{n} \ell_k \Pr(W=k) \ . \tag{4.6}$$

It remains to evaluate $\ell_k \ \forall \ k \in [0..n]$: from the construction of $t_{c1}$ and $t_{c2}$, we can see that $\ell_k$ is given by:

$\ell_k = A(k) + B$ where $A(k)$ is the sum of the two Hamiltonian paths and B the sum of the additional segments to modify the two paths into tours.

- for any choice of h (odd or even) B is easily seen to be $3r - \frac{r}{h}$
- using property (ii), we obtain:

$$A(k) = \underbrace{(2h+1)r}_{\text{horizontal lines}} + \underbrace{k\left(\frac{r}{h}\right)}_{\text{connections of point}} + \underbrace{2\left(r - \frac{r}{2h}\right)}_{\text{vertical side sections}}$$

hence $\ell_k = r\left[2h + \frac{1}{h}(k-2) + 6\right]$ (4.7)

replacing in (4.6) we get

$$EL_{t_p}\left(x^{(n)},p\right) \leq \frac{r}{2}\left[2h + \frac{1}{h}\left[E[W]-2\right] + 6\right] \tag{4.8}$$

and this bound is valid for any integer $h \geq 1$.

Let us find the best integer $h^*$ ($h^* \geq 1$):

Note that if $E[W]-2 \leq 0$ then $h^* = 1$.

If $E[W]-2 > 0$, $h^*$ should be the nearest integer to $\sqrt{(E[W]-2)/2}$, except in the case where $\sqrt{(E[W]-2)/2}$ is closer to 0 than to 1 since $h^* \geq 1$. The threshold will then be given by

$$\sqrt{(E[W]-2)/2} = 0.5 \qquad \text{or } E[W] = 2.5$$

Hence if:

- $E[W] \leq 2.5$, then $h^* = 1$ and in replacing $h=1$ and $E[W] = np$ in (4.8) we get part (ii) of Lemma 4.1
- $E[W] \geq 2.5$, $h^*$ satisfies $(h^* + \theta)^2 = \frac{E[W]-2}{2}$ where $\theta$ is a real number such that $|\theta| \leq \frac{1}{2}$. The right hand side of (4.8) becomes:

$$\frac{r}{2}\left[2h^* + \frac{1}{h^*}\,2(h^*+\theta)^2 + 6\right] = \frac{r}{2}\left[4(h^*+\theta + \frac{2\theta^2}{h^*} + 6\right]$$

and since $h^* \geq 1$ and $|\theta| \leq \frac{1}{2}$ we get:

$$EL_{t_p}\left(x^{(n)},p\right) \leq \frac{r}{2}\left[4\sqrt{(E[W]-2)/2} + 2\,\frac{1}{4} + 6\right].$$

Since $E[W] = np$, part (i) of Lemma 4.1 is verified.

Q.E.D.

Lemma 4.1 is concerned with arbitrary sequences in $[o,r]^2$ and thus under such conditions, the only valid lower bound on $EL_{t_p}(x^{(n)},p)$ is 0 $\forall$ p (take the sequence x consisting of points positioned at the exact same location). In the next subsection we are going to sharpen both the upper and the lower bound under the additional assumption that the sequence X is i.i.d. uniform over $[o,r]^2$ (this time, of course, considering $Eu[EL_{t_p}(X^{(n)},p)]$)

4.2.2 Case of a Uniform Sequence X

Lemma 4.2:

Let X be an infinite sequence of points independently and uniformly distributed over $[o,r]^2$ and p be the coverage probability for each point, then we have:

$$\text{(i)} \quad E_u[EL_{t_p}(X^{(n)},p)] \leq \left(\sqrt{(4/3)(np-3)} + \frac{11}{12} + \sqrt{2}\right) r \qquad \text{if } np \geq 3.75$$

$$\text{(ii)} \quad E_u[EL_{t_p}(X^{(n)},p)] \leq \left(\frac{np}{3} + \frac{2}{3} + \sqrt{2}\right) r \qquad \text{if } np \leq 3.75 \ .$$

Proof: here again the demonstration involves the construction of a tour through $X^{(n)}$ obtained by dividing the square $[o,r]^2$ in almost the same way as before; more precisely (see Figure 4.2 for a graphical illustration), divide the initial square into h equal-width (r/h) columns. Construct the tour $t_{c3}(X^{(n)})$ as follows: start from the point in the leftmost column having the largest vertical coordinate, then visit the point in the same column having the next lower vertical coordinate, and so on; from the lowest point of this column we go to the lowest point in the next column and visit the points in that column upward. We continue this

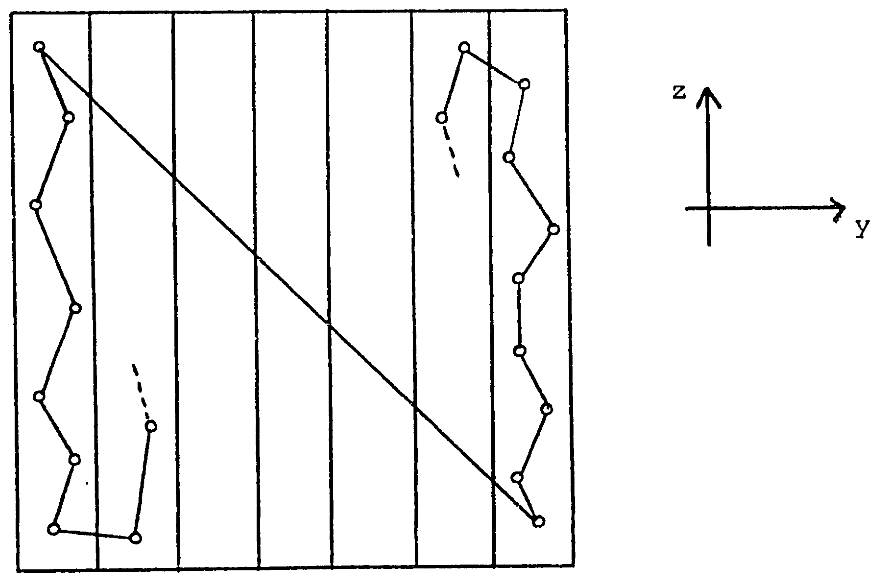



Figure 4.2: Construction a Tour

process until we reach the last point; we complete the tour by connecting the first and last point by a straight line (a similar construction was proposed and used for studying the TSP in Beardwood et al. [1959]; more specifically, this construction was essential in the proof of their lemmas 4 and 10). It is easy to see that $t_{c3}$ possesses a property similar to (b):

(b') given the number of points present in k then $Eu[L_{t_{c3}}(s_j)] = Eu[L_{t_{c3}}(s_{j'})]$ for all subsets $s_j$ and $s_{j'}$ of $X^{(n)}$ containing $\underline{\underline{k}}$ points. The reason for (b') is that $L_{t_{c3}}(s_j)$ and $L_{t_{c3}}(s_{j'})$ have exactly the same p.d.f. since both of them are tours through k points (even if different in the two subsets) that are distributed identically and independently over $[o,r]^2$.

So, finally, (a) and (b') imply that:

$$Eu[EL_{t_p}(X^{(n)},p)] \underset{\substack{\uparrow \\ t_p \text{ is the optimal} \\ \text{PTSP tour}}}{\leq} Eu[EL_{t_{c3}}(X^{(n)},p)] \underset{\substack{\uparrow \\ (a)}}{\equiv} Eu[E[L_{t_{c3}}(S)]$$

and $$Eu[E[L_{t_{c3}}(S)]] \underset{\substack{\uparrow \\ (b')}}{\equiv} \sum_{k=0}^{n} Eu[L_{t_{c3}}(k)] \Pr(W=k)$$

where $Eu[L_{t_{c3}}(k)]$ is the generic term that corresponds to $Eu[L_{t_{c3}}(s_j)]$ for any subset $s_j$ of $X^{(n)}$ of cardinality k.

It remains now to evaluate $Eu[L_{t_{c3}}(k)]$: first note that the last segment connecting the first and last point has an expected length less that $\sqrt{2}\, r$.

For the k-1 other links, it is convenient to project them on the horizontal Y-axis and vertical Z-axis.

- the projection of those k-1 links on the Y-axis gives a total expected length of $= \left(\frac{h-1}{h}\, r + \frac{k-h}{3h}\, r\right)$

$\left\{\begin{array}{l}\text{h-1 between column distances}\\ \text{of expected length } r/h\end{array}\right\}$ $\quad$ $\left\{\begin{array}{l}\text{k-h within column distances}\\ \text{of expected length } r/3h\end{array}\right\}$

- the projection on the Z-axis is bounded from above by hr

(h column and for each column an upper bound of r).

It follows that:

$$Eu\left[L_{t_{c3}}(k)\right] \leq \left(\frac{h-1}{h} + \frac{k-h}{3h} + h + \sqrt{2}\right) r =$$

$$= \left[\left(\sqrt{2} + \frac{2}{3}\right) + h + \frac{k-3}{3h}\right] r \quad .$$

Hence we finally obtain

$$Eu\left[EL_{t_p}\left(X^{(n)}, p\right)\right] \leq \left(\sqrt{2} + \frac{2}{3} + h + \frac{E[W]-3}{3h}\right) r \ . \qquad (4.9)$$

By following a procedure similar to the one described in the proof of Lemma 4.1 we then obtain the optimum value of the integer h*. Here the threshold corresponds to

$$\sqrt{\left(E[W]-3\right)/3} = 0.5 \quad \text{or} \quad E[W] = 3.75 \ .$$

Hence if:

- $E[W] \leq 3.75$ then $h^*=1$ and we get (using (4.9)) part (ii) of Lemma 4.2
- $E[W] \geq 3.75$ then $h^*$ satisfies

$$(h^* + \theta)^2 = \frac{E[W]-3}{3} \text{ where } \theta \text{ is a real number}/\left|\theta\right| \leq \frac{1}{2} \ .$$

The right hand side of (4.9) becomes:

$$r\left[\sqrt{2} + \frac{2}{3} + h^* + \frac{(h^*+\theta)2}{h^*}\right] = r\left[2(h^*+\theta) + \frac{\theta 2}{h^*} + \frac{2}{3} + \sqrt{2}\right]$$

$$\leq r\left(\sqrt{(4/3)\left(E[W]-3\right)} + \frac{11}{12} + \sqrt{2}\right) .$$

Q.E.D.

Let us now give a lower bound on $Eu\left[EL_{t_p}\left(X^{(n)},p\right)\right]$

Lemma 4.3:

Let X be a sequence of points independently and uniformly distributed over $[o,r]^2$. Then we have:

$$Eu\left[EL_{t_p}\left(X^{(n)},p\right)\right] \geq \left[\frac{5}{8}\left(E[\sqrt{W}] - n(1-p)^{n-1}\right) + \left(\Upsilon - \frac{5\sqrt{2}}{16}\right)n(n-1)p^2(1-p)^{n-2}\right]r$$

for all positive integers n and with $\Upsilon = \frac{2+\sqrt{2}}{15} + \frac{1}{3}\ \ell n(1+\sqrt{2}) \simeq 0.5214$.

Proof: First, since the strategy of reoptimizing the tour, (given that the subset $S=s_j$ of $X^{(n)}$ has to be visited on a particular instance) is the strategy corresponding to perfect information, it provides a lower bound on $EL_{t_p}\left(X^{(n)},p\right)$. We then have $EL_{t_p}\left(X^{(n)},p\right) \geq E\left[L_{t_1}(S)\right]$ (note that this inequality is concerned with random variables). The argument labeled (b') in the proof of Lemma 2 is still valid for $t_1$; hence we have

$$Eu\left[EL_{t_p}\left(X^{(n)},p\right)\right] \geq \sum_{k=0}^{n} Eu\left[L_{t_1}(k)\right] Pr(W=k)$$

where again $Eu\left[L_{t_1}(k)\right]$ is the generic term for $Eu\left[L_{t_1}(s_j)\right]$ for any subset $s_j$ of $X^{(n)}$ of cardinality k. [Formally $Eu\left[L_{t_1}(k)\right] = Eu\left[L_{t_1}(Y^{(k)})\right]$ where Y is i.i.d. uniform over $[o,r]^2$.]

Now it remains to evaluate $Eu[L_{t_1}(k)]$:

<u>for</u> k=0,1 $Eu[L_{t_1}(k)] = 0$ .

<u>for</u> k=2, $Eu[L_{t_1}(2)]$ is twice the expected distance between two points independently and uniformly distributed over $[o,r]^2$; this expected distance is given by:

$$\gamma r \equiv \left[\frac{2+\sqrt{2}}{15} + \frac{1}{3}\,\mathrm{Ln}\left(1+\sqrt{2}\right)\right] r \approx 0.5214r$$

(see Gilbert [1965], for example).

<u>for</u> k $\geq$ 3, one can use Lemma 3 of Beardwood et al. [1959] to obtain:

$$Eu[L_{t_1}(k)] \geq \frac{5}{8}\sqrt{k}\ r.$$

Since the argument used to derive this result for the TSP has been used over and over again for obtaining lower bounds for other Euclidean network design problem (see Papadimitriou [1978]), it might be worthwhile to sketch it briefly: the main observation here is that as the shortest tour through all points does not visit any point twice, the sum of the lengths of the two segments of this tour that terminate at a given point is at least as large as the sum of the distance from this point to its first and second nearest points. Let $Y^{(k)} = \{Y_1, Y_2, .., Y_k\}$ be the set of points uniformly and independently located over $[o,r]^2$; let $\ell_{1j}$ and $\ell_{2j}$ be the distances of $Y_j$ to its nearest and second nearest point (respectively). We then have $Eu[L_{t_1}(Y^{(k)})] \geq \frac{1}{2}\sum_{j=1}^{n} Eu[\ell_{1j} + \ell_{2j}]$

$$Eu[L_{t_1}(Y^{(k)})] \geq \frac{1}{2}\sum_{j=1}^{n} Eu[\ell_{1j} + \ell_{2j}] = \frac{n}{2}\,Eu[\ell_{11} + \ell_{21}] \text{ (by symmetry).}$$

We can evaluate $Eu[\ell_{11} + \ell_{21}]$ by noting that

$$\text{Prob}(\ell_{11} > \ell) = \left(1 - \frac{V_\ell}{r^2}\right)^{n-1}$$

where $V_\ell$ is the area of the intersection of the square $[o,r]^2$ and the disc of center $Y_1$ and radius $\ell$, and

$$\text{Prob}(\ell_{21} > \ell) = \left(1 - \frac{V_\ell}{r^2}\right)^{n-1} + n-1 \left(\frac{V_\ell}{r^2}\right)\left(1 - \frac{V_\ell}{r^2}\right)^{n-2} ,$$

so that $Eu[\ell_{11} + \ell_{21}] = \int_0^\infty (\Pr(\ell_{11} > \ell) + \Pr(\ell_{12} > \ell))d\ell.$

By using some tedious calculus, we can then obtain a lower bound on $Eu[\ell_{11} + \ell_{21}]$ (see Beardwood et al. [1959] p. 309-310) and get the desired result.

So finally by combining the three cases we get

$$Eu[EL_{t_p}(X^{(n)},p)] > [2\Upsilon \Pr(W=2) + \frac{5}{8} \sum_{k=3}^{n} \sqrt{k} \Pr(W=k)]\ r\ .$$

The right hand side can be written as

$$[\frac{5}{8} \sum_{k=0}^{n} \sqrt{k}\ \Pr(W=k) + (2\Upsilon - \frac{5\sqrt{2}}{8})\Pr(W=2) - \frac{5}{8} \Pr(W=1)]r$$

and since: $\sum_{k=0}^{n} \sqrt{k}\ \Pr(W=k) = E[\sqrt{W}],$

$$\Pr(W=2) = \frac{n(n-1)}{2} p^2(1-p)^{n-2}, \text{ and}$$

$$\Pr(W=1) = np\ (1-p)^{n-1},$$

we obtain the desired bound.

Q.E.D.

Note:

- $\sqrt{x} > \frac{x}{\sqrt{n}} \quad \forall \quad x \in [0,n]$ .

Hence $\quad E[\sqrt{W}] > E[\frac{W}{\sqrt{n}}] = \frac{E[W]}{\sqrt{n}} = p\sqrt{n}$

- $2\gamma - \frac{15}{24}\sqrt{2} > 0$ .

Hence Lemma 4.3 can be simplified to (but at the cost of a smaller lower bound):

$$Eu[EL_{t_p}(X^{(n)},p)] > \frac{5}{8}\, p\sqrt{n}\, [1 - \sqrt{n}\,(1-p)^{n-1}]r \quad .$$

4.3 Asymptotic Analysis

In section 4.2 we were interested in results valid for problems of finite size, i.e., lower and upper bounds for any values of n (for an arbitrary sequence x or uniform sequence X). Let us now concentrate on asymptotic results.

4.3.1 Asymptotic Behavior of the "Strategy of Reoptimizing"

We are interested, in this section, in the behavior of the random variable $E[L_{t_1}(S)]$ (that is, corresponding to the case in thich we construct the optimal TSP tour for each subset of points $s_j$ of $X^{(n)}$ that actually are present on a given instance) when $n \to \infty$. The following theorem is proved using the results of Beardwood et al. (presented in our introduction).

Theorem 4.1:

Let X be an infinite sequence of points independently and uniformly distributed over $[0,1]^2$ and p be the coverage probability for each point. Then we have:

$$\lim_{n\to\infty} \frac{E\left[L_{t_1}(S)\right]}{\sqrt{n}} = \beta\sqrt{p} \qquad \text{(a.s.)}$$

$\forall\, p \in [0,1]$, where $\beta$ is the "TSP constant".

Proof: • First note that, when p=1, $E\left[L_{t_1}(S)\right] \equiv L_{t_1}\left(X^{(n)}\right)$ and Theorem 4.1 corresponds to the result of Beardwood et al. [1959].

- Also, when p=0, Theorem 4.1 is trivially valid.
- Let us then consider the non-trivial case $0 < p < 1$.

From the Weak Law of Large Number, we have:

$$\begin{aligned} &\forall\, \varepsilon > 0\ \exists\, N\varepsilon:\ \forall\, n > N\varepsilon,\quad \Pr\left(\left|\frac{W}{n} - p\right| > \varepsilon\right) < \varepsilon \qquad \text{or,} \\ &\forall\, \varepsilon > 0\ \exists\, N\varepsilon:\ \forall\, n > N\varepsilon,\quad \Pr\left(n(p-\varepsilon) < W < n(p+\varepsilon)\right) > 1-\varepsilon \end{aligned} \tag{4.10}$$

by taking $\varepsilon$ small enough so that $\left|n(p+\varepsilon)\right| + 1 < n$, we can write (4.3) (in view of (4.10)):

$$\frac{E\left[L_{t_1}(S)\right]}{\sqrt{n}} = \sum_{k=0}^{\lceil n(p-\varepsilon)\rceil - 1} \frac{L_{t_1}(k)}{\sqrt{n}} \Pr(W=k) + \sum_{k=\lceil n(p-\varepsilon)\rceil}^{\lfloor n(p+\varepsilon)\rfloor} \frac{L_{t_1}(k)}{\sqrt{n}} \Pr(W=k)$$

$$+ \sum_{k=\lfloor n(p+\varepsilon)\rfloor+1}^{n} \frac{L_{t_1}(k)}{\sqrt{n}} \Pr(W=k)$$

where $$L_{t_1}(k) \equiv 1/\binom{n}{k} \sum_{\substack{s_j \\ |s_j|=k}} L_{t_1}(s_j) \tag{4.11}$$

$\lceil \alpha \rceil$ is the smallest integer greater or equal to $\alpha$

$\lfloor \alpha \rfloor$ is the largest integer smaller or equal to $\alpha$

from Lemma 4.1 (taking r=1, p=1), we can easily show that there exists a constant b such that

$$L_{t_1}(s_j) \leq b\sqrt{k} \tag{4.12}$$

for all $s_j$ such that $|s_j| = k$ $\quad k \in [0..n]$

(in fact any $b \geq 3.5$ would do).

(4.10), (4.11), and (4.12) imply:

$$\forall\ \varepsilon > 0\ \exists N_\varepsilon :\ \forall\ n \geq N_\varepsilon$$

$$\sum_{k=\lceil n(p-\varepsilon)\rceil}^{\lfloor n(p+\varepsilon)\rfloor} \frac{L_{t_1}(k)}{\sqrt{n}} \Pr(W=k) \leq \frac{E[L_{t_1}(S)]}{\sqrt{n}} \leq \sum_{k=\lceil n(p-\varepsilon)\rceil}^{\lfloor n(p+\varepsilon)\rfloor} \frac{L_{t_1}(k)}{\sqrt{n}} \Pr(W=k) + b\varepsilon \ . \tag{4.13}$$

From Beardwood et al. [1959], we have:

$$\forall\ \varepsilon > 0\ \ \exists\ K_\varepsilon :\ \forall\ k \geq K_\varepsilon\ \ (\beta-\varepsilon)\sqrt{k} \leq L_{t_1}(s_j) \leq (\beta+\varepsilon)\sqrt{k} \tag{4.14}$$

(a.s.)

for all $s_j$ such that $|s_j| = k$ .

Hence (4.10), (4.13), and (4.14) imply:

$$\forall\ \varepsilon > 0 \qquad M_\varepsilon = \max\left\{N_\varepsilon, \frac{K_\varepsilon}{p-\varepsilon}\right\} : \forall\ n > M_\varepsilon$$

$$(1-\varepsilon)(\beta-\varepsilon)\ \sqrt{(p-\varepsilon)} < \frac{E\left[L_{t_1}(S)\right]}{\sqrt{n}} < (\beta+\varepsilon)\ \sqrt{p+\varepsilon} + b\varepsilon \qquad \text{(a.s.)} \qquad (4.15)$$

since $\varepsilon$ is arbitrarily small the theorem is proved.

Q.E.D.

Corollary 4.1

Under the condition of Theorem 4.1, we have:

$$\lim_{n\to\infty} \frac{E_u\left[E[L_{t_1}(S)]\right]}{\sqrt{n}} = \beta\ \sqrt{p}$$

Proof:

We have $0 < \dfrac{E[L_{t_1}(S)]}{\sqrt{n}} < \dfrac{EL_{t_p}\left(x^{(n)},p\right)}{\sqrt{n}}$ .

From Lemma 4.1, $\dfrac{EL_{t_p}\left(x^{(n)},p\right)}{\sqrt{n}}$ is bounded from above by a constant independent of n, so the almost sure convergence of Theorem 4.1 implies (by the dominated convergence theorem)

$$\lim_{n\to\infty} \frac{E_u\left[E[L_{t_1}(S)]\right]}{\sqrt{n}} = \beta\ \sqrt{p} \quad .$$

Let us now turn to the asymptotic behavior of the optimal PTSP tour and present one of the most important theoretical results of this chapter.

### 4.3.2 Expected Length of the Optimal PTSP Tour When $n \to \infty$

Theorem 4.2:

Let X be an infinite sequence of points independently and uniformly distributed over $[0,1]^2$ and p be the coverage probability for each point. Then there exists a constant c(p) such that:

$$\forall\ p \in [0,1] \quad \lim_{n \to \infty} \frac{EL_{t_p}\left(X^{(n)},p\right)}{\sqrt{n}} = c(p) \qquad \text{(a.s.)} .$$

Proof: We are going to use a very interesting recent result by Steele [1981a]. Steele proved the following:

Let $\phi$ be any real-valued function of the finite subsets of $R^2$ (i.e., a valuation, mapping finite sets to the reals) with the following properties:

(a) $\phi$ is Euclidean; i.e. linear and invariant under translation

(b) $\phi$ is monotone; i.e., $\phi\left(\{z\} \cup A\right) > \phi\left(A\right)$ for any $z \in R^2$ and finite subsets A of $R^2$

(c) $\phi$ has bounded variance, under the uniform distribution; i.e., $\mathrm{var}\left[\phi\left(X^{(n)}\right)\right] < \infty$ whenever the points of $X^{(n)}$ are independent and uniformly distributed in $[0,1]^2$

(d) $\phi$ is subadditive; i.e. if $\left\{Q_i\right\}_{i=1}^{m^2}$ is a partition of the unit square $[0,1]^2$ into squares with edges parallel to the axis and of length $1/m$ and if $rQ_i \equiv \{\eta : \eta = r\xi,\ \xi \in Q_i\}$, then there exists a constant $C > 0$ such that for all positive integers m and positive reals r one has:

$$\phi\left(X^{(n)} \quad [0,r]^2\right) < \sum_{i=1}^{m^2} \phi\left(X^{(n)} \cap rQ_i\right) + Crm .$$

Then, if X is a sequence of points independently and uniformly distributed over $[0,1]^2$, there exists a constant $\beta(\phi)$ such that

$$\lim_{n\to\infty} \frac{\phi\left(x^{(n)}\right)}{\sqrt{n}} = \beta(\phi) \qquad \text{(a.s.)}$$

(Steele proved this theorem for $R^d$, $d \geq 2$).

Since no confusion can arise throughout the proof of Theorem 4.2, let us simplify our notation:

Let $\phi_p\left(x^{(n)}\right) \equiv EL_{t_p}\left(x^{(n)},p\right)$; i.e., the expected length of the optimal PTSP tour going through $x^{(n)} = \{x_1,x_2,\ldots,x_n\}$ when each point is present only with probability p.

Let us show that $\phi_p$ verifies the axioms (a), (b), (c), (d): first note that, when p=1, $\phi_1\left(x^{(n)}\right)$ is nothing but the length of the optimal TSP tour through $x^{(n)}$ and we know that Theorem 4.2 is valid (Beardwood et al. [1959]); also, for p=0, $\phi_o\left(x^{(n)}\right) \equiv 0 \;\forall\; x$ hence Theorem 4.2 is trivially valid.

Let us then concentrate on $\phi_p$ with $0 < p < 1$.

- It is trivial to verify that

$$\phi_p\left(\alpha x^{(n)}\right) = \alpha\phi_p\left(x^{(n)}\right) \qquad \forall \text{ sequence of point } x, \;\; \forall \text{ real number } \alpha$$

where $\alpha x^{(n)}$ is the sequence of points $(\alpha x_1,\ldots,\alpha x_n)$; (i.e., each points of the sequence has its coordinates magnified by the real $\alpha$); indeed by changing the coordinates of each point by a factor $\alpha$, we multiply each distance between two points by $\alpha$ and since $\phi_p$ is function of those distances in a linear way we get the result.

Also $\phi_p\left(x^{(n)}+\xi\right) \equiv \phi_p\left(x^{(n)}\right) \quad \forall\; \xi \in R^2$ .

Hence (a) is satisfied and $\phi_p$ is Euclidean.

- It is also obvious that (b) is satisfied; indeed given we have the optimal PTSP tour through n points, its expected length will always be less than or equal to the expected length of the optimal PTSP tour through the union of the same n previous points with an additional one (by contradiction).

- Again (c) is obviously satisfied by $\phi_p$ for any p.

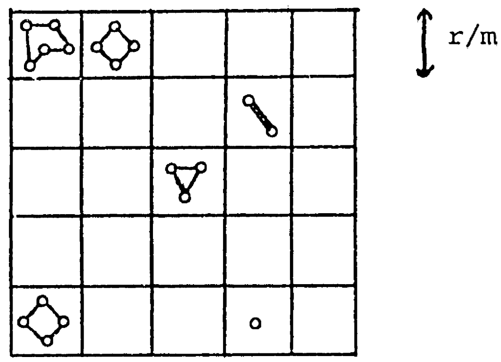
- For the verification of (d) we need a more elaborate analysis:

We have seen in the proof of Theorem 4.1 that as a consequence of Lemma 4.1 there exists a constant b such that $L_{t_1}(X^{(n)}) \leq b\sqrt{n}\, r$ for any sequence x within $[o,r]^2$ (in fact any constant $b \geq 3.5$ can do); hence we have

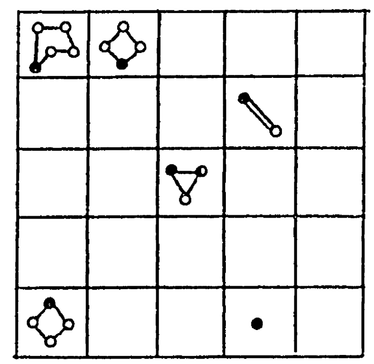
$$\forall\, n \quad \phi_1(x^{(n)}) \equiv L_{t_1}(x^{(n)}) \leq b\, \sqrt{n}\, r \quad . \tag{4.16}$$

Now consider the following tour through the sequence $x^{(n)} \cap [o,r]^2$ (see Figure 4.3 for an illustration): first construct the optimal PTSP tours in each "subsquare" through $x^{(n)} \cap r\, Q_i$, each of expected length $\phi_p(x^{(n)} \cap rQ_i)$ $(i \in [1..m^2])$. Then, in each square $rQ_i$ where $x^{(n)} \cap rQ_i$ is not the empty set, choose one point from $x^{(n)}$ $rQ_i$ and consider this point to be always present (we turned a "white" point into a "black" point, using the terminology of Chapters 2 and 3); if we let $\phi'_p(x^{(n)} \cap rQ_i)$ denote the new expected length of the PTSP tour initially constructed in $rQ_i$, we have:

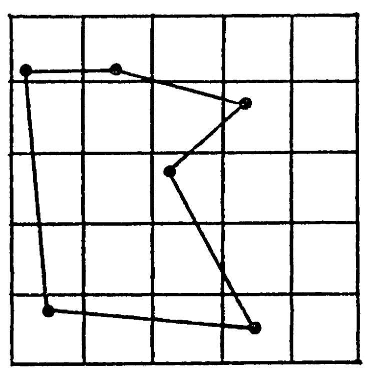
$\phi'_p(x^{(n)} \cap rQ_i) \geq \phi_p(x^{(n)} \cap rQ_i)$ (since now, one point always has to be visited)



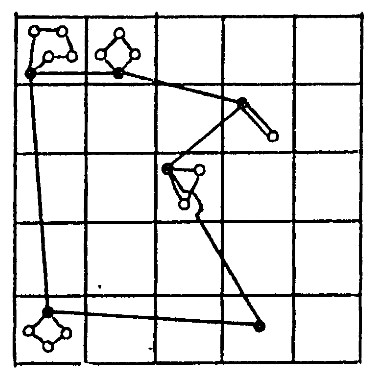

(1) The PTSP tours through $x^{(n)} \cap r\,Q_i$



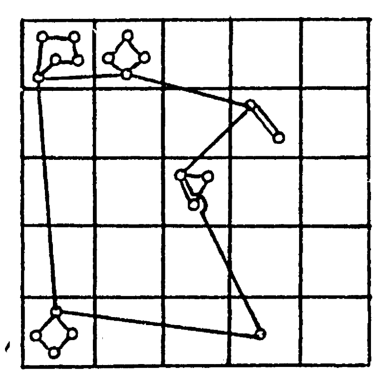
(2) turn one point in "black" in each subsquare



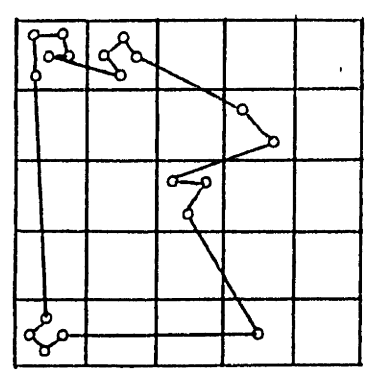
(3) construction of the TSP through black points



(4) combination of (2) and (3)



(5) A <u>spanning walk</u> of expected length less than in (4)



(6) A <u>tour</u> of expected length less than in (5)

<u>Figure 4.3</u>: the subadditivity of $\phi_p$

In each square $rQ_i$ where $x^{(n)} \cap rQ_i$ is not the empty set we have a black point so that we have at most a total of $m^2$ such points from $x^{(n)} \cap [o,r]^2$.

Construct a <u>TSP tour</u> through these points; from (4.16) with $n = m^2$ we have:

$$\phi_1\big(\text{black points}\big) < b\sqrt{m^2}\, r = bmr \quad . \tag{4.17}$$

Now the expected length of the combination of the TSP and the PTSP tours is (since we have black points, the expectation of the combination is the sum of the expectations of each part):

$$\sum_{i=1}^{m^2} \phi'_p\big(x^{(n)} \cap rQ_i\big) + \phi_1\big(\text{black points}\big)$$

(and this corresponds to a spanning walk).

Now if we turn each black point back into a white point, the expected length of the spanning walk decreases, and so does it if we transform this spanning walk into a tour; but then this tour goes through all points of $x^{(n)} \cap [o,r]^2$, hence its expected length cannot be smaller than $\phi_p\big(x^{(n)} \cap [o,r]^2\big)$ (Optimum PTSP). Hence we have

$$\phi_p\big(x^{(n)} \cap [o,r]^2\big) < \sum_{i=1}^{m^2} \phi'_p\big(x^{(n)} \cap rQ_i\big) + \phi_1\big(\text{black points}\big)$$

and using (4.17)

$$\phi_p\big(x^{(n)} \cap [o,r]^2\big) < \sum_{i=1}^{m^2} \phi'_p\big(x^{(n)} \cap rQ_i\big) + bmr \quad . \tag{4.18}$$

We need a last result: assume that the optimal PTSP through the points of $x^{(n)}$ $rQ_i$ is $(x_1, x_2, \ldots, x_{k_i}, x_1)$ where $k_i = |x^{(n)} \cap rQ_i|$ (assume of course that $k_i \neq 0$). Now consider $x_1$ to be always present.

From Lemma 3.8 part b we have:

$$\phi_p'\left(x^{(n)} \cap rQ_i\right) \leq \phi_p\left(x^{(n)} \cap rQ_i\right) + 2d^*(1-p) \qquad (4.19)$$

where $d^* = \max\limits_{1<j\leq k_j} d(x_1, x_j)$.

Since $rQ_i$ has an edge of $r/m$

$$d^* \leq \sqrt{2}\,(r/m) \qquad \text{(diagonal of } rQ_i\text{)}$$

so finally

$$\phi_p'\left(x^{(n)} \cap rQ_i\right) \leq \phi_p\left(x^{(n)} \cap rQ_i\right) + 2\sqrt{2}\,\frac{r}{m}\,(1-p) \quad . \qquad (4.20)$$

This is true for any $i \in [1..m^2]$, so combining (4.18) and (4.20) we obtain:

$$\phi_p\left(x^{(n)} \cap [o,r]^2\right) \leq \sum_{i=1}^{m^2} \phi_p\left(x^{(n)} \cap rQ_i\right) + \left((1-p)2\sqrt{2}\,\frac{r}{m}\right)m^2 + bmr \qquad (4.21)$$

so by choosing $C = (1-p)\,2\sqrt{2} + b$ in (4.21) the subadditivity requirement is verified.

Q.E.D.

For all similar asymptotic results (the TSP, the matching problem, the spanning tree problem) it is interesting to note that the respective limiting constants are unknown and only bounds have been established:

- the best bounds for the TSP ($\equiv\beta$) are found in Beardwood et al. [1959] $0.625 \leq \beta \leq 0.9204$ and $\beta$ has been estimated to be $\approx 0.765$ (Stein [1977]) (it was previously estimated to be 0.75 in Beardwood et al. [1959])
- for the matching problem Papadimitriou [1978] has established that $0.25 \leq \mu \leq 0.401$ and estimated $\mu$ to be $\approx 0.35$

- for the minimum spanning tree problem Gilbert [1965] showed that $0.5 \leq \beta_{MST} \leq 0.707$ and estimated $\beta_{MST}$ to be $\approx 0.68$ (in Bentley and Saxe [1980] $\beta_{MST}$ has been estimated to be 0.66)

Our problem is no exception and the next section will be concerned with bounds on c(p).

First, however, we present a corollary to Theorem 4.2 similar to the one given for Theorem 4.1.

<u>Corollary</u> 4.2

Under the condition of Theorem 4.2 we have:

$$\lim_{n\to\infty} \frac{E_u\left[EL_{t_p}\left(x^{(n)},p\right)\right]}{\sqrt{n}} = c(p)$$

<u>Proof</u>: Identical to the one for Corollary 4.1.

### 4.3.3 <u>Bounds on the Constant c(p)</u>

The lemmas of section 4.2 give immediately bounds for c(p):

- From Lemma 4.1 we have $\limsup_{n\to\infty} \dfrac{EL_{t_p}\left(x^{(n)},p\right)}{\sqrt{n}} \leq \sqrt{2p}$ for an arbitrary sequence x lying in $[0,1]^2$; together with Theorem 4.2 this result implies that $c(p) \leq \sqrt{2p}$ . (4.22)

- From Lemma 4.2 and Corollary 4.2 we can improve this upper bound; indeed we got $\limsup_{n\to\infty} \dfrac{E_u\left[EL_{t_p}\left(x^{(n)},p\right)\right]}{\sqrt{n}} \leq \sqrt{(4/3)\ p}$ which implies (Corollary 4.2) that $c(p) \leq \sqrt{(4/3)\ p}$ . (4.23)

- From Lemma 4.3 and Corollary 4.2 we get (see end of proof of Lemma 4.3)

  Lemma 4.3: => $\liminf_{n\to\infty} \dfrac{E_u\left[EL_{t_p}\left(x^{(n)},p\right)\right]}{\sqrt{n}} \geq \frac{5}{8}\ p$ .

  Corollary 4.2: => $c(p) \geq \frac{5}{8}\ p$ . (4.24)

In fact during the proof of Lemma 4.3 we got a sharper bound when n is very big; indeed we showed that

$$E_u\left[EL_{t_p}\left(X(n),p\right)\right] \geq \frac{15}{24}\left[E[\sqrt{W}] - n(1-p)^{n-1}\right] \tag{4.25}$$

(for X sequence of points independently and uniformly distributed within $[0,1]^2$);

By Weak Law of Large Numbers, $\frac{W}{n} \to p$ in probability

and since $g(x) = \sqrt{x}$ is continuous $\sqrt{W/n} \to \sqrt{p}$ in probability. (4.26)

Since $\{\sqrt{W/n}\}$ is a sequence of uniformly bounded random variables, convergence in probability implies convergence in quadratic mean (Lukacs [1968]) which in turn implies that:

$$\lim E\left[\sqrt{W/n}\right] = \sqrt{p} \quad . \tag{4.27}$$

Hence (4) (6) and Corollary 4.2 imply that: $c(p) \geq \frac{5}{8}\sqrt{p}$ . (4.28)

Finally, with the best of (4.22), (4.23), (4.24) and (4.28), we obtain:

$\forall\ 0 \leq p \leq 1$

$$0.625\sqrt{p} = \frac{5}{8}\sqrt{p} \leq c(p) \leq \sqrt{4/3}\sqrt{p} \simeq 1.155\sqrt{p} \quad . \tag{4.29}$$

This result was obtained directly from section 4.2; let us now present the best bounds for c(p):

Lemma 4.4:

The constant c(p) of the asymptotic result given by Theorem 4.2 is bounded as follows:

$$\beta\sqrt{p} \leq c(p) \leq \min\{\beta, 0.9204\sqrt{p}\} \qquad \forall\, p \in [0,1]$$

where $\beta$ is the "TSP constant".

Proof:

- since the variables $\frac{E[L_{t_1}(S)]}{\sqrt{n}}$ are lower bounds on $\frac{EL_{t_p}(X^{(n)},p)}{\sqrt{n}}$

for any n (see proof of Lemma 4.3), Theorem 4.1 implies $c(p) \geq \beta\sqrt{p}$. [Note that the best known lower bound on $\beta$ is 0.625 and thus this gives the same result as in (4.29); however $\beta\sqrt{p}$ is certainly a better lower bound than in (4.29) since any improvement on $\beta$ would directly be of interest for c(p)].

- it remains to verify the upper bounds: the construction given in Lemma 4.2 was used in the proof of Lemma 10 of Beardwood et al. [1959] from which one can infer the following property: (using our notation)

$$\limsup_{n\to\infty} \frac{E_u[L_{t_{c3}}(X^{(n)})]}{\sqrt{n}} \leq 0.9204 \qquad (4.30)$$

where X is a sequence of points independently and uniformly distributed over $[0,1]^2$

This result comes from two major parts in Beardwood et al. (the interested reader is referred to Beardwood et al. [1959], Lemma 5 and Lemma 10 for a precise statement of the proof):

(1) Consider a Poisson point process $\pi$ in $R^2$ with uniform intensity parameter 1; call $\pi([o,r]^2)$ the random set of points in $[o,r]^2$ and let

$t_1\left(\pi\left([o,r]^2\right)\right)$ be the TSP tour through $\pi\left([o,r]^2\right)$, then the TSP constant $\beta$ is the limiting value (as $r \to \infty$) of the expected distance between successive points on $t_1(\pi\left([o,r]^2\right)$.

(2) Using the same construction as in Lemma 4.2, we obtain $t_{c_3}\left(\pi\left([o,r]^2\right)\right)$ (here we divide the square into r/m equal-width (m) columns and we keep m fixed as $r \to \infty$); by showing that the total expected length of the "within-column" distances and the expected length of the return link from the end of the last column to the beginning of the first contribute a negligible amount to the total path length of $t_{c_3}\left(\pi\left([o,r]^2\right)\right)$ as $r \to \infty$, we obtain (4.30) by computing $\lambda(m)$ - the expected distance between two successive points (drawn from $\pi$) in an infinite column of side m - and then by choosing the best value m*. (The obtention of 0.9204 resorts to numerical quadrature and special function calculations).

Now as mentioned in our proof of Lemma 4.2 (properties (a) and (b)') we have:

$$E_u\left[EL_{t_{c3}}\left(X(n),p\right)\right] \equiv \sum_{k=0}^{n} E_u\left[L_{t_{c3}}(k)\right] \Pr(W=k) \tag{4.31}$$

By using (4.30), (4.31), and the fact that

$\frac{W}{n} \to p$ (a.s.) (Strong Law of Large Number), one can prove (using the same techniques as in the proof of Theorem 4.1) that:

$$\lim_{n\to\infty} \sup \frac{E_u\left[L_{t_{c3}}\left(X(n),p\right)\right]}{\sqrt{n}} \leq 0.9204 \sqrt{p} \quad .$$

This implies that $c(p) \leq 0.9204 \sqrt{p}$.

To obtain the other upper bound it suffices to note that:

$$E_u[EL_{t_p}(x^{(n)},p)] \leq E_u[E[L_{t_1}(S)]] \leq E_u[L_{t_1}(x^{(n)})]$$

and that $\lim_{n\to\infty} \frac{E_u[L_{t_1}(x(n))]}{\sqrt{n}} = \beta$

from Beardwood et al. [1959].

Q.E.D.

This completes section 4.3. In the next section we shall present generalizations of the results developed in sections 4.2, 4.3: in the first part we present similar results to the slightly different problems where one (or more) of the points must always be visited (depot), and to the PHPP problem (see Chapter 2). We then show in a second subsection that the results of section 4.2 and 4.3 can be extended to the cases of bounded Lebesgue measurable sets in d-dimensional Euclidean space $R^d$. Finally we mention the use of other metrics.

### 4.3.4 Generalizations

A. Extensions to problems with n white points and m black points: So far we have been concerned with sequence (x) of points which are present only with a probability p, independently of each other; let us generalize our result to the case where some of the points always require a visit. To do so, we introduce another sequence of points y of "always present" (black) points (x remains a sequence of "white" points); we will consider problem instances with n white points $(x^{(n)})$ and m black points $(y^{(m)})$; the union of the n+m points will be written $x^{(n)} \cup y^{(m)}$; we will assume m to be a non-decreasing function of n (possibly a constant) mapping the set of natural integers into itself.

In a first part we will present straightforward extensions of sections 4.2 and 4.3 to the case of n white points and 1 black point (a depot or a customer). Then in a second part we will mention similar extensions of our finite size results to the case of m black points and we will present generalizations of Theorem 4.1 and Theorem 4.2.

(1) Case of n white points and one black point:

It is very easy to extend lemmas 4.1 - 4.3 to this case:

- We obtain the following bounds for $EL_{t_p}\left(x^{(n)} \cup y^{(1)}, p\right)$:

Lemma 4.1 =>

$$\left(\sqrt{2(np-1)} + \frac{13}{4}\right) r \qquad \text{if } np \geq 1.5$$

$$\left(\frac{np}{2} + \frac{7}{2}\right) r \qquad \text{if } np \leq 1.5 \quad .$$

- Under the condition of Lemma 4.2 $E_u\left[EL_{t_p}\left(x^{(n)} \cup y^{(1)}, p\right)\right]$ is bounded from above by:

$$\left(\sqrt{(4/3)(np-3)} + \frac{11}{2} + 3\sqrt{2}\right) r \qquad \text{if } np \geq 3.75$$

$$\left(\frac{np}{3} + \frac{2}{3} + 3\sqrt{2}\right) r \qquad \text{if } np \leq 3.75 \quad .$$

- Lemma 4.3 gives the same lower bound.

For the asymptotic results of section 4.3 it is easy to check that Theorem 4.1 and Theorem 4.2 not only remain valid, but lead (respectively) to the same asymptotic constant; this can best be seen from:

$$\frac{\sqrt{n}}{\sqrt{n+1}} \frac{E\left[L_{t_1}(S)\right]}{\sqrt{n}} \leq \frac{E\left[L_{t_1}(S \cup y(1))\right]}{\sqrt{n+1}} \leq \frac{\sqrt{n}}{\sqrt{n+1}} \left(\frac{E\left[L_{t_1}(S)\right]}{\sqrt{n}} + \frac{2\sqrt{2}}{\sqrt{n}}\right)$$

(which also holds for $EL_{t_p}\left(x^{(n)} \cup y^{(1)}, p\right)$)

together with Theorem 4.1 (Theorem 4.2) and with $\lim_{n\to\infty} \frac{2\sqrt{2}}{\sqrt{n}} = 0$.

Before turning to the general case of m black points one may remark that we assumed the position of $y^{(1)}$ to be fixed (i.e. a depot); we can obtain the same extensions if the position is random (we even get slightly better bounds for Lemma 4.2 and Lemma 4.3).

(2) Case of n white points and m black points:

Since it is straightforward to extend the lemmas of section 4.2 to this case, we will only be concerned with the extensions of Theorem 4.1 and Theorem 4.2:

Theorem 4.3: (generalization of Theorem 4.1)

Let X and Y be two infinite sequences of points independently and uniformly distributed over $[0,1]^2$, each point of X being present only with a probability p, each point of Y being always present, then we have:

$$\lim_{n\to\infty} \frac{E\left[L_{t_1}\left(S \cup Y^{(m)}\right)\right]}{\sqrt{n+m}} = \begin{cases} \beta\sqrt{p} & \text{if } m = o(n) \\ \beta & \text{if } n = o(m) \\ \beta\sqrt{(p+\xi)/(1+\xi)} & \text{if } \lim_{n\to\infty} \frac{m}{n} = \xi \end{cases}$$

with probability 1 ((a.s.)) .

Proof: Using the same argument as in the proof of Theorem 4.1, one can show that:

$$\forall\ \varepsilon > 0 \quad N_\varepsilon : \forall\ n > N_\varepsilon$$

$$(\beta-\varepsilon)(1-\varepsilon)\sqrt{\frac{n}{n+m}(p-\varepsilon)+\frac{m}{n+m}} < \frac{E\left[L_{t_1}\left(S \cup Y^{(m)}\right)\right]}{\sqrt{n+m}}$$

$$< (\beta+\varepsilon)\sqrt{\frac{n}{n+m}(p-\varepsilon)+\frac{m}{n+m}} + b'_\varepsilon \quad . \qquad \text{(a.s.)}$$

Since $\varepsilon > 0$ is arbitrarily small, the three cases of Theorem 4.3 can be obtained.

Q.E.D.

Theorem 4.4: (generalization of Theorem 4.2)

Let X and Y be two infinite sequences of points independently and uniformly distributed over $[0,1]^2$, each point of X being present only with a probability p, each point of Y being always present, then we have:

$$\lim_{n\to\infty} \frac{EL_{t_p}\left(X^{(n)} \cup Y^{(m)}, p\right)}{\sqrt{n+m}} = \begin{cases} c(p) & \text{if } m = o(n) \\ \beta & \text{if } n = o(m) \\ c'(p,\xi) & \text{if } \lim\limits_{n\to\infty} \frac{m}{n} = \xi \end{cases}$$

with probability 1 $\big((a.s.)\big)$.

Proof: Let us consider the three cases separately:

(i) $m = o(n)$:

consider $t_p\left(X^{(n)}\right)$; by assuming one of the points of $X^{(n)}$ to be always present, one can then show (following an argument presented in the proof of the subadditivity of $\phi_p$ in Theorem 4.2) that

$$EL_{t_p}\left(X^{(n)} \cup Y^{(m)}, p\right) \leq EL_{t_p}\left(X^{(n)}, p\right) + 2\sqrt{2} + b\sqrt{m+1} \tag{4.32}$$

(where b is the same constant as presented in the proof of Theorem 4.2).

Also $EL_{t_p}\left(X^{(n)} \cup Y^{(m)}, p\right) \geq EL_{t_p}\left(X^{(n)}, p\right)$. (4.33)

The first part of Theorem 4.4 is proved by using Theorem 4.1, the fact that $m=o(n)$, and the following inequality from (4.32) and (4.33)

$$\sqrt{\frac{n}{n+m}}\ \left[\frac{EL_{t_p}(X^{(n)},p)}{\sqrt{n}}\right] \leq \frac{EL_{t_p}(X^{(n)} \cup Y^{(m)},p)}{\sqrt{n+m}} \leq \sqrt{\frac{n}{n+m}}\ \left[\frac{EL_{t_p}(X^{(n)},p)}{\sqrt{n}}\right.$$

$$\left. + \frac{2\sqrt{2}}{\sqrt{n}} + \frac{b\sqrt{m+1}}{\sqrt{n}}\right] .$$

(ii) if $n = o(m)$:

the result of Theorem 4.4 is obtained directly from Theorem 4.3 and from the classical result of Beardwood et al. [1959] after noticing that:

$$\frac{E[L_{t_1}(S \cup Y^{(m)})]}{\sqrt{n+m}} \leq \frac{EL_{t_p}(X^{(n)} \cup Y^{(m)},p)}{\sqrt{n+m}} \leq \frac{L_{t_1}(X^{(n)} \cup Y^{(m)})}{\sqrt{n+m}} .$$

(iii) if $\lim\limits_{n\to\infty} \frac{m}{n} = \xi$:

for this last case one can proceed exactly as in the proof of Theorem 4.2 to show the existence of a constant $c'(p,\xi)$.

Q.E.D.

Using the same techniques as in the proof of Lemma 4.4, one can easily show that

$$\beta\sqrt{(p+\xi)/(1+\xi)} \leq c'(p,\xi) \leq \min\left\{\beta,\ 0.9204\ \sqrt{(p+\xi)/1+\xi)}\right\} . \tag{4.34}$$

Finally one should note that all the previous results are also valid (with the same asymptotic constant) for the PHPP (as defined in Chapter 2) since for any set of points K in $[0,1]^2$ (K can be any combination of black and white points) one has (with the obvious extension of notation):

$$\frac{E\left[L_{t_1}(K)\right]}{\sqrt{|K|}} - \frac{\sqrt{2}}{\sqrt{|K|}} \leq \frac{E\left[L_{h_1}(K)\right]}{\sqrt{|K|}} \leq \frac{E\left[L_{t_1}(K)\right]}{\sqrt{|K|}}$$

$$\frac{EL_{t_p}(K,p)}{\sqrt{|K|}} - \frac{\sqrt{2}}{\sqrt{|K|}} \leq \frac{EL_{h_p}(K,p)}{\sqrt{|K|}} \leq \frac{EL_{t_p}(K,p)}{\sqrt{|K|}} \quad . \tag{4.35}$$

B. <u>d-dimensional Euclidean Space $R^d$ and Lebesgue measurable sets;</u>

We will be concerned with generalizations as $n \to \infty$. So far we presented all our results for sequences of points lying in a square $[o,r]^2$ of a 2-dimensional Eudlidean Space $R^2$

(i) In fact all our results can be extended to the case of sequence of points x, lying in a d-cube $[0,1]^d$ of a d-dimensional Euclidean Space $R^d$: for example

- Lemma 4.1 gives: (for $d \geq 3$)

$$\text{as } n \to \infty \quad EL_{t_p}\left(x^{(n)},p\right) \leq \frac{\sqrt{d}}{\left(2(d-1)\right)^{(d-1)/2d}} \left(E[W]\right)^{\frac{d-1}{d}} + o\left(E[W]^{\frac{d-2}{d}}\right) \tag{4.36}$$

(the proof is a generalization of the method of section 4.1 and is based on a construction that was suggested in Few [1955] for the TSP).

- Theorem 4.1 gives: (using a similar extension in Beardwood et al. [1959])

$$\lim_{n \to \infty} \frac{E\left[L_{t_1}(S)\right]}{n^{(d-1)/d}} = \beta(d) \; (p)^{\frac{d-1}{d}} \qquad \text{(a.s.)} \quad . \tag{4.37}$$

- Theorem 4.2 remains valid; indeed using the general result for Lemma 4.1 (given above) and the generalization for d-dimensions of Steele's result, the proof of Theorem 4.2 remains unchanged and we get:

$$\lim_{n\to\infty} \frac{EL_{t_p}\left(X^{(n)},p\right)}{n^{(d-1)/d}} = c_d(p) \qquad \text{(a.s.).} \tag{4.38}$$

One can also generalize Lemma 4.4 and the bounds on $c_d(p)$. We obtain:

$\forall\, p \in [0,1]$, $\forall\, d \geq 3$:

$$\beta(d)\, p^{\frac{d-1}{d}} \leq c_d(p) \leq \left(\frac{d}{6}\right)^{\frac{1}{2}}\, 12^{\frac{1}{2d}}\, p^{\frac{d-1}{d}} \tag{4.39}$$

where $\left(\frac{d}{6}\right)^{\frac{1}{2}} \cdot 12^{\frac{1}{2d}}$ is obtained through a generalization of the construction of $t_{c_3}$ to a d-dimensional space. (all results of 4.3.4 A. can also be generalized)

(ii) It is also true that the extensions given in (i) remain valid when the sequence of points lies in any bounded Lebesgue measurable set A with Lebesgue measure (≃ d-volume) v(A) provided we multiply each right hand side by $v(A)^{1/d}$; this is true for Theorem 4.1 due to a similar extension in Beardwood et al. [1959]; for Lemma 4.1 it comes from a general result from Mathematical Analysis (see for example von Neumann [1950]) saying that we can cover A with a finite number (say $m(\varepsilon)$) of disjoint d-cubes $[o,r_i]^d$ such that $\sum_{i=1}^{m(\varepsilon)} v\left([o,r_i]^d\right) \leq v(A) + \varepsilon$. This fact together with the subadditivity of the functional (see proof of Theorem 4.2) imply the claim; then for Theorem 4.2 it follows from the previous extension of Lemma 4.1 and the result from Steele [1981a]. (Again similar extensions are valid for the results of section 4.3.4 A.)

c. Other metrics:

It is interesting to note that all the asymptotic results of this chapter are not restricted to the Euclidean metric; all results remain valid (of course with different asymptotic constant) if one chooses the $L_\infty$ metric (rectilinear metric) or more generally the $L_k$ metric ($k \geq 1$)

$$\left(\text{i.e., } \|x-y\|_k = \left[ \sum_{i=1}^{d} (x_i-y_i)^k \right]^{1/k} \text{ where } d \text{ is the dimension of the space considered}\right).$$

4.4 Conclusion:

In addition to its theoretical interest, Theorem 4.2 should prove to be extremely useful in assessing the "goodness" of heuristics for the PTSP through a probabilistic analysis (in the same fashion as the result of Beardwood et al. [1959] for the TSP); besides this algorithmic application, Theorem 4.2, together with an estimation $\tilde{c}(p)$ of the constant $c(p)$, provides an approximation formula useful to predict with high probability the expected length of an optimal PTSP tour, if the number of points is large.

Also the approximation formulas based on Theorem 4.1 and Theorem 4.2 (assuming an estimation $\tilde{c}(p)$) provide an important practical "by-product" (take $\beta = 0.765$):

$\left(\tilde{c}(p) - 0.765 \sqrt{p}\right) \sqrt{na}$ represents an estimation of the penalty one has to pay when n customers have to be served within a region R of area a (each of them present only with some fixed probability p) and when the route is not reoptimized for each realization of the problem.

(Note that, from Theorem 4.3 and Theorem 4.4, the two strategies - "PTSP or reoptimization" - are asymptotically equivalent when $n = o(m)$, i.e. when the number of white points is negligible compared to the number of black points).

# CHAPTER 5

# ALGORITHMIC INVESTIGATION

## 5.1 Introduction

### 5.1.1 Focus of Chapter 5

Chapters 2, 3, and 4 have been mainly concerned with theoretical aspects of the PTSP; in Chapter 2 we derived closed-form expressions that efficiently give the expected length (in the PTSP sense) of a given tour of a graph G under various conditions; then, partly based on Chapter 2, we derived, in Chapter 3, general combinatorial properties of the PTSP; finally we presented, in chapter 4, several bounds and an asymptotic analysis for the PTSP in the plane - one of the main theoretical results being the derivation of strong limit laws for the PTSP (i.e. concerned with almost sure convergence).

It is now time to use this extensive theoretical investigation to develop solution procedures for solving the PTSP; this will be the main concern of Chapter 5.

Our emphasis is on the conceptualization of solution procedures built upon the understanding of the problem that has been obtained through the previous chapters. As such, all the proposed algorithms are based on a theoretical foundation and all contain some kind of rationale behind their design. We are, however, not concerned with the practical implementation of those algorithms here (i.e., with computer codes), neither with experimental testing of their validity and relative merits; we will provide, on the other hand, discussions (whenever possible) of those procedures based on the theoretical results of the previous chapters and on some additional theoretical results presented during the development in Chapter 5.

In conclusion, our aim, here, is principally methodological and is motivated by the desire to apply the analytical developments in order to present diverse solution strategies.

Finally, one should note at the outset that the PTSP is certainly "harder" than its deterministic counterpart (TSP) which itself belongs to the notorious class of NP-hard problems (in fact it is very easy to show formally that the recognition version of the PTSP is NP-complete; see Garey and Johnson [1979] for a good introduction to the NP-complete notion). There is strong evidence to suggest that problems from this class of difficult combinatorial problems cannot be solved optimally with an algorithm which is guaranteed to run in polynomial time. Our approach to the large-scale PTSP will then be to seek fast heuristic algorithms with polynomial time bounds. Nevertheless, for problems of smaller size we will attempt to give some guidance on the design of exact procedures.

### 5.1.2 Contents of Chapter 5

First, we briefly provide, at the end of this subsection, some clarifications on the notation adopted throughout the development of this chapter; this is then followed by the main sections; our results are divided into two parts: in section 5.2 we are concerned with exact optimization methods for solving the PTSP; we first show how one can formulate the PTSP (for each case, $m=0$, $m=1$, $m \geq 2$) as an integer nonlinear programming problem (with a nonlinear objective function - a polynomial of order n, the number of white nodes - and linear constraints); then we successively transform this formulation first, to a mixed integer linear program, and, finally, to a pure integer linear program. In a third subsection, we discuss the relative merit of these three formulations and propose a Branch-and-Bound procedure (believed to be one of the best exact methods for tackling this specific problem) built upon similar procedures for the traditional TSP. We conclude section 5.2 by showing that the relationship between the PTSP and TSP is not as simple as one could imagine based on the first part of that section. Indeed, we will show that a seemingly natural extension of the dynamic programming formulation of the TSP does not solve the PTSP, and that, in fact, one cannot use dynamic programming approaches to provide an exact solution procedure for this problem. The second main section (5.3) contains an exposition of heuristic procedures (that is, not guaranteed to obtain an optimal solution) after providing a brief discussion on the necessity of developing such procedures for a complex problem like the PTSP, we first present some theoretical preliminaries on which some of the proposed procedures will be directly built. (For example, we introduce easier problems for which we derive worst case ratios between their optimal solutions and the optimal PTSP solution). Based on those preliminaries, we present a host of procedures under the generic term of Tour construction procedures (a term originally used for the TSP); we present an extension of the Clarke and Wright savings approach (see Golden et al. [1980], for example) and label it the Supersaving Algorithm; we also introduce the "Almost Nearest Neighbor" Algorithm and finally conclude this section on Tour construction procedures by listing several "insertion" procedures. In a second subsection, 5.3.3, we briefly mention the use of "hill-climbing" methods (similar to the "2-opt" or "3-opt" heuristics proposed for the TSP). Finally, in a third subsection we turn our attention to the case of the PTSP in the plane. Based mainly on results from Chapter 4, this section will analyze a recent heuristic for the TSP based on spacefilling curves (see Platzman and Bartholdi [1983]) and we will also look at procedures based on partitionning approaches (see Karp [1977]).

We conclude Chapter 5 with a review of the most interesting results and most promising approaches proposed to solve the PTSP.

Before presenting our results, let us first make some additional comments on the notation and conventions used throughout this chapter. All notation introduced in Chapters 2, 3, and 4 is valid and will be used hereafter; more specifically, with the exception of the section concerning the PTSP in the plane (5.3.4) which is based on Chapter 4, all other sections are based on the graph-theoretic concepts of Chapters 2 and 3. Most of our analysis is conducted assuming a general p.m.f. for W (the few results concerned with the binomial case will be explicitly mentioned) and with only one exception, we will be mainly concerned with the <u>variation</u> of the PTSP in which we have <u>one black</u> node and <u>n white</u> nodes (the extension to $m=0$, or $m \geq 2$ being readily obtainable, see discussion at the beginning of section 5.2.3). We recall that the distance matrix D is assumed to satisfy the <u>triangular inequality</u>; in general D is assumed to be asymmetric except in some cases for which a symmetric D will be assumed explicitly. By convention, we will assume (especially for the formulations given in 5.2) that $d(i,i) = +\infty$ for each node of G.

## 5.2 Exact Optimization Methods

As mentioned in the introduction the emphasis in this section is on the discussion of optimization-based methods for solving the PTSP (as opposed to heuristic procedures which are the subject of section 5.3).

Almost indispensable for such a goal is a mathematical representation of the problem.

A very concise formulation of the problem for a graph $G=(N,A,D)$ with m black nodes and n white nodes is simply given by:

$$\min_{t \in T} \left\{ E[L_t] \equiv \sum_{r=0}^{n} \alpha_r L_{m,t}^{(r)} \right\} \tag{5.1}$$

where T is the set of all hamiltonian circuits or tours of the graph G, and where the other terms of (5.1) were defined in Chapters 2 and 3; although very convenient to represent the problem, (5.1) does not suggest algorithmic strategies. Our first concern is then to formulate the problem as a <u>mathematical programming problem</u>; that is the minimization of an objective function subject to constraints. Here it is apparent that the constraints are expressed in (5.1) under the form $t \varepsilon T$, hence the same constraints as for the TSP (namely, the optimal solution has to be a tour). This immediately suggests an assignment-based formulation of the problem based on the introduction of the same decision variables as for the TSP; that is let:

$$x_{ij} = \begin{cases} 1 & \text{if arc}(i,j) \text{ is in the tour } t \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

The familiar formulation of the TSP (see for example Golden and Magnanti [1980]) based on those decision variables is given by (assuming $d(i,i) = +\infty$ for any node of G, and $|N|=n+m$ to be consistent with our notation):

$$\text{Minimize} \quad \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} d(i,j)x_{ij}$$

$$\text{subject to:} \quad \sum_{i=1}^{n+m} x_{ij} = 1 \qquad (j \in [1..n+m])$$

$$\sum_{j=1}^{n+m} x_{ij} = 1 \qquad (i \in [1..n+m]) \qquad (5.3)$$

$$X = (x_{ij}) \in S$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i,j \in [1..n+m])$$

where the first two sets of constraints represent the traditional constraints of an assignment problem and merely ensure, for this problem, that each node of G has one and only one arc coming into it and one and only one arc going out of it; S represents any restrictions that prohibits subtour solutions satisfying the other constraint. For example S can be defined as:

$$S = \left\{(x_{ij}) : \sum_{i \in Q} \sum_{j \in N-Q} x_{ij} \geq 1 \text{ for every nonempty } Q \subset N\right\} \qquad (5.4)$$

This is of course not the only way to express the set S; indeed, alternative forms for expressing S frequently constitute the major difference among various algorithms (see for example Parker and Rardin [1983]).

Now that the constraints $t \in T$ of (5.1) have been expressed in function of $x_{ij}$ (see formulation (5.3)), it remains to express $E[L_t]$ as a function of $x_{ij}$ to obtain a mathematical programming formulation of the PTSP. This will be the main concern of our first subsection.

### 5.2.1 Integer nonlinear programming formulation of the PTSP

We will distinguish the case where G has no black node (i.e., m=0) from the cases where G has at least one black node (i.e., m≥1)

A. Case of a graph G with no black node: let us state our main result:

**Lemma 5.1**

Let G = (N,A,D) be a graph with no black node and with n white nodes, let t be a given tour of G, and let $x_{ij}$ be the 0-1 decision variables defined in (5.2). Then one can express the expected length $E[L_t]$ of the tour t as follows:

$$E[L_t] = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=0}^{n-2} \alpha_r \, d(i,j) \, x_{ij}^{(r)}$$

where:

- $x_{ij}^{(o)} \equiv x_{ij}$

- $x_{ij}^{(r)} \equiv \sum_{k_1,\dots,k_r=1}^{n} x_{ik_1} x_{k_1k_2} \cdots x_{k_rj} \quad \forall\, r \in [1..n-2]$

Proof:

From Chapter 3 we have $E[L_t] = \sum_{r=0}^{n-2} \alpha_r L_{0,t}^{(r)}$ where $L_{0,t}^{(r)}$'s have been defined in (2.3). We noted that $L_{0,t}^{(o)}$ was simply the length of the tour t; hence from (5.3) we have:

$$L_{0,t}^{(o)} \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} d(i,j) \, x_{ij} \qquad (5.5)$$

For $r \in [1..n-2]$ define the variables $x_{ij}^{(r)}$ as:

$$x_{ij}^{(r)} = \begin{cases} 1 & \text{if } d(i,j) \text{ is part of } L_{0,t}^{(r)} \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

Then we have:

$$L_{0,t}^{(r)} \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} d(i,j)\ x_{ij}^{(r)} \tag{5.7}$$

But now from the construction of the $L_{0,t}^{(r)}$ we have the following equivalences (see Chapter 2):

$\{d(i,j)$ is part of $L_{0,t}^{(r)}\}$

<=> {there exist r distinct nodes $k_1,..,k_r$ such that the path $(i,k_1,..,k_r,j)$ belongs to the tour t}

<=> {there exist r distinct nodes $k_1,..,k_r$ such that $x_{ik_1}=1$, $x_{k_1k_2}=1$, ..., <u>and</u> $x_{k_rj}=1$}

so finally (5.6) can be expressed as:

$$x_{ij}^{(r)} = \begin{cases} 1 & \text{if there exists } k_1,..,k_r \text{ such that } x_{ik_1}=1,.., \text{ and } x_{k_rj}=1 \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$

It is then a matter of "boolean logic" to express $x_{ij}^{(r)}$ in function of $x_{ik_1}$, .., $x_{k_rj}$ (the expression "and" corresponding to a multiplication, and "or" corresponding to an addition):

$$x_{ij}^{(r)} = \sum_{k_1,..,k_r=1}^{n} x_{ik_1}\ x_{k_1k_2} \cdots x_{k_rj} \tag{5.9}$$

Using (5.7) and (5.9) we obtain the desired result.

Q.E.D.

Note:

(5.9) allows us to readily verify Fact 3.2 (i) and Fact 3.3 (iii); for example for Fact 3.3 (iii) we have:

$$\sum_{j=1}^{n} x_{ij}^{(r)} = \sum_{j=1}^{n} \sum_{k_1,\ldots,k_r=1}^{n} x_{ik_1} \cdot\cdot x_{k_r j}$$

$$= \sum_{k_1,\ldots,k_r=1}^{n} x_{ik_1} \cdot\cdot x_{k_{r-1}k_r} \sum_{j=1}^{n} x_{k_r j}$$

$$= \sum_{k_1,\ldots,k_r=1}^{n} x_{ik_1} \cdot\cdot x_{k_{r-1}k_r} \quad \text{since} \sum_{j=1}^{n} x_{k_r j}=1 \text{ by (5.3)}$$

and so successively applying this fact we obtain for all $r \in [0..n-2]$

$$\sum_{j=1}^{n} x_{ij}^{(r)} = 1 \qquad \forall\, i \in [1..n]$$

$$\sum_{i=1}^{n} x_{ij}^{(r)} = 1 \qquad \forall\, j \in [1..n] \tag{5.10}$$

Lemma 5.1 together with the constraints of (5.3) constitute an Integer nonlinear programming formulation of the PTSP with a nonlinear objective function which is polynomial in $x_{ij}$ of order n-1 and with linear constraints.

The following fact reduces the order of the polynomial by half:

Fact 5.1

Given a graph G = (N,A,D) and a tour t, the variables $x_{ij}^{(r)}$ defined in Lemma 5.1 are such that:

$$x_{ij}^{(n-2-r)} = x_{ji}^{(r)}$$

(for D not necessarily symmetric)

Proof:

To prove this fact one has to show that:

$$x_{ij}^{(n-2-r)} = 1 \quad \Rightarrow \quad x_{ji}^{(r)} = 1$$

and $$x_{ij}^{(n-2-r)} = 0 \quad \Rightarrow \quad x_{ji}^{(r)} = 0$$

Of course, this is equivalent to showing that:

$$x_{ij}^{(n-2-r)} = 1 \quad \Leftrightarrow \quad x_{ji}^{(r)} = 1$$

Following the proof of Lemma 5.1, we showed that:

$$x_{ij}^{(n-2-r)} = 1 \quad \Leftrightarrow \quad \left\{ \begin{array}{l} \text{there exist } n-2-r \text{ distinct nodes } k_1,\ldots,k_{n-2-r} \\ \text{such that the path } (i,k_1,\ldots,k_{n-2-r},j) \text{ belongs to} \\ \text{to the tour } t \end{array} \right\}$$

By definition of a tour this, in turn, is equivalent to saying that there exist r distinct nodes along the tour between j and i, and this is again equivalent to:

$$x_{ji}^{(r)} = 1$$

Q.E.D.

Note: A consequence of Fact 5.1 is the following:

- if $n-2 = 2q$

$$E[L_t] = \sum_{i=1}^{n} \sum_{j=1}^{n} d(i,j)\Big(\sum_{r=0}^{q-1} \big(\alpha_r\, x_{ij}^{(r)} + \alpha_{n-2-r}\, x_{ji}^{(r)}\big) + \alpha_q\, x_{ij}^{(q)}\Big)$$

- if $n-2 = 2q+1$

$$E[L_t] = \sum_{i=1}^{n} \sum_{j=1}^{n} d(i,j)\Big(\sum_{r=0}^{q} \big(\alpha_r\, x_{ij}^{(r)} + \alpha_{n-2-r}\, x_{ji}^{(r)}\big)\Big)$$

Hence we need only consider the $x_{ij}^{(r)}$ up to $\lfloor\frac{n-2}{2}\rfloor$; that is, $E[L_t]$ is a polynomial in $x_{ij}$ of order $\lfloor\frac{n-2}{2}\rfloor + 1$.

Let us now consider the more complicated cases where the grapn G contains black nodes.

B. Case of a graph G with $m \geq 1$ black nodes:

For this section it will be necessary to distinguish between black and white nodes; we will present our result for a general m ($m \geq 2$) and will briefly indicate at the end the specificity of the case m=1. Without loss of generality we will assume that the n white nodes are labeled 1,2,...,n and m black nodes n+1,...,n+m. Our main result is contained in the following lemma:

Lemma 5.2

Let G = (N,A,D) be a graph with m black nodes and n white nodes (assume the white nodes are labeled 1,2,..,n), let t be a given tour of G and let $x_{ij}$ be the 0-1 decision variables defined in (5.2). Then:

$$E[L_t] = \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} \sum_{r=0}^{n} \alpha_r\, d(i,j)\, y_{ij}^{(r)}$$

where:

(1) $y_{ij}^{(o)} \equiv x_{ij}$ for all i,j

(2) $$y_{ij}^{(r)} \equiv \begin{cases} x_{ij}^{(r)} & \text{if } 1 \leqslant i,j \leqslant n \\ \sum_{k=0}^{r} (r+1-k)\ x_{ij}^{(k)} & \text{if } n+1 \leqslant i,j \leqslant n+m \\ \sum_{k=0}^{r} x_{ij}^{(k)} & \text{otherwise} \end{cases}$$

$\forall\ r \in [1..n-1]$

(3) $$y_{ij}^{(n)} \equiv \begin{cases} \sum_{k=0}^{n} x_{ij}^{(k)} & \text{if } n+1 \leqslant i,j \leqslant n+m \\ 0 & \text{otherwise} \end{cases}$$

and $$x_{ij}^{(r)} = \sum_{k_1,\ldots,k_r=1}^{n} x_{ik_1}, x_{k_1k_2} \cdot\cdot x_{k_rj}$$

$\forall\ r \in [1..n]$

<u>Proof</u>:

We proceed as with Lemma 5.1; we have

$$E[L_t] = \sum_{r=0}^{n} \alpha_r\ L_{m,t}^{(r)} \tag{5.11}$$

For $r \in [1..n]$ define the variables $x_{ij}^{(r)}$ and $y_{ij}^{(r)}$ as:

$$x_{ij}^{(r)} = \begin{cases} 1 & \text{if there exist exactly r white nodes between i and j along the tour t.} \\ 0 & \text{otherwise} \end{cases} \tag{5.12}$$

$$y_{ij}^{(r)} = \{\text{number of times } d(i,j) \text{ is a part of } L_{m,t}^{(r)}\} \tag{5.13}$$

We then have (by definition of $L_{m,t}^{(r)}$):

$$L_{m,t}^{(r)} \equiv \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} d(i,j)\ y_{ij}^{(r)} \tag{5.14}$$

From the proof of Lemma 5.1 we know that (5.12) can be expressed as:

$$x_{ij}^{(r)} = \sum_{k_1,\ldots,k_r=1}^{n} x_{ik_1},\ldots x_{k_r j} \tag{5.15}$$

To conclude the proof of Lemma 5.2 it remains to establish the relationships between $y_{ij}^{(r)}$ and $x_{ij}^{(r)}$ for the different cases (note that we will implicitly use results from Chapter 3):

(1) r=0; by definition of $L_{m,t}^{(o)}$, $y_{ij}^{(o)} = \begin{cases} 1 & \text{if } (i,j) \text{ belongs to the tour} \\ 0 & \text{otherwise} \end{cases}$

hence $y_{ij}^{(o)} = x_{ij}$

(2) $r \in [1..n-1]$; we have three subcases:

- if i and j are white nodes; then d(i,j) appears at most once in $L_{m,t}^{(r)}$; is appears only when i and j are separated by r white nodes: hence $y_{ij}^{(r)} = x_{ij}^{(r)}$
- if i and j are black nodes; then d(i,j) can appear up to (r+1) times in $L_{m,t}^{(r)}$; in fact it appears exactly k times if i and j are separated exclusively by r+1-k white nodes along the tour t, that is when $x_{ij}^{(r+1-k)} = 1$; hence,

$$y_{ij}^{(r)} = \sum_{k=1}^{r+1} k\ x_{ij}^{(r+1-k)} \tag{5.16}$$

and by setting $r+1-k \equiv u$ we get the desired result.

- if one node is white and one black; then $d(i,j)$ appears at most one and it does appear when $i$ and $j$ are separated exclusively by not more than $r$ white nodes; hence

$$y_{ij}^{(r)} = \sum_{k=0}^{r} x_{ij}^{(k)} \tag{5.17}$$

(3) $r=n$; then again $d(i,j)$ appears at most once and only one when $i$ and $j$ are both black and if they are separated exclusively by white nodes (up to $n$ of them). Hence,

$$y_{ij}^{(n)} = \begin{cases} \sum_{k=0}^{n} x_{ij}^{(k)} & \text{if } n+1 \leqslant i,j \leqslant n+m \\ 0 & \text{otherwise} \end{cases}$$

Q.E.D.

Note:

It is easy (but somewhat cumbersome) to verify that:

$$\sum_{i=1}^{n+m} y_{ij}^{(r)} = \begin{cases} 1 & \text{if } 1 \leqslant j \leqslant n \\ r+1 & \text{otherwise} \end{cases}$$
$$\sum_{j=1}^{n+m} y_{ij}^{(r)} = \begin{cases} 1 & \text{if } 1 \leqslant i \leqslant n \\ r+1 & \text{otherwise} \end{cases} \tag{5.18}$$

for all $r \in [0..n-1]$

For the case of only one black node Lemma 5.2 still holds, but we only have to consider case (1) and case (2).

Also one should note that when $m \geq 1$, there is <u>no</u> analogous result to Fact 5.1 obtained for m=0.

In conclusion, we have seen that every case of the PTSP (i.e., m=0, m≥1) can be formulated as an integer nonlinear programming problem with a nonlinear objective function and with the same linear constraints as for the traditional TSP formulation.

In our next subsection we will present two possible linearizations of the formulation.

### 5.2.2 Linearizations:

We will present our results for the case m=0 and will briefly mention how to extend them to the cases m≥1.

A. Graph G with no black node:

(1) first linearization: mixed integer linear programming formulation:

In the previous section we expressed $x_{ij}^{(r)}$ given by (5.8) as a function of products of $x_{ij}$'s (see (5.9)); however, this is not the only way of expressing this boolean expression, and an alternative method is presented in the next lemma. First we need to state the following simple fact (the simple proof is omitted):

Fact 5.2:

Let $X_1$, $Y_1$, $X_2$, $Y_2$ be four boolean variables (i.e., 0-1 variables) then the variable

$$Z = \begin{cases} 1 \text{ if } (X_1=1 \text{ and } Y_1=1) \text{ or } (X_2=1 \text{ and } Y_2=1) \\ 0 \text{ otherwise} \end{cases}$$

can be expressed by the following inequalities:

$$X_1 + Y_1 - 1 \leq Z \leq 1$$

$$X_2 + Y_2 - 1 \leq Z \leq 1$$

$$0 \leq Z \leq X_1 + X_2$$

$$0 \leq Z \leq X_1 + Y_2$$

$$0 \leq Z \leq Y_1 + X_2$$

$$0 \leq Z \leq Y_1 + Y_2$$

We can now state our main result:

<u>Lemma 5.3</u>:

The variables $x_{ij}^{(r)}$ defined in (5.8) can be expressed by the following set of inequalities ($r \in [1..n-2]$)

$$x_{ik}^{(r-1)} + x_{kj} - 1 \leq x_{ij}^{(r)} \leq 1 + x_{ik}^{(r-1)} - x_{kj} \qquad 1 \leq k \leq n$$

<u>Proof</u>:

First let us prove Lemma 5.3 for r=1. One can express (5.8) in the following form:

$$x_{ij}^{(1)} = \begin{cases} 1 & \text{if } (x_{i1}=1 \text{ and } x_{1j}=1) \text{ or } (x_{i2}=1 \text{ and } x_{2j}=1) \text{ or } \ldots\ldots \\ & \qquad \ldots\ldots\ldots (x_{in}=1 \text{ and } x_{nj}=1) \\ 0 & \text{otherwise} \end{cases}$$

From Fact 5.2 the set of inequalities

$$x_{ik} + x_{kj} - 1 \leq x_{ij}^{(1)} \leq 1 \qquad\qquad 1 \leq k \leq n$$

force $x_{ij}^{(1)}$ to be 1 if there exists a node $k \in [1..n]$ such that $x_{ik} = 1$ and $x_{kj} = 1$. To express that $x_{ij}^{(1)}$ has to be zero if there is no such node k, according to Fact 5.2 we would need many inequalities of the form

$$0 \leq x_{ij}^{(1)} \leq \sum_{k\epsilon S} x_{ik} + \sum_{k\notin S} x_{kj} \qquad (5.19)$$

for any subset S of $N=\{1,2,\ldots,n\}$

In fact we know that ((5.10)):

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad \text{for all } j$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad \text{for all } i$$

Hence the right-hand side inequality of (5.19) is equivalent to:

$$\sum_{k\epsilon S} x_{ik} + 1 - \sum_{k\epsilon S} x_{kj}$$

$$= 1 + \sum_{k\epsilon S} (x_{ik} - x_{kj}) \text{ for any subset } S \subseteq N$$

From there it is easy to see that it suffices to consider <u>only</u> the subset S containing one element (i.e. $S=\{k\}$ $k \in N$) since $x_{ij}^{(1)} = 0$ implies we <u>must</u> have one k such that

$x_{ik} = 0$ <u>and</u> $x_{kj} = 1$.

$$\text{hence } 0 \leq x_{ij}^{(1)} \leq 1 + x_{ik} - x_{kj} \qquad 1 \leq k \leq n \qquad (5.20)$$

are a sufficient set of inequalities to force $x_{ij}^{(1)}$ to 0, in case $x_{ij}^{(1)} \neq 1$.

This has proved Lemma 5.3 for r=1.

For $r > 1$ one can note that we have:

$$x_{ij}^{(r)} = \sum_{k=1}^{n} x_{ik}^{(r-1)} x_{kj} \qquad (5.21)$$

$$\sum_{i=1}^{n} x_{ij}^{(r-1)} = 1 \qquad \text{for all } j$$

$$(5.22)$$

$$\sum_{j=1}^{n} x_{ij}^{(r-1)} = 1 \qquad \text{for all } i$$

Hence one can proceed exactly as before to obtain the desired result.

Q.E.D.

Consequence:

Lemma 5.3 allows us to formulate the PTSP (with m=0) as a mixed integer linear program; the decision variables are the 0-1 variables $x_{ij}$ and the continuous variables $x_{ij}^{(r)}$.

The constraints are, in addition to the ones of the integer nonlinear program given previously, those given in Lemma 5.3; the objective function is still given by Lemma 5.1.

Compared to the integer nonlinear programming formulation, this new formulation is linear but adds $n^2(n-2)$ new continuous variables ($x_{ij}^{(r)}$ $r \in [1..n-2]$) and for each of them we have 2n constraints, i.e. a total of $2n^3(n-2)$ new constraints. (Note: one can use Fact 5.1 to reduce the number of new variables and additional constraints respectively to $n^2\left(\frac{n-2}{2}\right)$ and $n^3(n-2)$).

In any case we have $O(n^3)$ additional continuous variables and $O(n^4)$ additional constraints compared to the original formulation.

Let us present another linearization that will require fewer constraints and that will give a pure integer linear programming formulation of the PTSP.

(2) Second linearization: pure integer linear programming formulation: The result of this section is contained in the following lemma:

Lemma 5.4:

The variables $x_{ij}^{(r)}$ defined in (5.8) can be expressed by the following set of inequalities ($r \in [1..n-2]$)

- $x_{ij}^{(r)} \leq 1 - z_{ij}^{(r-1)}$
- $x_{ij}^{(r)} \leq 1 + z_{ij}^{(r-1)}$
- $x_{ij}^{(r)} \geq \alpha_{ij}^{(r)} + \beta_{ij}^{(r)} - 1$
- $\alpha_{ij}^{(r)} \geq \frac{1}{n}\left[1 + (n-1)\, z_{ij}^{(r-1)}\right]$
- $\beta_{ij}^{(r)} \geq \frac{1}{n}\left[1 - (n-1)\, z_{ij}^{(r-1)}\right]$
- $x_{ij}^{(r)}, \alpha_{ij}^{(r)}, \beta_{ij}^{(r)} \in \{0,1\}$

where $z_{ij}^{(r-1)} = \left(\sum_{k=1}^{n} k(x_{ik}^{(r-1)} - x_{kj})\right)/(n-1)$

Proof:

As in the proof of Lemma 5.3, let us first consider the case $r=1$:

The main idea is the following: if one assigns a distinct number (say $\xi_k$) to each node k of G, then the quantity $\sum_{k=1}^{n} \xi_k(x_{ik}-x_{kj})$ will take on the value 0 only if there exists a $k \in [1..n]$ such that $x_{ik}=1$ and

$x_{kj}=1$; indeed from (5.3) we know that:

- there exists a unique $\ell$ such that $x_{i\ell} = 1$ for a given i.
- there exists a unique $\ell'$ such that $x_{\ell' j} = 1$ for a given j.

Hence:

$\sum_{k=1}^{n} \xi_k(x_{ik}-x_{kj}) = \xi_\ell - \xi_{\ell'}$ which is zero only if $\ell=\ell'$.

The quantity $z_{ij}$ defined in Lemma 5.4 simply corresponds to the choice $\xi_k=k$ (the natural number k) and is divided by n-1 for normalization (so that $-1 \leq z_{ij} \leq 1$).

Now we have the following logical relationships:

$$\text{if } z_{ij} = 0 \Rightarrow x_{ij}^{(1)} = 1 \tag{5.23}$$

$$\text{if } z_{ij} \neq 0 \Rightarrow x_{ij}^{(1)} = 0 \tag{5.24}$$

- Assuming that $x_{ij}^{(1)} \in \{0,1\}$ then (5.24) can be expressed by:

$$x_{ij}^{(1)} \leq 1 - z_{ij}$$

$$x_{ij}^{(1)} \leq 1 + z_{ij}$$

- To express (5.23) one has to introduce two new 0-1 variables $\alpha_{ij}^{(1)}$ and $\beta_{ij}^{(1)}$ as follows:

$$\text{if } z_{ij} > 0 \Rightarrow \alpha_{ij}^{(1)} = 1 \tag{5.25}$$

$$\text{if } z_{ij} < 0 \Rightarrow \beta_{ij}^{(1)} = 1 \tag{5.26}$$

Now (5.23) can be equivalently stated as:

$$\text{if } \alpha_{ij}^{(1)} = 1 \text{ \underline{and} } \beta_{ij}^{(1)} = 1 \Rightarrow x_{ij}^{(1)} = 1 \tag{5.27}$$

Finally (5.25), (5.26), and (5.27) can be expressed by the following inequalities:

$$\alpha_{ij}^{(1)} \geq \frac{1}{n} [1 + (n-1)\ z_{ij}]$$

$$\beta_{ij}^{(1)} \geq \frac{1}{n} [1 - (n-1)\ z_{ij}]$$

$$x_{ij}^{(1)} \geq \alpha_{ij}^{(1)} + \beta_{ij}^{(1)} - 1$$

The case r=1 is now proved. The general case ($r > 1$) follow immediately (for the same reasons as given in Lemma 5.3).

Q.E.D.

Consequence:

Lemma 5.4 leads to a pure integer linear programming formulation for the PTSP (m=0); the decision variables are the 0-1 variables $x_{ij}^{(r)}$ ($r \in [0..n-2]$) $\alpha_{ij}^{(r)}$, $\beta_{ij}^{(r)}$ ($r \in [1..n-2]$); the constraints are, in addition to the traditional TSP formulation's constraints, those given in Lemma 5.4; the objective function is expressed as in Lemma 5.1.

Compared to the integer nonlinear programming formulation, one now has $3n^2(n-2)$ new 0-1 variables ($x_{ij}^{(r)}$, $\alpha_{ij}^{(r)}$, $\beta_{ij}^{(r)}$ $r \in [1..n-2]$) and $8n^2(n-2)$ addtional constraints.

Before going to the next section, let us briefly look at the linearizations of the PSTP when $m \geq 1$.

B. Graph G with black nodes:

We saw in part A. that $x_{ij}^{(r)} = \sum_{k=1}^{n} x_{ik}^{(r-1)} x_{kj}$ allowed us to obtain first the linearizations of $x_{ij}^{(1)}$ and then to obtain linearizations of $x_{ij}^{(r)}$.

One can see that the $y_{ij}^{(r)}$ defined in Lemma 5.2 can be obtained by similar recursive relationships; for example for $r \in [1..n-1]$ one has:

$$y_{ij}^{(r)} = \begin{cases} \sum_{k=1}^{n} y_{ik}^{(r-1)} x_{kj} & 1 \leq i,j \leq n \\ y_{ij}^{(r-1)} + x_{ij} + \sum_{k=1}^{n} y_{ik}^{(r-1)} x_{kj} & n+1 \leq i,j \leq n+m \\ y_{ij}^{(r-1)} + \sum_{k=1}^{n} y_{ik}^{(r-1)} x_{kj} & \begin{cases} 1 \leq i \leq n \\ n+1 \leq j \leq n+m \end{cases} \\ y_{ij}^{(r-1)} + \sum_{k=1}^{n} x_{ik} y_{kj}^{(r-1)} & \begin{cases} n+1 \leq i \leq n+m \\ 1 \leq j \leq n \end{cases} \end{cases}$$

So as before, one has to linearize $y_{ij}^{(1)}$ for the case $r=1$ only, and then generalize it to $r > 1$.

For $y_{ij}^{(1)}$ one can use similar linearization ideas as the ones developed for $m=0$; we will however not present any detailed derivations since they would not add anything new compared to previous results.

Let us now turn our attention to the discussion of exact solution methods.

5.2.3 Discussion of Exact Methods:

Based on the mathematical programming formulations of the previous two subsections (5.2.1 and 5.2.2), we shall analyze in the remainder of section 5.2 what seem to be the more promising strategies for finding optimal solutions to the PTSP. This subsection is organized in three parts: we provide first a discussion of the relative merits of the three formulations; this is followed by the presentation of a Branch-and-Bound approach which we consider one of the few "feasible" approaches; finally in the last part we will show how a seemingly natural dynamic programming formulation of the PTSP is in fact not valid.

For the rest of section 5.2 (except part A) we will concentrate our analysis on the case in which the graph G has n white nodes and one black node (a depot); based on 5.2.1 and 5.2.2 the other cases can be similarly treated. One of the reasons for our choice is that the case of one black node and n white nodes is the first in the series of PTSPs for which we have a depot (which is almost always the case in practical applications).

A. General Discussion:

For this part, where general comments are sought, it is more convenient to place our discussion in the context of m=0; this is done, however, without loss of generality.

For completeness the three mathematical programming formulations are given in Appendix F and we will refer to them as follows:

$F_1$ : Integer Nonlinear Programming Formulation

$F_2$ : Mixed Integer Linear Programming Formulation

$F_3$ : Pure Integer Linear Programming Formulation.

Also included in Appendix F is a numerical table giving for some chosen values of n respectively the number of variables and constraints for $F_1$, $F_2$, and $F_3$.

As one can see from this numerical table, there is little hope of solving exactly problems of size of more than 6 nodes using a general integer programming code on $F_2$ and $F_3$. In both these formulations the numbers of constraints and of integer variables become rapidly too large; moreover for problems of such small size, complete enumeration would possibly be faster.

On the other hand, $F_1$ has the same set of constraints as the TSP and this has a double advantage (over $F_2$ and $F_3$):

1. those constraints have been analyzed over and over again by many authors and they define a feasible region (polytope) of which we begin to have some understanding (see Klee [1980] for a general discussion).

2. a large number of papers have been published on the development of exact solution procedures for the TSP and the accumulated experience within this area should be taken advantage of.

Moreover in formulations $F_2$ and $F_3$ we lose an important advantage, namely that whenever a feasible set of $x_{ij}$ is obtained, the other variables are <u>all</u> automatically determined; this is of course also achieved by our linear constraints, but only indirectly (and "inefficiently").

In summary, a general integer linear programming code (for $F_2$ or $F_3$) has to be automatically discarded because, given the current state-of-the-art, it could solve problem sizes so small that either the results from Chapter 3 would have already told us that the TSP tour solves the problem, or the problem could be solved by complete enumeration. Hence one should concentrate on specially designed algorithms; for $F_1$ one should benefit from results for the TSP; as for $F_2$ and $F_3$ no immediate special structure seems to appear (one may simply note that $F_3$ is a 3-dimensional assignment problem with additional constraints, and this does not help much!). Based on this brief discussion it seems natural to concentrate on $F_1$; this is the concern of the next section.

B. Branch-and-Bound Approaches:

Following the argument in part A, we will analyze in somewhat more detail the formulation $F_1$; as stated, we shall concentrate on the variation of the PTSP with one black node and n white nodes.

$F_1$ is an Integer Nonlinear programming problem with a nonlinear objective function polynomial of order n in the $x_{ij}$'s (see (5.8) and Lemma 5.2) and with a set of linear constraints identical to the TSP formulation (see (5.3)).

The general idea (adapted from Lawler [1963]) can be expressed as follows (let $X = (x_{ij})$):

Suppose that our objective function f(X) (see Lemma 5.2) can be decomposed as the sum of a low-order polynomial function g(X) and another function h(X) that is "less important" than g(X). Suppose a lower bound (of course independent of X) on h(X) is available - say h - and suppose $\overline{X}$ is a known feasible solution to the problem. Then one can reduce the feasible region (i.e., defined by the set of all tours) by limiting our search for an optimal PTSP tour among the tours such that $S = \{X : g(X)+h \leq f(\overline{X})\}$; in other words one can discard all tours such that

$$g(X) + h > f(\overline{X}) \qquad (5.28)$$

A branch-and-bound procedure would then be based on the minimization of g(X), a task certainly easier than the minimization of f(X).

Let us see how one can apply this general strategy in the context of the PTSP. Let $X = \{x_{ij}\}$ with $x_{ij}$ the 0-1 variables defined in (5.2); from Lemma 5.2 one can express the objective function as:

$$L(X) \equiv E[L_t] = \sum_{r=0}^{n-1} \alpha_r L_{1,t}^{(r)} \equiv \sum_{r=0}^{n-1} \alpha_r L^{(r)}(X)$$

where $L^{(r)}(X)$ is a polynomial of order r+1.

Now we saw in Chapter 3 that the weights $\alpha_r$ form a non-increasing sequence (i.e., $\alpha_{r+1} < \alpha_r \; \forall \; r \in [0..n-3]$); so, depending on the choice of W, only the first terms of L(X) are likely to be important.

We will limit our investigation to the very particular cases where $\alpha_0 L^{(o)}(X)$ (i.e. the linear term) is dominant; for example if W is a binomial random variable corresponding to a coverage probability p=0.9, then $\alpha_0=0.81$, $\alpha_1=0.081$, $\alpha_2=0.0081$ etc. These are cases which we believe can be solved by directly using techniques developed for the TSP (i.e. Held and Karp [1970] [1971] for symmetric problems, and Balas and Christofides [1981] for asymmetric problems) for problems of relatively modest size.

Comment:

When the cases considered necessitate consideration of the quadratic term as well (or higher-order polynomials) we will not be able to use exact methods developed for the TSP and entirely new methods should then have to be considered. This is outside of the scope of the thesis as argued in the introduction. Moreover it is very likely that not one but several different schemes have to be developed to solve PTSPs of non-trivially small size, depending on the distribution of W and the relative values of the objective function's weights. When one notes that the TSP is a PTSP that corresponds to a very specific case for W (namely Pr(W=n)=1) and that the most efficient exact methods for this problem have been built on more than two decades of intensive research, one realizes that this area must be left as a topic for future research.

(end of comment)

Coming back to cases of the PTSP which are in some sense close to the TSP (i.e. $\alpha_0 L_{1,t}^{(o)}$ is dominant), one should be able to obtain good results; those will depend on the quality of the lower bound on $\sum_{r=1}^{n-1} \alpha_r L_{1,t}^{(r)}$ .

From Chapter 3 we have two possibilities based either on Lemma 3.1 or Lemma 3.2; the bound developed using Lemma 3.2 might be better than the one from Lemma 3.1 but it requires obtaining the length of the corresponding optimal TSP tour (its length) first.

Let us conclude this section on exact methods by showing that some procedures proposed to solve the TSP cannot be applied to the PTSP, even for cases "close" to the TSP (i.e. p close to 1).

### C. The Inadequacy of a Dynamic Programming Approach for the PTSP

In previous sections, recognizing that the PTSP could be formulated in exactly the same manner as the TSP apart from a different objective function, we took advantage of this relationship to propose some procedures based on the most efficient known exact method for the TSP.

Here we will provide an illustration of the truly complex relationship between the TSP and the PTSP and will show that one cannot always extend results developed for the TSP (which is a special case of the PTSP) to the general PTSP.

We have seen in Chapter 2 that the expected length of a tour t can be computed using the recursive relationships (2.8). These relationships suggest, at first glance, a straightforward extension of the dynamic programming formulation for the TSP to the PTSP; unfortunately as we will soon see this approach does not work.

(i) Let us first briefly present for completeness the classical dynamic programming formulation for the TSP (see Held and Karp [1962]) in terms of our notation:

Let $G = (N \equiv N_1 \cup N_2, A, D)$ where $N_1=\{1\}$ $N_2=\{2,3..,n+1\}$.

Given a set $Q \subseteq N_2$ and $k \in Q$, we let $L(Q,k)$ be the length of the shortest path starting from node k, visiting all the nodes in Q, and ending at node 1; the length of an optimal TSP tour will then be given by $L(N_2,1)$.

We begin by finding $L(Q,k)$ for $|Q|=0$, which is simply

$$L(\phi,k) = d(k,1) \;\forall\; k \in N$$

To calculate $L(Q,k)$ for $|Q| > 0$, we argue that the corresponding shortest path is obtained by first going to a node $\ell$ ($\ell \in Q$) and then by looking up $L(Q-\{\ell\},\ell)$ in our preceeding table. Thus

$$L(Q,k) = \min_{\ell \in Q} [d(k,\ell) + L(Q-\{\ell\}, \ell)] \tag{5.29}$$

It is clear that one has to calculate $L(Q,k)$ for all sets Q of a given size and for each possible node of N-Q, before one can go to the next step. Hence if we count each value of $L(Q,k)$ as one storage location, we need space equal to:

$$\sum_{q=0}^{n} (n+1-q)\binom{n}{q} = 2^n + n2^{n-1} = (n+2)2^{n-1} = O(n2^n)$$

and a number of additions and comparisons equal to

$$\sum_{q=0}^{n} (n+1-q)(n-q)\binom{n}{q} = n \sum_{q=0}^{n-1} (n+1-q)\binom{n-1}{q}$$

$$= n[2^n+(n-1)2^{n-2}] = O(n^2 2^n)$$

(ii) Assume now that $N_1=\{1\}$ corresponds to a black node and $N_2$ is the set of n white nodes; also assume for simplicity that W is a binomial random variable corresponding to a coverage probability p. To solve this corresponding PTSP, it seems natural (based on the recursive relationship (2.8)) to extend the dynamic programming formulation of the TSP as follows:

Given, as before, a set $Q \subseteq N_2$ and $k \notin Q$,

let E(Q,k) be the expected length (in the PTSP sense) of the optimal path starting from node k, visiting all the nodes in Q, and ending at node 1; also let F(Q,k) be the node following immediately k on such an optimal path. Then the value of the optimal solution for the PTSP is given by $E(N_2,k)$.

Now it is true that $E(\phi,k) = d(k,1)$ for all $k \in N$

$$F(\phi,k) = 1$$

For Q such that $|Q| > 0$, one could argue as in the derivation of (2.8), and claim that the following recursive relationships hold (note that they are indeed natural extensions of (5.29), see Figure 5.1 for an illustration):

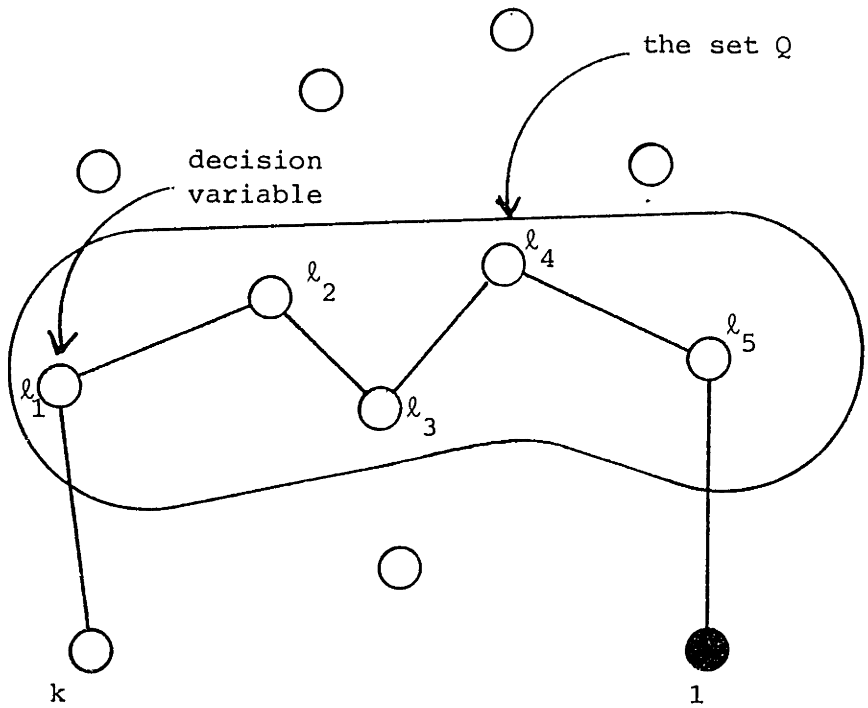



Figure 5.1: An illustration of a D.P. Approach for the PTSP

$$E(Q,k) = \underset{\ell_1 \varepsilon Q}{\text{Min}} \left\{ p \sum_{r=1}^{|Q|} (1-p)^{r-1} \left( d(k,\ell_r) + E\left(Q - \bigcup_{j=1}^{r} \{\ell_j\}, \ell_r\right)\right) + (1-p)^{|Q|} d(k,1) \right\} \tag{5.30}$$

where $\ell_r = F\left(Q - \bigcup_{j=1}^{r-1} \{x_j\},\ \ell_{r-1}\right)$ for $r \varepsilon [2..|Q|]$

This approach is seen to have the same space requirement as for the TSP (i.e. $O(n2^n)$) and an order $O(n)$ more of additions and comparisons (i.e. $O(n^3 2^n)$.

However, this approach does not work; to be more precise the recursive relationships (5.30) are not valid. (We provide in Appendix H a numerical example demonstrating this.)

(5.30) is based on Bellman's famous principle of optimality (Bellman [1957]), which states: "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". At first glance this seems to be a paradox, but in fact it can be explained by the fact that this principle applies only to problems that can be decomposed into stages, with a decision required at each stage. Indeed the PTSP is not such a problem and we have already at least two earlier results that point out this fact. One can simply note that the objective function is not separable in the $x_{ij}$'s as shown in Lemma 5.2 (which also explains why the DP approach works for the TSP). In fact, more generally, one <u>cannot</u> solve the PTSP by <u>any</u> stage-decomposition approach, since the underlying structure of the problem is exactly in the opposite direction; namely the optimal solution can be discovered only by looking at the problem as a whole (see for example

Lemma 3.12).

This concludes the discussion of exact optimization methods to solve the PTSP; as indicated in the introductory section the main goals of this section were to provide several mathematical programming formulations for the problem (this, by itself, requiring some elaborate analysis) and to discuss possible directions for valid approaches.

As could be expected for a thesis whose main emphasis is not on algorithmic design, the discussion was general and the main message to be drawn is that the PTSP is an extremely difficult combinatorial problem for which one should definitely not count on any unique exact scheme to solve problems of even very modest size for all possible p.m.f.'s for W. For some cases (i.e., p close to 1 in the binomial case) we are, however, confident that procedures developed for the TSP can be used profitably to tackle our problem.

We now turn our attention to heuristic procedures.

## 5.3 Heuristic Procedures

As mentioned in the introductory section (5.1.1) and as demonstrated in section 5.2, the PTSP belongs to a class of difficult combinatorial optimization problems whose exact solutions seem to require a non-polynomial (as a function of the input) amount of time. Hence for problems of medium and large size it seems reasonable to concentrate on procedures which, although not guaranteed to find the optimal solution, will provide "good" approximate solutions. This is the concern of this section. Before presenting and analyzing such procedures, let us first give some theoretical preliminaries upon which some procedures will be based. Also, let us recall that we shall concentrate on the PTSP with one black node and n white nodes.

Finally, to complete the preliminaries, recall that from Lemma 3.9 we also know that, for any graph G of size up to n=3, we have $t_1 \equiv t_p$ for any p.m.f. for W, provided that the distance matrix D is symmetric; for asymmetric D the same result holds up to n=2.

Hence, from now on, we will assume that $n \geq 3$ ($n \geq 4$ if D symmetric).

### 5.3.1 Theoretical Preliminaries

A. Solving the PTSP by solving easier problems; worst-case ratios.

We have $G=(N_1 \cup N_2, D, A)$ $|N_1|=1$, $|N_2|=n$. Let t be a given tour of G; we know from Chapters 2 and 3 that the expected length (in the PTSP sense) of t is given by:

$$E[L_t] = \sum_{r=0}^{n-1} \alpha_r L_{1,t}^{(r)} \tag{5.31}$$

In Chapters 3 and 4 we already introduced two particular tours, namely:

- the optimal TSP tour $t_1$ of G, which solves $\min_t \{\alpha_0 L_{1,t}^{(o)}\}$
- the optimal PTSP tour $t_p$ of G, which solves $\min_t \{E[L_t] = \sum_{r=0}^{n-1} \alpha_r L_{1,t}^{(r)}\}$

(note that from Fact 3.2 in Chapter 3 we know that $L_{1,t}^{(n-1)}$ is tour-independent, hence $t_p$ solves also $\min_t \{\sum_{r=0}^{n-2} \alpha_r L_{1,t}^{(r)}\}$).

In fact (5.31) suggests the introduction of (n-1) different optimization problems of increasing difficulty (we will be more precise later on), each of them seeking a specific tour, the TSP and PTSP being too extremes (respectively the easiest and the hardest of them); this can be seen as follows:

For $k \in [0..n-2]$ define $v_k$ to be a tour that solves the following minimization problem:

$$\min_{t}\{\sum_{r=0}^{k} \alpha_r L_{1,t}^{(r)}\} \tag{5.32}$$

Then we have $v_o \equiv t_1$ and $v_{n-2} \equiv t_p$. Each intermediate problem and its solution (i.e., $v_k$, $k \in [1..n-3]$) does not possess a physical interpretation (as the TSP or PTSP do), but they can certainly be of great help with respect to finding heuristic procedures for the PTSP. In fact the strategy of solving the PTSP by finding $v_k$ ($k \in [0..n-3]$) already provides n-2 different heuristics (of increasing difficulty as k increases); if one considers in turn that each of these intermediate problems can be solved either exactly or by heuristic procedures, we then have at hand a great number of possibilities for designing heuristic procedures for the PTSP.

Before discussing in more detail which strategy to choose (this will of course depend on the p.m.f. for W), let us first derive upper bounds on "how far" from the optimal value of the PTSP tour a tour $v_k$ can be; that is, we assume that we are given the tour $v_k$, and we want to compute the largest percentage above optimality that may be attained with $v_k$ (i.e., the worst case ratio); we already derive such a bound for $v_o \equiv t_1$: in Chapter 3 (Theorem 3.1) we obtained the following result

$$\frac{E[L_{t_1}] - E[L_{t_p}]}{E[L_{t_p}]} < \frac{1-E[W]/n}{E[W]/n} \tag{5.33}$$

It seems natural to expect this upper bound to improve (i.e., become smaller) for tour $v_k$, $k > 1$. Indeed $t_1$ is the tour that minimizes the first element of $E[L_t]$, while $v_1$ minimizes the two first elements of $E[L_t]$ and so on.

To illustrate this point we will formally obtain an upper bound on $\frac{E[L_{v_1}]-E[L_{t_p}]}{E[L_{t_p}]}$ for a general pmf for W and we will briefly mention similar bounds for $v_k$, $k > 1$.

Let us present our result for $v_1$ (recall that D is assumed to satisfy the triangular inequality):

Lemma 5.5

Let G be a given graph with n white nodes and one black node; then we have:

$$\text{(i)} \quad \frac{E[L_{v_1}]-E[L_{t_p}]}{E[L_{t_p}]} \leq \frac{(\alpha_0+\alpha_1)(1-2\alpha_1)-\alpha_0 E[W]/n}{\alpha_0(E[W]/n+\alpha_1)}$$

where $\alpha_0 \equiv (E[W^2]-E[W])/n(n-1)$

$\alpha_1 \equiv \alpha_0-(E[W^3]-3E[W^2]+2E[W])/n(n-1)(n-2)$

(ii) for the binomial case (i) becomes:

$$\frac{E[L_{v_1}]-E[L_{t_p}]}{E[L_{t_p}]} \leq \frac{2(1-p)^2}{p}$$

Proof:

- $v_1$ is a tour minimizing $\alpha_0 L_{1,t}^{(o)} + \alpha_1 L_{1,t}^{(1)}$, hence we have:

$$E[L_{t_p}] = \sum_{r=0}^{n-1} \alpha_r L_{1,t_p}^{(r)} \geq \alpha_0 L_{1,v_1}^{(o)} + \alpha_1 L_{1,v_1}^{(1)} + \sum_{r=2}^{n-1} \alpha_r L_{1,t_p}^{(r)} \qquad (5.34)$$

from Chapter 3 (Lemma 3.2) we have:

$$L_{1,t}^{(r)} \geq L_{1,t_1}^{(o)} \quad \forall\, r \in [0..n-1] \quad \text{for any given tour } t$$

(5.34) can then be rewritten as follows:

$$E[L_{t_p}] > \alpha_0 L^{(o)}_{1,v_1} + \alpha_1 L^{(1)}_{1,v_1} + L^{(o)}_{1,t_1} \left[\sum_{r=2}^{n-1} \alpha_r\right] \tag{5.35}$$

- From Lemma 3.3 we have:

$$E[L_{v_1}] < \alpha_0 L^{(o)}_{1,v_1} + \alpha_1 L^{(1)}_{1,v_1} + L^{(o)}_{1,v_1} \left[\sum_{r=2}^{n-1} (r+1)\ \alpha_r\right] \tag{5.36}$$

- If we combine (5.35) and (5.36) we obtain:

$$E[L_{v_1}] - E[L_{t_p}] < L^{(o)}_{1,v_1} \left[\sum_{r=2}^{n-1} (r+1)\ \alpha_r\right] - L^{(o)}_{1,t_1} \left[\sum_{r=2}^{n-1} \alpha_r\right] \tag{5.37}$$

- If remains, now, to compare $L^{(o)}_{1,v_1}$ with $L^{(o)}_{1,t_1}$:

By definition of $t_1$ we know that:

$$L^{(o)}_{1,t_1} < L^{(o)}_{1,v_1} \tag{5.38}$$

(5.38) is however of no help here, since we need to have $L^{(o)}_{1,v_1}$ bounded from above by a function of $L^{(o)}_{1,t_1}$. This is done as follows:

By definition of $v_1$:

$$\alpha_0 L^{(o)}_{1,t_1} + \alpha_1 L^{(1)}_{1,t_1} > \alpha_0 L^{(o)}_{1,v_1} + \alpha_1 L^{(1)}_{1,v_1} \tag{5.39}$$

From Lemma 3.3

$$\alpha_0 L^{(o)}_{1,t_1} + \alpha_1 L^{(1)}_{1,t_1} < (\alpha_0 + 2\alpha_1)\ L^{(o)}_{1,t_1} \tag{5.40}$$

From Lemma 3.2

$$\alpha_0 L^{(o)}_{1,v_1} + \alpha_1 L^{(1)}_{1,v_1} \geq \alpha_0 L^{(o)}_{1,v_1} + \alpha_1 L^{(o)}_{1,t_1} \tag{5.41}$$

Finally from (5.39), (5.40), and (5.41) we have:

$$L^{(o)}_{1,t_1} \geq \frac{\alpha_0}{\alpha_0+\alpha_1} L^{(o)}_{1,v_1} \tag{5.42}$$

also

$$L^{(1)}_{1,v_1} \geq L^{(o)}_{1,t_1} \tag{5.43}$$

- In conclusion using (5.35), (5.37), (5.42), and (5.43) we obtain the desired results. The expressions given for $\alpha_0$ and $\alpha_1$ are obtained through a straightforward computation from the definition of $\alpha_r$ (see Chapter 3, 3.1.2 B.). When W is a binomial random variable corresponding to a coverage probability p, we know from Chapter 2 that $\alpha_0=p^2$, $\alpha_1=p^2(1-p)$, and $E[W]/n=p$. It suffices then to substitute for these quantities in the general formula.

Q.E.D.

Notes:

(1) Lemma 5.5 is also valid for $m \geq 1$. (the proof is identical)

(2) As argued, the upper bound given by Lemma 5.5 improves on the one given by Theorem 3.1; for example for the binomial case, we have $U_1 = \frac{2(1-p)^2}{p} \leq U_0 = \frac{1-p}{p}$ (for p=0.9, $U_1 \simeq 0.022$, $U_0 \simeq 0.111$).

(3) One should however be warned that Theorem 3.1 and Lemma 5.5 together imply <u>only</u> that "in the worst-case sense" finding $v_1$ is "better" than finding $v_0 \equiv t_1$ as a heuristic procedure for the PTSP; indeed one can

easily construct examples where $E[L_{t_1}] \leq E[L_{v_1}]$. Theorem 3.1 and Lemma 5.5 show that $v_1$ is <u>guaranteed</u> to be "closer" than $t_1$ to the optimal PTSP tour.

(4) We can derive the same kind of bound for $v_k$, $k \geq 2$. The derivation is in fact identical to the one leading to Lemma 5.5 and will be omitted. The quality (i.e. the tightness) of the upper bounds (say for $v_k$) depends very much on the quality of the established bounds between $L^{(o)}_{1,v_k}$ and $L^{(o)}_{1,t_1}$ (see (5.42) for the case of $v_1$) and one can show that for general $v_k$, we have:

$$L^{(0)}_{1,t_1} \geq \left(\alpha_0 / \left(\alpha_0 + \sum_{r=1}^{k} r\alpha_r\right)\right) L^{(o)}_{1,v_k} \tag{5.44}$$

(5) General discussion:

It is now clear that solving (5.32) for a given k (as a heuristic procedure for the PTSP) is an increasingly difficult task as k increases, but a task guaranteed (in terms of worst-case) to give a solution increasingly close to the optimal PTSP tour; the best choice in this tradeoff will of course depend on W and its p.m.f. To illustrate the discussion, consider the binomial case with a coverage probability p; then if p is close enough to 1 (for example p=0.95) finding $v_0 \equiv t_1$ (i.e., the TSP tour) would be good enough with a percentage above optimality less than or equal to $\frac{0.05}{0.95} \sim 6\%$.

Clearly for p not so close to 1 (for example p=0.8) it might be worthwhile to look for $v_1$ instead (for p=0.8 the upper bounds for the worst-case ratio for $v_0$ and $v_1$ are respectively $\frac{0.2}{0.8} \sim 25\%$ and $\frac{2\text{x}0.04}{0.8} \sim 10\%$). This is however based on relative worst case ratios so that the best strategy for this case would be to find both $v_0$ and $v_1$ and then choose the tour with the smallest expected length (note that if we register a major decrease from $E[L_{v_0}]$ to $E[L_{v_1}]$, this denotes a case where the TSP tour is not at all adequate as a solution to the PTSP). One can argue the same way for an even smaller p by considering $v_2$ and so on.

Of course the worst-case ratios assume that we solve for $v_k$ through exact method; based on our discussion of section 5.2 we know that, as soon as $k > 1$, solving for $v_k$ is not an easy task (for $v_1$ the objective function is already quadratic in $x_{ij}$, so that either the branch-and-bound approach can be applied or one of the linearizations for very small size); in fact except for $v_0$ (i.e. the TSP) for which one can count on exact procedures for up to 200 nodes, one probably has to rely on heuristics to solve for $v_k$, as well.

We shall return to this point in Section 5.3.2 where we concentrate on procedures for finding $v_1$ that can be extended easily to $v_k$ $k>1$ (sometimes at some cost in computational speed, but, depending on W, with possibly greater savings in routing). Before doing this, let us present two interesting results.

B. <u>The expected length of subtours of a graph G</u>:

In this section we would like to find the expected length in the PTSP sense of a subtour $t_1$ (that is, not including every node) of a graph G containing a black node and n white nodes; let n+1 be the black node and suppose $t_1=(n+1,1,2,\ldots,n_1,n+1)$ is the subtour ($n_1 < n$). If W corresponds to the binomial case for which each white node is present with probability p, independently of each other, it is then obvious that $E[L_{t_1}]$ is obtained using Theorem 2.2 by defining the $n_1$ quantities $L_{t_1}^{(r)}$

$r \in [0..n_1-1]$. On the other hand if W has a general p.m.f., as defined in section 2.3.4 and used throughout Chapter 3 (that is, defined on the entire set of white nodes), the problem is not as trivial and we need a more elaborate analysis:

Lemma 5.6

Given a graph G with n white nodes, one black node (say node n+1); consider a general p.m.f. for W - the number of white nodes present - ; consider a subtour $t_1$ containing only $n_1$ white nodes ($n_1 \leq n$):

$$t_1 = (n+1, 1, 2, \ldots, n_1, n+1)$$

then $E[L_{t_1}]$ can be computed atternatively as follows:

(i) obtain $Pr(W_1=k)$ [where $W_1$ is the number of present white nodes out of the set of $n_1$ nodes] from $Pr(W=k)$, then obtain the corresponding weights $\alpha'_r$ and compute $E[L_{t_1}]$ by the traditional formula

$$E[L_{t_1}] = \sum_{r=0}^{n_1-1} \alpha'_r \, L_{1,t_1}^{(r)}$$

(ii) "work directly with Pr(W=k)" and apply the following formula:

$$E[L_{t_1}] = \sum_{r=0}^{n_1-2} \alpha_r \, L_{1,t_1}^{(r)} + L_{1,t_1}^{(n_1-1)} \left( \sum_{r=n_1-1}^{n-1} \alpha_r \right)$$

Proof:

- For (i) it suffices to note that W is defined on the set of all n white nodes and, as we are concerned with only $n_1$ of them in the subtour $t_1$, we simply have to define the random variable $W_1$ corresponding to this subset of white nodes if we want to use results from Chapter 2; in fact it is easy to show that

$$Pr(W_1=j) = \binom{n_1}{j} \sum_{k=j}^{n-n_1+j} \left(\binom{n-n_1}{k-j}/\binom{n}{k}\right) Pr(W=k) \qquad \text{for } j \in [0..n_1]$$

• To prove (ii) the easiest way is to create a new graph $G'=(N',A,D')$ obtained from $G=(N,A,D)$ as follows:

First create a copy of the black node (i.e. the depot) - say n+2 - connected to all the other nodes belonging to $t_1$ (except the depot) exactly as was the original depot (with the same distances) and connected to the remaining nodes (including the depot) by arcs of length 0. Then modify D so that every distance between nodes not in the subtour $t_1$ (or between the depot and each one of those nodes) is set to 0.

Consider now the following tour through each node of G'

$$t = (n+1,1,2,...,n_1,n+2,n_1+1,...,n,n+1)$$

It is then obvious that by construction

$$E[L_{t_1}] \equiv E[L_t]$$

But now, $E[L_t]$ can be obtained by using Theorem 2.3 with m=2:

$$E[L_t] = \sum_{r=0}^{n} \alpha_r L_t^{(r)}$$

The lemma is then a consequence of:

$$L_t^{(r)} = L_{t_1}^{(r)} \qquad \forall\, r \in [0..n_1-1]$$

$$L_t^{(r)} = L_{t_1}^{(n_1-1)} \qquad \forall\, r \in [n_1..n-1]$$

$$L_t^{(n)} = d(n+1,n+2)+d(n+2,n+1) = 0$$

Q.E.D.

Note:

Lemma 5.6 can be easily extended to the PTSP with a number of black nodes greater than 1 (i.e., $m > 1$).

C. Merging subtours:

This section will be concerned with the demonstration of the following result (see Figure 5.2):

Lemma 5.7:

Given a graph G with n white nodes and one black node (say node n+1); consider three subtours of G having only the black node in common and spanning (together) every node of G; assume every node is relabeled so that the three subtours are:

$$t_1 = (n+1, 1, 2, \ldots, i_1, n+1)$$

$$t_2 = (n+1, i_1+1, \ldots, i_2, n+1)$$

$$t_3 = (n+1, i_2+1, \ldots, n, n+1)$$

Consider now the "merging" of $t_1$ and $t_2$ in a subtour $t_{12}$ as follows:

$$t_{12} = (n+1, 1, 2, \ldots, i_1, i_1+1, \ldots, i_2, n+1)$$

Then we have:

$$E[L_{t_1}] + E[L_{t_2}] - E[L_{t_{12}}] = \sum_{r=0}^{i_2-2} \alpha_r \, s^{(r)}$$

with:

$$s^{(r)} = \sum_{k=\ell}^{u} s(i_1-k, i_1+1+r-k)$$

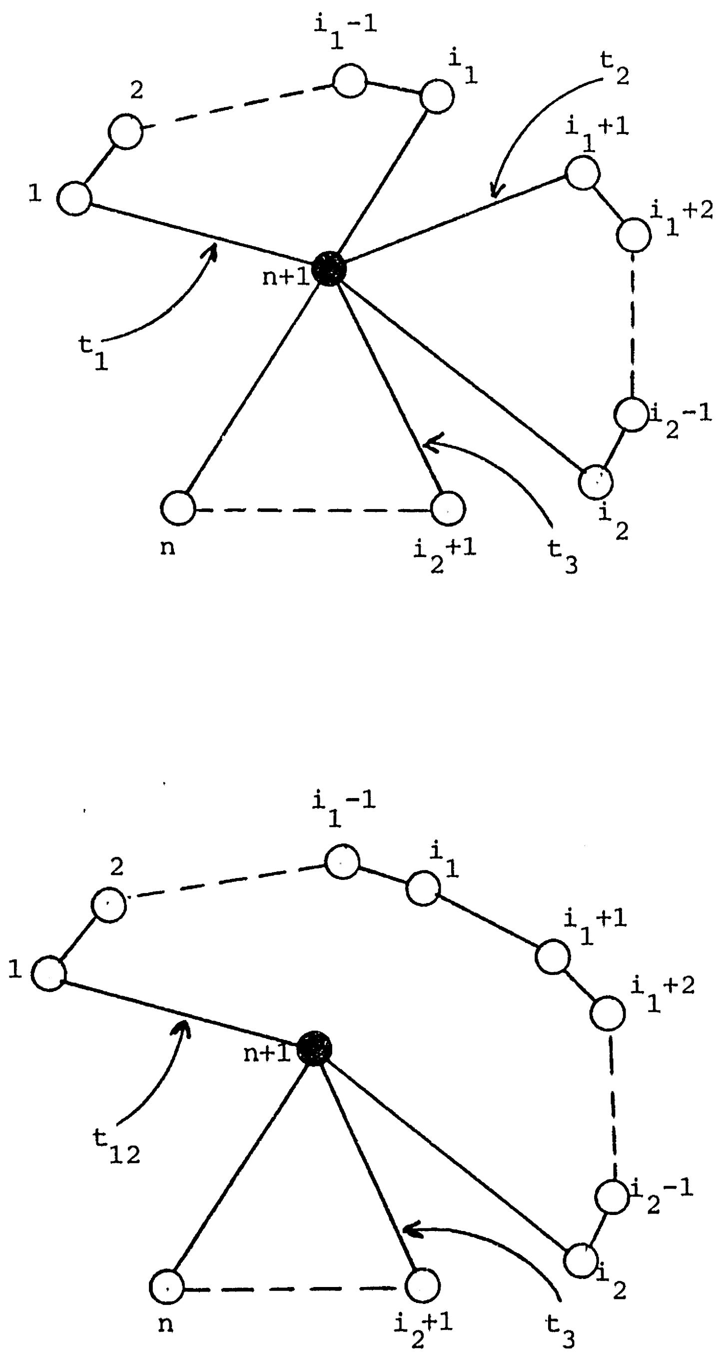



Figure 5.2: Merging of subtours

where

$$s(i,j) \equiv d(i,n+1)+d(n+1,j)-d(i,j)$$

$$\ell = \max\{0, i_1+1+r-i_2\}$$

$$u = \min\{r, i_1-1\} \qquad (\text{and } \ell \leq u)$$

Proof:

Lemma 5.7 is a direct consequence of the previous lemma (using part (ii)). Indeed we have:

$$E[L_{t_1}] = \sum_{r=0}^{i_1-2} \alpha_r L_{1,t_1}^{(r)} + L_{1,t_1}^{(i_1-1)} \sum_{r=i_1-1}^{n-1} \alpha_r \tag{5.45}$$

$$E[L_{t_2}] = \sum_{r=0}^{i_2-i_1-2} \alpha_r L_{1,t_2}^{(r)} + L_{1,t_2}^{(i_2-i_1-1)} \sum_{r=i_2-i_1-1}^{n-1} \alpha_r \tag{5.46}$$

$$E[L_{t_{12}}] = \sum_{r=0}^{i_2-2} \alpha_r L_{1,t_{12}}^{(r)} + L_{1,t_{12}}^{(i_2-1)} \sum_{r=i_2-1}^{n-1} \alpha_r \tag{5.47}$$

Now as

$$L_{1,t_{12}}^{(i_2-1)} \equiv L_{1,t_1}^{(i_1-1)} + L_{1,t_2}^{(i_2-i_1-1)}$$

we finally get from (5.45), (5.46) and (5.47) (after some calculus)

$$E[L_{t_1}]+E[L_{t_2}]-E[L_{t_{12}}] = \sum_{r=0}^{i_2-2} \alpha_r S^{(r)}$$

where $S^{(r)}$ is defined in the statement of the lemma.

Q.E.D.

Note: $S^{(o)} = s(i_1,i_1+1)$ (i.e., the traditional savings)

$S^{(1)} = s(i_1-1,i_1+1) + s(i_1,i_1+2)$ and so on

We are now in a position to present a host of procedures aimed at giving some "good" solutions to the PTSP. As already mentioned we consider the case of a graph G with one black node and n white nodes; these procedures will be developed for a general p.m.f. for W except in some cases where a binomial random variable will be assumed explicitly; also D will be assumed to be symmetric (the case of non-symmetric D being readily solvable by similar treatment).

In the first two sections we will present heuristics following a classification often adopted for the TSP (see Golden and Magnanti [1980]), namely tour construction procedures and tour improvement procedures. In a third subsection we will analyze the PTSP in the plane by briefly looking at a Spacefilling Curve heuristic and Partitioning algorithms.

### 5.3.2 Tour Construction Procedures

In this subsection we present procedures that generate an approximately optimal tour from the distance matrix as opposed to the next subsection where we mention procedures that attempt to find a better tour, given an initial tour.

#### A. Savings Approaches:

This part is a direct consequence of Lemma 5.7 and, based on the discussion of 5.3.1 A., will lead to the "Supersavings Algorithms".

The principle is based on the Clarke-Wright "savings" approach which is widely used in solving vehicle routing problems; its basic idea is simple (we give the general idea in the context of the TSP, see Clarke and Wright [1964] for other routing contexts.): suppose that, to begin with, each node is linked in a subtour containing two arcs to the depot (see Figure 5.3a). The total tour length is then $2 \sum_{j=1}^{n} d(n+1,j)$ $\left( \sum_{j=1}^{n} \left(d(n+1,j) + d(j,n+1)\right) \text{ if D is asymmetric}\right)$. If now we link two nodes - say i and j - we achieve savings of s(i,j) in travel distance equal to (see Figure 5.3b)

$$s(i,j) = d(n+1,i) + d(n+1,j) - d(i,j)$$

$(d(i,n+1) + d(n+1,j) - d(i,j)$ if D is asymmetric)

For every possible pair of nodes there is a corresponding savings; we order these savings from largest to smallest and starting from the top of the list we link nodes i and j which are end points, i.e. adjacent to n+1, with maximum savings $s_{ij}$.

The reasons for the popularity of this method is its flexibility in including various constraints during the linking process (for example constraints on maximum capacity of a vehicle, maximum distance etc.).

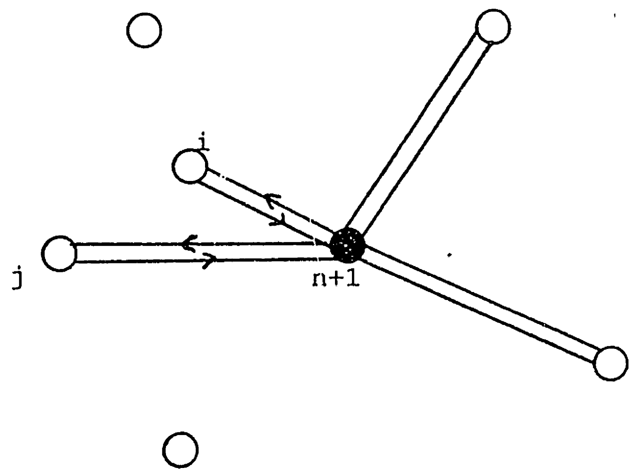
Lemma 5.7 allows us to generalize this idea to a PTSP; the results given in this lemma indicate the savings in expected length that can be achieved by linking two end-points nodes $i_1$ and $i_1+1$, i.e.:

$$\sum_{r=0}^{i_2-2} \alpha_r \, s^{(r)}$$

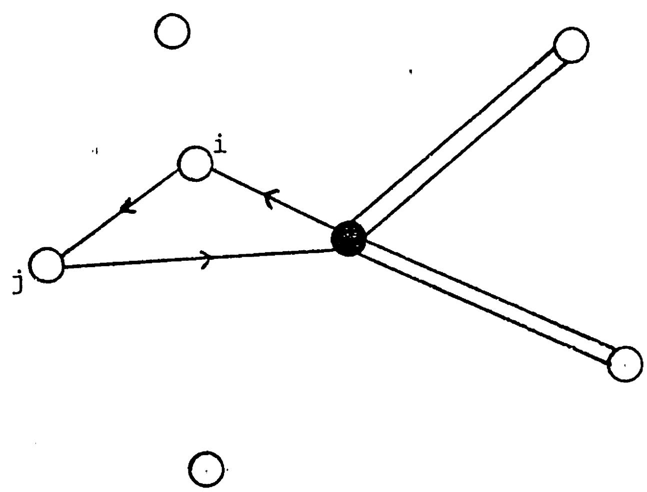
- Note that this expression reduces to the traditional C.W. savings when $\alpha_0=1$, $\alpha_r=0 \; \forall \; r > 0$ (i.e. each point always has to be visited, the case of the TSP); indeed we then obtain:

$$E[L_{t_1}]+E[L_{t_2}]-E[L_{t_{12}}]=S^{(o)} \equiv \sum_{k=0}^{0} s(i_1-k,i_1+1-k) = s(i_1,i_1+1).$$

- Also if each of the two subtours that are merged contains only one node (besides the depot) then Lemma 5.2 gives




a) initial setup




b) nodes i and j have been linked

Figure 5.3: (Tour aggregation)

$$E[L_{t_1}] + E[L_{t_2}] - E[L_{t_{12}}] = \alpha_0 \; S^{(0)} \tag{5.48}$$

$$= \alpha_0 \; s(i_1, i_1+1)$$

which again makes sense.

However to obtain an operational heuristic we need to rely on the discussion given in part A of 5.3.1; indeed as clearly indicated by Lemma 5.7 the savings in expected length obtained when linking two end points i and j do not depend on i and j only (as was the case for the TSP), but also on <u>all</u> nodes that were part of the two subtours under consideration; this prevents one from defining <u>a-priori</u> savings for pairs of nodes (since these savings depend, as well, on the other nodes contained in the subtours). In fact one would have to compute $\sum_{r=0}^{i_2-2} \alpha_r \; S^{(r)}$ for <u>each</u> possible merging; this is a formidable task since it would require first the computation of every s(i,j) ($n^2$ such elements), then for each couple of nodes (i,j) the determination of every possible savings depending on the other nodes present in the subtours containing i and j.

If on the other hand we want to find $v_1$ instead of $t_p$, the savings associated with linking two end point nodes will be (from Lemma 5.7):

$$\alpha_0 \; S^{(o)} + \alpha_1 S^{(1)} \tag{5.49}$$

where $S^{(o)} = s(i,j)$

and $S^{(1)} = s(i-1,j) + s(i,j+1)$

where i-1 is the node preceeding i on its subtour

and j+1 is the node following j on its subtour

[<u>Note</u>:

If i-1 is in fact the depot then, as:

$$s(n+1,j) = d(n+1,n+1) + d(n+1,j) - d(n+1,j) = 0$$

$S^{(1)}$ reduces to $S^{(1)} = s(i,j+1)$]

The idea is then the following:

(1) We start by forming subtours containing, in addition to the depot (that belongs to all subtours), exactly two nodes; we will then have $\lfloor \frac{n}{2} \rfloor$ such subtours and possibly one additional subtour containing a single node if n is odd. (A possibly way to form these subtours is to use the Clarke-Wright savings algorithm by imposing the additional constraints, during the linking process, that we cannot form subtours of more than two nodes.)

(2) Once each node (except possibly one) has been associated with a "companion" we are in a position to compute what we call the supersavings which are (for linking node j and node k):

$$S_{ij,k\ell} = \alpha_0 \, s(j,k) + \alpha_1(s(i,k) + s(j,\ell)) \tag{5.50}$$

For each pair of nodes (i,j) and (k,$\ell$) we need to compute four such savings when D is symmetric (i can be linked to k or $\ell$, the same being true to j).

(3) We order then these supersavings from largest to smallest and starting at the top of this list we form larger subtours by linking nodes until a tour is formed.

Let us derive the number of computations involved in this procedure: The calculation of s(i,j) requires about $cn^2$ operations for some constant c; their ordering requires $cn^2 \ell gn$ comparisons and interchanges; the last step of forming pairs involves at most $n^2$ operations; hence the first step of our algorithm requires $O(n^2 \ell gn)$ computations. The second and third steps together will in fact require the same order of computations so that the overall procedure requires on the order of $n^2 \ell gn$ computations, hence the same as for the Clarke-Wright algorithm for the TSP.

Of course we have many other possibilities based on the same principle: for example, instead of forming, in the first step, subtours with two nodes, one could form subtours with three nodes and then compute in the second step the following supersavings (again based on Lemma 5.7):

$$S_{ijk,\ell gr} = \alpha_0 s(k,\ell)+\alpha_1(s(j,\ell)+s(k,g))+\alpha_2(s(i,\ell)+s(j,g)+s(k,r)) \tag{5.51}$$

(for each merging of two "triplets" ijk and $\ell$gr we need to consider again only four savings since only i,k,$\ell$, and r can be linked). The overall computational effort is still of order $n^2 \ell gn$. (The first step would still be accomplished through the Clarke-Wright savings algorithm by requiring subtours of three nodes). One can also consider first the formation of subtours containing four nodes or more.

The tradeoff is the following: forming larger clusters of nodes during the first step (based only on s(i,j)) "penalizes" the PTSP savings criteria, but on the other hand, in the second step this allows us to compute more accurate supersavings.

Note that if one wishes to form subtours of 4 nodes in the first step, one can obtain better results by breaking this first step into two steps as follows:

(1) form subtours of two nodes using the C.W. savings approach.

(2) define supersavings as before; then, based on those supersavings, form subtours of four nodes.

In conclusion the best strategy for solving the PTSP according to savings criteria would be as follows: (The Supersavings Algorithm)

(A) Based on the discussion of part A section 5.3.1 determine, according to the p.m.f. of W (and the corresponding weights $\alpha_r$), k* such that $v_{k*}$ is a good approximation for $t_p$.

(B) To determine the best tour corresponding to $v_{k*}$ choose the best among the k*+1 tours obtained as follows:

(1) heuristic tour for $v_0$: use Clarke-Wright savings algorithm

(2) heuristic tour for $v_1$:

- form subtours of two nodes (using C.W. savings alg.)
- compute supersavings
- form the tour based on these supersavings

(3) heuristic tour for $v_2$:

- form subtours of three nodes (using C.W. savings alg.)
- compute supersavings
- form the tour based on these supersavings

(4) heuristic tour for $v_3$:

- form subtours of two nodes (using C.W. savings alg.)
- compute supersavings
- form subtours of four nodes (based on previous supersavings)
- computer supersavings
- form the tour based on the last supersavings

etc.

This overall procedure has $O(k^*n^2\ell gn)$ computations.

B. "Almost" Nearest Neighbor Algorithms:

The tour construction procedure developed in this section is again closely related to a similar heuristic for the TSP usually referred to as the Nearest Neighbor algorithm; Let us briefly describe for completeness the Nearest Neighbor method. Under this algorithm a tour is constructed as follows:

1. Start with an arbitrary node.

2. Find the node not yet on the path which is closest to the node last added and add to the path the arc connecting these two nodes.

3. When all nodes have been added to the path, join the first and last nodes.

This algorithm requires on the order of $n^2$ computations and if the method is repeated every node of the graph as a starting node, the overall procedure would run in an amount of time proportional to $n^3$.

It will become clear later that one useful way of interpreting the approach of this heuristic is the following: Suppose we are given a path containing j nodes in addition to the depot - say (n+1,1,2,...,j). If we add a node - say j+1 - to this path by adding the arc (j,j+1), then the increase in length will be d(j,j+1). The Nearest Neighbor algorithm consists simply of choosing the minimum increase in length possible at each step.

By finding the increase in expected length (in the PTSP sense) of adding a node to a path, we can then define a strategy based on the minimization of this expression. This rationale leads to algorithms that we will call, for obvious reason, Almost Nearest Neighbor algorithms.

Here again, the discussion in part A of 5.3.1 implies that schemes with increasing computational requirements but also improved performance can be developed:

Let us first look at the optimization problem corresponding to $v_1$; by adding node j+1 to the path (n+1,1,2,...,j) the increase in the objective function is:

$$\alpha_0 d(j,j+1)+\alpha_1 d(j-1,j+1) \tag{5.52}$$

since the increase will be equal to d(j,j+1) for $L^{(o)}$ and to d(j-1,j+1) for $L^{(1)}$.

More generally the increase in the objective function corresponding to $v_k$ will be:

if $k < j$

$$\alpha_0 d(j,j+1)+\alpha_1 d(j-1,j+1)+\ldots+\alpha_k d(j-k,j+1)$$

(5.53)

if $k > j$

$$\alpha_0 d(j,j+1)+\ldots+\alpha_{j-1} d(1,j+1)+\alpha_j d(n+1,j+1)$$

Based on previous results one can then propose a procedure as follows:

(1) choose k* as before (see savings methods)

(2) starting from node n+1 (i.e., the depot) add the nearest node

(3) use the minimization of increase in expected length as a criterion as given by (5.53) to add successive nodes until the tour is formed.

Note: As before the choice of k* depends strongly on W and the number of nodes n.

In fact we are going to "improve" on this procedure on the basis of the following fact:

For a given tour, each arc belonging to the tour is weighted equally (in fact each weight is 1) with respect to the TSP problem (i.e., minimization of $L_t^{(o)}$). For the PTSP with no black node, this is again true (the weight being $\alpha_0$). However as shown in Chapter 3 this is not true anymore when the number of black nodes m>0 (see Appendix A). For m=1 (our concern, here) we know, see proof of Fact 3.2, that for a given tour (n+1,1,2,3,...,n,n+1) (n+1 being still the black node), d(n+1,1) will appear in each $L_t^{(r)}$ $r \varepsilon [0..n-1]$, and its weight will be $\sum_{r=0}^{n-1} \alpha_r$. The same is true of d(n,n+1). On the other hand d(j,j+1) for $j \varepsilon [1..n-1]$) has weight $\alpha_0$ only.

From this observation one can give a modified version of the previous method in which:

- step (2) is replaced by: (2)' starting from node n+1 add its nearest node and call it node 1 so that we have d(n+1,1); then find the second nearest node, call it n and connect n to n+1.

- and (3) is replaced by: (3)' proceed as in 3 in order to connect 1 to n.

(The idea is that since the two arcs adjacent to the depot have more weight in the total objective function, one should minimize them first).

One can extend this modified version to the problem of finding $v_k$ for a general k; here again one can proceed first by "growing two trees" out of the depot until both of them contain k nodes, then by joining the two trees according to the previous method (i.e. step (3)' based on 5.53).

C. Insertion methods:

Before concluding this section on tour construction procedures let us briefly mention that, again, one can extend the rationale of a number of insertion methods, proposed for the TSP, to the PTSP. A good description of these insertion methods - Nearest Insertion, Cheapest Insertion, Arbitrary Insertion, Farthest Insertion, - is given in Golden et al. [1980]. Such a method takes a subtour on j nodes at iteration j and attempts to determine which node not in the subtour should join the subtour next (the selection step) and then decides where in the subtour it should be inserted (the insertion step). The reader is referred to Golden et al. [1980] for details on the different insertion schemes; for our purpose one should simply notice that the insertion step (and sometimes the selection step) is based on the consideration of the already introduced notion of "savings" associated with the insertion, i.e., if node $\ell$ is to be included in between nodes i and j this will produce an increase in the length of:

$$d(i,\ell) + d(\ell,j) - d(i,j)$$

The following fact simply gives the increase corresponding to $v_k$.

Fact 5.3

Let G=(N,A,D) be a graph with one black node and n white nodes and $t = (n+1,1,\ldots,n_1,n+1)$ a subtour containing $n_1$ white nodes. Assume we next include node j between node i and i+1. We then have, calling this new subtour t':

$$E[L_{t'}]-E[L_t] = \sum_{r=0}^{n_1-1} \alpha_r\left(L_{t'}^{(r)} - L_t^{(r)}\right) + \left(\sum_{r=n_1}^{n-1} \alpha_r\right) 2d(n+1,j)$$

Based on this fact and on the discussion in part A of 5.3.1 one can obtain a host of insertion procedures for our problem.

5.3.3 Tour improvement procedures

As indicated in Golden et al. [1980], branch exchange heuristics are among the best-known heuristics for the symmetric TSP. They work as follows:

1. Find an initial tour - Generally, this tour is chosen randomly from the set of all possible tours.

2. Improve the tour by a $\ell$-change.

3. Continue 2 until no additional improvement can be made.

A $\ell$-change of a tour consists of the deletion of $\ell$ arcs in the tour and their replacement by $\ell$ other arcs to form a new tour. (A tour is then called $\ell$-optimal if it is not possible to improve the tour via a $\ell$-change). One of the most effective implementations of this idea is given in Lin and Kernighan [1973].

In fact tour improvement procedures belong to a general approach to solve combinatorial optimizations problems sometimes referred to as "Local Search" (see Papadimitriou and Steiglitz [1982]). One can possibly think of similar procedures for the PTSP, although they would certainly not be as effective (or fast) as for the TSP; indeed in the case of the TSP one can check immediately if exchanging 2 arcs by 2 other arcs leads to a decrease in the objective function (if $(i,j)$ and $(k,\ell)$ are exchanged by $(i,k)$ and $(j,\ell)$ we have an improvement if $d(i,k) + d(j,\ell) < d(i,j)+d(k,\ell)$, assuming D is symmetric). For the same case under the PTSP we have to recompute the entire expected length of the new tour ($O(n^2)$ step). One can, of course, based as before on the discussion contained in part A of 5.3.1, apply these tour improvement procedures to solve for $v_k$ (see (5.32)). For example assume that we want to find $v_1$ - i.e. the tour that minimizes $\alpha_0 L^{(o)}_{1,t} + \alpha_1 L^{(1)}_{1,t}$ - through a 2-change procedure; assume the initial tour t is given by $t=(n+1,1,2,3,..,n,n+1)$ and that $(i,i+1)$ and $(j,j+1)$ $(1\leqslant i\leqslant n-3, i+2\leqslant j\leqslant n-1)$ are replaced by $(i,j)$ and $(i+1,j+1)$ (note that this is the only possibly exchange; the choice of $(i,j+1)$ and $(i+1,j)$ would break the initial tour in two subtours). The new tour t' is then such that:

$$L^{(o)}_{1,t'} - L^{(o)}_{1,t} = d(i,j)+d(i+1,j+1)-d(i,i+1)-d(j,j+1)$$

$$\begin{aligned} L^{(1)}_{1,t'} - L^{(1)}_{1,t} &= d(i-1,j)+d(i,j-1)+d(i+1,j+2)+d(i+2,j+1) \\ &\quad - d(i-1,i+1)-d(i,i+2)-d(j-1,j+1)-d(j,j+2) \end{aligned}$$

and it will be accepted only if

$$\alpha_0 \left(L^{(o)}_{1,t'}-L^{(o)}_{1,t}\right) + \alpha_1 \left(L^{(1)}_{1,t'}-L^{(1)}_{1,t}\right) \leqslant 0 \qquad (5.54)$$

This gives a procedure to approximate $v_1$ that can be carried out in $O(n^2)$ computations; by extending this idea to obtain an approximate solution for $v_k$ by a $\ell$-change one has a procedure that is an $O(kn^{\ell})$ polynomial algorithm (hence to approximate the PTSP, the procedure is of order n more costly than for approximating the TSP).

This concludes the discussion of heuristic methods for solving the PTSP on a given graph G; we based most of our results on the theoretical investigations of Chapters 2,3. Before concluding Chapter 5 let us now look at the case of the PTSP in the plane.

### 5.3.4 The PTSP in the plane

This section is basically inspired by Chapter 4 and as such, the notation and terminology used in that chapter are assumed known and will not be repeated here (see especially section 4.1).

We already mentioned in Chapter 4 the great interest of results like that of Beardwood et al. [1959] for algorithmic applications (i.e., development and probabilistic evaluation of heuristic methods for combinatorial optimization problems (see Karp [1976] and Karp [1977]). For our purpose Theorem 4.2 and its generalization Theorem 4.4 provide the results necessary for a probabilistic analysis of heuristics.

This section contains two parts: in the first part we first briefly present a heuristic based on spacefilling curves that has been recently proposed to solve the TSP in the plane (Platzman and Bartholdi [1983]); we then discuss the main characteristics that may make this heuristic a good candidate for solving the PTSP in the plane. Finally in a second part we mention the possibility of using partitioning algorithms for solving large size PTSP problems in the plane, in much the same way they have been proposed for the TSP in a seminal paper by Karp (Karp [1977]).

#### A. The Spacefilling Curve Heuristics:

We will not present the heuristic in detail; the interested reader is referred to the original paper (Platzman and Bartholdi [1983]) for a detailed treatment of the subject.

In general, spacefilling curves are by definition continuous mappings $\psi$ from the closed unit interval [0,1] onto a set A of dimension 2 or more; now, as pointed out in Platzman and Bartholdi [1983], if the mappying $\psi$ is such that $\psi(0) = \psi(1)$, then $\psi(\alpha)$ ($\alpha \in [0,1]$) traces a "tour" of all the points of A as $\alpha$ varies from 0 to 1. The general idea is then as follows: given such a spacefilling curve $\psi$ and given a sequence $x^{(n)}$ of n points in A, construct a tour by sequencing the points as they appear along the spacefilling curve; that is, for each point $x_j$ of the sequence $x^{(n)}$ of A to be visited, compute a $\alpha_j$ such that $\psi(\alpha_j)=x_j$ and then sort the points by their corresponding $\alpha_j$'s. (This idea corresponds to solving the TSP in [0,1] instead of A.)

Platzman and Bartholdi define a particular spacefilling curve over the unit square $[0,1]^2$ and, showing how its inverse is computed, they provide a specific heuristic to solve the TSP in the plane; this curve is constructed recursively by dividing the square into four identical subsquares, and filling each with a spacefilling curve that we link to form a circuit. As noted by the authors this construction resembles (in its principle) the partitioning algorithm of Karp (Karp [1977]).

Let us now list, using our notation the main interesting points of this specific heuristic (obtained in Platzman and Bartholdi [1983], refer to that paper for proofs); before, let us simply mention that the heuristic tour obtained by this spacefilling curve will be denoted as $t_f$:

(1) This heuristic, since it is based on <u>sorting</u>, has the desirable property (already met by most of our constructions in Chapter 4) that given any subset $s_j$ of $x^{(n)}$ ($j \in [1..2^n]$) the construction of $t_f(s_j)$ through the spacefilling curve heuristic gives the <u>same</u> tour as the resulting process of skipping in the PTSP sense the points of $(x^{(n)}-s_j)$ from $t_f(x^{(n)})$. (see property (a) given in the proof of Lemma 4.1)

(2) The heuristic requires $O(n\ell gn)$ operations, hence it is very fast.

(3) for any arbitrary sequence of points x in $[0,r]^2$

$$L_{t_f}(x^{(n)}) \leq 2\sqrt{n}\, r$$

(4) for a uniform sequence of points X over $[0,1]^2$

$$\limsup_{n\to\infty} \frac{E_u[L_{t_f}(X^{(n)})]}{\sqrt{n}} \leq \sqrt{\pi}$$

$$\liminf_{n\to\infty} \frac{E_u[L_{t_f}(X^{(n)})]}{\sqrt{n}} \geq \beta$$

(where $\beta$ is the TSP constant)

(5) it has been estimated that, actually,

$$\limsup_{n\to\infty} \frac{E_u[L_{t_f}(X^{(n)})]}{\sqrt{n}} \leq 1.006$$

$$\liminf_{n\to\infty} \frac{E_u[L_{t_f}(X^{(n)})]}{\sqrt{n}} \geq 0.906$$

(The authors indicate that the tour will be approximately 25% longer than the optimal tour.)

Property (1) is of course very interesting with respect to the PTSP. Based on this property one can analyze $t_f$ very easily and our results are summarized in the following lemma:

Lemma 5.8

Let x be an arbitrary sequence of points in $[0,1]^2$ and p be the coverage probability for each point. Then the expected length of $t_f(x^{(n)})$ (in the PTSP sense) satisfies:

(i) $EL_{t_f}(x^{(n)},p) \leq 2E[\sqrt{W}]$

If the sequence is independently and uniformly distributed over $[0,1]^2$

$$\text{(ii)} \quad \limsup_{n\to\infty} \frac{E_u[EL_{t_f}(X^{(n)},p)]}{\sqrt{n}} \leq 1.006 \sqrt{p}$$

$$\liminf_{n\to\infty} \frac{E_u[EL_{t_f}(X^{(n)},p)]}{\sqrt{n}} \geq 0.906 \sqrt{p}$$

Proof:

A consequence of properties 1,3, and 5 together with similar proofs as in Lemma 4.1, Lemma 4.3 and section 4.3.3.

Consequences:

(1) (i) does not improve on the result given in Lemma 4.1 for the upper bound on the expected length of the optimal PTSP tour; nevertheless it is a worst-case guarantee for the expected length of the heuristic tour.

(2) From Corollary 4.2 and Lemma 4.4 one can induce that the spacefilling curve heuristic tour is approximately within 25% of optimally for points uniformly distributed over $[0,1]^2$.

(3) As already discussed in Chapter 4 and because of similar generalities obtained in Platzman and Bartholdi, point (2) above remains valid for any Lebesque set of dimension 2.

In conclusion this analysis points to the spacefilling curve heuristic as an interesting solution procedure for the PTSP especially because of its speed. One, however, has to realize that this heuristic gives the same tour no matter what the p.m.f. of W is; in other words the determination of the tour is W-independent and this is certainly not very desirable compared to our previous procedures in which the specific p.m.f. of W was taken into account. In summary this heuristic is interesting for obtaining very quickly a tour with some interesting, although limited properties.

Let us conclude this section on the PTSP in the plane by briefly mentioning the possibility of using partitioning algorithms, as well.

B. Partitioning Algorithms:

Following the paper by Karp (Karp [1977]), in which a partitioning algorithm for the TSP in the plane has been devised, analyzed probabilistically, and tested, there has been a proliferation of similar analyses for other combinatorial optimization problems (see for example Haimovitch and Rinnooy Kan [1983], Fisher and Hochbaum [1980], Hochbaum and Steele [1982], Papadimitriou [1978], etc.)

Although the practicability of some of those procedures has been questioned (see Psaraftis [1984]) the partitioning algorithm devised by Karp for the TSP has been successful in solving very-large problems. From a practical point of view, this procedure is a decomposition algorithm for which the main idea is to partition a large rectangle (in which the points to be visited are lying) into a number of subrectangles and solve a TSP in each subrectangle. Let us briefly give an outline of the procedure:

1. Let n be the number of nodes to visit and t an upper bound on the number of nodes in a subrectangle. Define $k=\lceil \ell g((n-1)/(t-1))\rceil$.

2. Divide the original rectangle into subrectangles using the following strategy: iteratively divide the subrectangles (containing,

say, m nodes) at the $\lceil \frac{m}{2} \rceil$th node along the long side of the rectangle until we get $2^k$ subrectangles, each containing no more than t of the nodes.

3. Solve the TSP in each subrectangle.

4. Convert the Eulerian cycle obtained in 3 to a Hamiltonian cycle by virtue of the triangle inequality.

If step 3 is solved exactly this procedure can be shown to be asymptotically optimal almost surely in the sense that: For any $\varepsilon>0$ there is an algorithm $A(\varepsilon)$ based on partitioning such that

(i) it runs in time $C(\varepsilon)n + O(n \ell gn)$ where $C(\varepsilon)$ is a constant

(ii) with probability 1 this algorithm yields a tour costing not more than $(1+\varepsilon)$ times the cost of an optimal tour.

If step 3 is solved by a heuristic this procedure remains a valid one even for extremely large problems.

It is easy to see that one can adopt the same procedure for the PTSP, replacing TSP by PTSP in step 3. Based on the results of Chapter 4 one should be able (using the same analysis as in Karp [1977], and using Theorem 4.2 in the place of the Beardwood et al.'s theorem ) to show that this procedure is asymptotically optimal if step 3 is solved exactly. Using heuristic procedures proposed in the previous sections one should be able to handle nicely quite large problems through this generic procedure.

## 5.4 Conclusion

This chapter has presented some of the algorithmic consequences of the set of theoretical investigations carried out in Chapters 2,3, and 4. It contains results that are not only based on intuitive considerations but on theoretical foundations as well.

We have concluded, that, although exact methods seem a little too ambitious for a general PTSP, a Branch-and-Bound procedure presented in section 5.2.3 could be appropriate for solving reasonable-size PTSP problems that are in some sense "close" to the traditional TSP problems (e.g., a coverage probability p close to 1, or a dominant number of black nodes).

It is however toward heuristic methods that we should turn our attention for problems of practical size. With respect to those procedures we presented a considerable variety by extending the rationale behind the development of similar procedures for the TSP. The framework introduced in section 5.3.1 A. together with some tour construction procedures seem the more appealing and promising directions. For very large problems partitioning algorithms should also be considered seriously. Finally if speed (and not so much accuracy) is sought the spacefilling curve heuristic should provide reasonable tours.

We conclude by making a somewhat intuitive and positive point on procedures based on partitioning (such as Karp's, or even some construction tours proposed in Chapter 4). These are such that they eliminate the possibility of obtaining tours similar in shape to the star-shaped example and thus eliminate tours that are, although of reasonable length, behaving badly in terms of the PTSP. (by similar in shape we mean a tour whose "interior" (shaded) has the following form; see Figure 5.4).

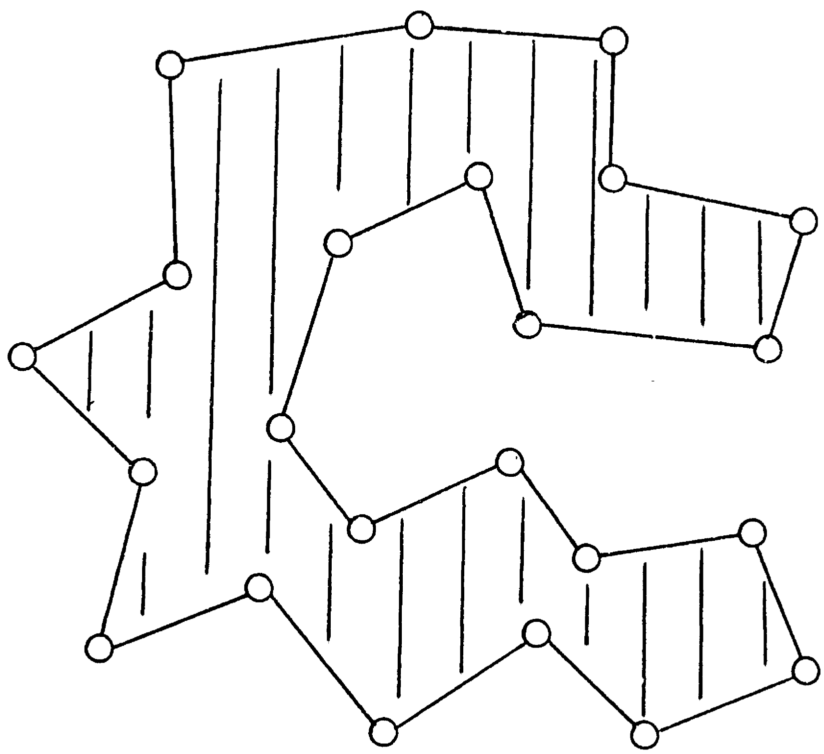


Figure 5.4: "Non-Desirable" Type of Tour

# CHAPTER 6

# FINAL REMARKS AND CONCLUSIONS

In this last chapter of the thesis we present first a summary of the principal contributions of this dissertation and then list topics for further research.

## 6.1 Summary of Results

### 6.1.1 Methodological (Theoretical) Contributions

Perhaps this summary should start with the initial contribution of this work; namely the definition and formulation, in Chapter 1, of the PTSP, a problem which provides a conceptualization of many practical situations likely to be encountered in various forms in several application areas. Following introduction of the PTSP, we then presented theoretical and algorithmic results for this new problem. We first reduced in Chapter 2 the computational requirements for obtaining the expected length of any tour through n points from $O(n2^n)$ (naive enumeration) to $O(n^2)$ via the derivation of closed-form expressions. Based on this fundamental chapter, we then derived properties of optimum PTSP tours together with useful bounds on their expected lengths. By comparing the TSP and PTSP we have also shown how the introduction of probabilistic elements into well-known combinatorial optimization problems may change drastically the behavior of their optimal solutions and of their combinatorial properties.

Next, motivated by theoretical results like that of Beardwood et al. [1959] for algorithmic applications, we presented an analysis of the PTSP in the plane. The main theoretical results include bounds on the expected length of the optimal PTSP tour through n points lying in a square, asymptotic analysis of the "strategy of reoptimizing", and finally, perhaps most importantly, the following strong limit law: the expected length of the optimal PTSP tour through n points drawn from a uniform distribution in the unit square is almost surely (with probability 1) asymptotic to a constant times $\sqrt{n}$; this result being valid for any general subset A of a d-dimensioned Euclidean space of measure v(A) and different metrics, $\sqrt{n}$ being replaced by $v(A)^{1/d}\, n^{(d-1)/d}$.

Based on the theoretical results obtained in Chapters 2, 3, and 4 we finally suggested several solution procedures to solve the PTSP. After providing several mathematical programming formulations for the problem, we proposed a branch-and-bound scheme believed to be one of the best exact methods for tackling this specific problem; we also demonstrated the inadequacy of dynamic programming approaches for the PTSP. Recognizing the limitations of exact optimization methods for a complex problem like the PTSP, we also presented a variety of heuristic procedures. Based on the previous theoretical work and on the analysis of the rationale behind some heuristics proposed for the TSP, we have been able to design, using similar reasoning, several heuristics including: tour construction procedures (Supersavings algorithm, Almost Nearest Neighbor algorithm); "hill-climbing" methods; and partitioning algorithms including a space-filling approach.

### 6.1.2 Discussion

We turn next to a discussion of the contributions of this work. This section is divided in two parts: modeling first and applications next.

A. Modeling:

As already discussed in Chapter 2 the results obtained in this thesis cover situations in which we have m black nodes (i.e., depot or "always present" customers), n white nodes (i.e., "uncertain" customers) with a general p.m.f. for W - the number of white nodes present. We also discussed the possibility of using our results for instances in which we have special cases of node-specific probabilities; that is, cases where each white node j has a coverage probability $p_j$, independently of each other, $p_j$ being of the form $1-(1-p)^{k_j}$ for a common p ($k_j$ representing the number of superimposed nodes with coverage probability p).

For the <u>general</u> node-specific probability cases, we mentioned the use of recursive relationships to compute the expected length of tours. In fact by choosing p sufficiently small, one can see that $f(k) = 1-(1-p)^k$, $k = 0,1,2,..$ can approximate (arbitrarily closely) any given set of probabilities $p_j$, $j \in [1..n]$. This, together with the superimposing trick, imply that all the results (except for the asymptotic analysis of Chapter 4), obtained in this dissertation, can be used for cases with general node-specific probabilities.

In conclusion the different cases covered by the results of this thesis offer a wide range of possibilities for modeling purposes.

B. Applications:

The PTSP has been introduced in the context of tactical routing, that is, in the context of <u>designing</u> of optimal routes under some uncertainty. This is, however, not the only context in which the PTSP methodology can be used. Let us present two other areas of application:

(i) "Strategic" Routing: Preliminary Planning Of Distribution Systems

In preliminary planning of urban collection and delivery systems we are very often interested in "sizing up" the requirements for vehicle fleets, estimating the number of points that can be served with given resources, etc. Those decisions are usually embedded in a hierarchical planning process (as a first stage). The following situation arises very frequently in practice: given a region, we have information on the set of all potential customers for a specific service (the total number of potential customers and their locations, but we know that on a daily basis only a fraction of those customers will have to be served; the problem is to obtain an estimate of the routing costs a company would face if it decided to implement such a service. (This is a strategic decision as opposed to the tactical decision of effectively routing the vehicles that have been purchased.)

The usual way of handling this problem is by using an approximation formula based on the Beardwood et al.'s theorem (already encountered many times along the development of the thesis; see, for example, Chapter 1, Section 1.2.1). More specifically, assuming that the region has area A, and that the average number of customers we would expect on a daily basis per vehicle is n, we use $0.765\sqrt{nA}$ to obtain an estimate of the routing distance for a vehicle on a daily basis.

This approach, however, has major drawbacks:

* n has to be sufficiently large so that the asymptotic result obtained in Beardwood et al. gives a good approximation.

* the distance between points has to follow the euclidean metric; in urban areas this is often far from true, since the distances may not

even be symmetric due to one-way streets, etc. (Of course Beardwood et al.'s theorem can be extended to the rectilinear or other metrics, but then the constant is different and one would need an estimate of it).

* the customers have to be "nicely" distributed within the area (i.e., independently and identically distributed over the area considered). A consequence of this is that, to be valid, the use of the asymptotic formula necessarily imposes the condition that in estimating n, one must assume independence between potential customers and an identical behavior with respect to their being present or not.

* finally this formula does not provide any "guarantees" (lower or upper bounds) regarding the eventual size of the actual routing costs.

However, by defining a probability mass function on the actual number of customers to be visited on a specific day, this problem can be modeled as a PTSP on the set of all potential customers. Hence by determining a good PTSP tour (see Chapter 5 for solution procedures) through all the customers, we can then use expressions from Chapter 2 to compute the expected length of this tour; this, in turn, provides a good upper bound on the size of actual routing costs the company will, on the average, face every day. Moreover, this method does not impose any of the restrictions necessitated by the traditional one. In other words, as previously discussed, this approach can be used for cases where:

- the probability of coverage is not identical among customers
- there is possible dependence between the presence of these customers
- the distance matrix can be arbitrary (not necessarily following any particular metric)
- the number of customers can be arbitrary.

(ii) Location of Facilities

We assume that the reader has some familiarity with facility location problems (see Larson and Odoni [1981] for an excellent introduction to such problems). We place our discussion, here, in the context of one of the most basic of these problems, namely, the problem of locating a single new facility so as to serve a set of n potential customers (most of the comments remain, however, valid for more general problems as well).

The traditional approach is to locate this single facility under the assumption that each time a customer places a request, he will be served exclusively by one vehicle who visits this customer and then comes back to the facility immediately; in other words we minimize the average distance to or from the facility for the population of users considering "radial" distances (that is, a weighted average of the distances between the facility and each of the customers).

One can, however, find many practical situation where vehicles will serve more than one customers per trip. The problem of locating a facility under this condition has been introduced in Burness and White [1976]; they considered the rather "formidable" task of locating the facility assuming the construction of a traveling salesman tour through each possible subset of the original n customers. The optimal location required an iterative procedure (whose convergence was not proved) during which at each step $2^n-1$ traveling salesman problems had to be solved!

The merging of a highly strategic decision (location of a facility) and highly tactical decision (the design of a TSP for each subset of customers) is very ambitious and might not in fact reflect adequately the different nature of strategic and tactical decisions.

Instead of this, one can see that an alternative approach is to design an a-priori tour through all n customers assuming the PTSP methodology, that is to locate the facility in order to minimize the expected length (in the PTSP sense) of an a-priori tour through all potential customers.

## 6.2 Further Research

"...a highly finished paper, with all its theorems carefully proved, all avenues explored, and all loose ends carefully snipped, may arouse one's admiration; but its very perfection drains it of vitality, and there is little one can do with it except file it. Papers are more entertaining if they are still rich in conjectures.." (Hammersley [1972]).

We hope that we have convinced the reader that a rigorous mathematical analysis of combinatorial optimization problems containing probabilistic elements can be performed, but that these problems deserve very often special treatment due to the drastic change of their properties when uncertainty is introduced.

One can add that the topics explored in this thesis appear to be so rich in fruitful possibilities that a section on subjects for further research is mandatory. Let us briefly list some important directions:

### 6.2.1 Algorithmic Implementation

It is important to implement the different solution strategies given in Chapter 5. Besides the obvious reason of comparing the various solution procedures proposed, this will allow one to obtain an estimate (via the solution of a range of problems) of the constant c(p) introduced in Chapter 4.

### 6.2.2 Theoretical Investigation

There are several theoretical open-problems which were encountered during the development of the thesis that need to be addressed in the future:

(A) Chapter 3:

- Lemmas 3.4 and 3.5:

We conjecture but have not proved that the lower bound on $E[L_t]$ for m=0 and n not prime is the same as the one obtained for n prime.

- Section 3.4:

Prove or disprove our conjecture on the worst-case behavior of an optimum TSP tour when used as a solution to the corresponding PTSP problem (see Lemmas 3.10 and 3.11 in particular).

(B) Chapter 4:

- It is easy to see that Theorems 4.1 and 4.3 can be generalized to the case in which the points are independently and identically distributed (and not necessarily uniformly) over $[0,1]^2$.

- To show that this is true also for Theorems 4.2 and 4.4 one has to verify additional properties (given in Steele [1981a]) for $\phi_p$. We have already proved that $\phi_p$ verifies properties A6 and A7 of Steele's paper but have not been able to prove or disprove A8, the final step.

(C) Chapter 5:

The investigation of algorithms has been of an introductory nature and more has to be done in terms of exact procedures and design of heuristics. A deeper analysis of the relationships among the intermediate problems (i.e., finding $v_k$) introduced in section 5.3.1.A should be undertaken.

### 6.2.3 Applications of the PTSP Methodology

It is important to investigate in more detail uses of the PTSP methodology in different areas of application such as preliminary planning of distribution systems and location of facilities.

### 6.2.4 Further Research on the Integration of Uncertainty Into Combinatorial Optimization Problems

Let us simply mention two other routing problems whose probabilistic treatment might be conducted along lines similar to those developed in this thesis:

A. The m-PTSP:

The m-PTSP is the natural extension of the PTSP (as the m-TSP was for the TSP) corresponding to the design of a prespecified number m, of distinct tours that collectively visit each of the nodes in a set exactly once, while having a black node (depot) in common. The objective is to minimize the expected total length (in the PTSP sense) covered by the m tours.

B. The Probabilistic Vehicle Routing Problem (PVRP)

This problem can be formulated as follows: consider a standard VRP (see Chapter 1) but with demands which are probabilistic in nature, rather than deterministic; the problem then is to determine a <u>fixed</u> (before realization of the random variables) set of routes of minimal expected total distance; this expected total distance corresponds to the total distance of the fixed set of routes (as in the deterministic case) plus the expected value of the <u>extra</u> distance that might be required by a particular realization of the random variables. The extra distances will be due to the fact tha. demand on some routes may occasionally exceed the capacity of the vehicles assigned to these routes and force vehicles to go back to the depot before continuing on their routes. This problem differs from the SVRP introduced in Chapter 1 (see Stewart and Golden [1983]) in the sense that here we are concerned only with routing costs without the introduction of additional parameters. We have already made some progress in analyzing this problem by showing that randomness can introduce surprising complications that go well beyond the additional computational complexity resulting from the need to consider multiple-valued variables (hence rendering doutful direct applications of techniques developed for the VRP).

A safe general conclusion would be to assert that much ha yet to be discovered in this very exciting and almost unexplored area of research.

## REFERENCES


Aho, A., J. Hopcroft, J. Ullman, (1974), The Design and Analysis of Computer Algorithms, Addison-Wesley, Reading, MA.

Aho, A., J. Hopcroft, J. Ullman, (1983), Data Structures and Algorithms, Addison-Wesley, Reading, MA.

Akl, S., (1981), On the Expected Number of Optimal and Near-Optimal Solutions to the Euclidean Traveling Salesman Problem, J. Comput. Appl. Math., 7, 287-288.

Badrikian, A., J. F. C. Kingman, J. Kuelbs, (1976), Ecole d'Ete de Probabilites de Saint-Flour V-1975, Lecture Notes in Mathematics, Springer-Verlag. Berlin. Heidelberg. New York.

Balas, E., N. Christofides, (1981), A Restricted Lagrangian Approach to the Traveling Salesman Problem, Math. Progr., 21, 19--46.

Ball, M., M. Magazine, (1981), The Design and Analysis of Heuristics, Networks, 11, 215--219.

Beardwood, J., J. Halton, J. Hammersley, (1959), The Shortest Path Through Many Points, Proc. Camb. Phil. Soc., 55, 299--327.

Bellman, R., (1957), Dynamic Programming, Princeton Univ. Press.

Bellmore, M., G. Nemhauser, (1968), The Traveling Salesman Problem: A Survey, Oper. Res., 16, 538--558.

Bentley, J., J. Saxe, (1980), An analysis of two heuristics for the Euclidean Traveling Salesman Problem, Proc. 18th Annual Conf. Comm. Contr. Computing, 41--49.

Bodin, L., B. Golden, (1981), Classification in Vehicle Routing and Scheduling, Networks, 11, 97--108.

Bodin, L., B. Golden, A. Assad, M. Ball, (1983), Routing and Scheduling of Vehicles and Crews, The State of the Art, Comput. and Oper. Res., 10, 69--211.

Bollobas, B., (1973), The Optimal Arrangement of Producers, J. London Math. Soc., 6, 605--613.

Burness, R., J. White, (1976), The Traveling Salesman Location Problem, Transp. Science, 10, 348--360.

Christofides, N., (1975) Graph Theory: An Algorithmic Approach, Academic Press.

Christofides, N., (1976), Worst-Case Analysis of a New Heuristic for the Traveling Salesman Problem, Technical Report, Carnegie-Mellon University, Pittsburgh, PA.

Christofides, N., S. Eilon, (1969), Expected Distances in Distribution Problems, Op. Res. Quart., 20, 437--443.

Clarke, G., J. Wright, (1964), Scheduling of Vehicles from a Central Depot to a Number of Delivery Points, Oper. Res., 12, 568--581.

Crowder, H., M. Padberg, (1980), Solving Large-Scale Symmetric Traveling Salesman Problems to Optimality, Man. Science, 26, 495--509.

Dempster, M., J. Lenstra, A. Rinnooy Kan, eds., (1982), Deterministic and Stochastic Scheduling, NATO Adv. Study Series C: Math and Phys. Sciences.

Efron, B., C. Stein, (1981), The Jackknife Estimate of Variance, Ann. Statist., 9, 586--596.

Eilon, S., C. Watson-Gandy, N. Christofides, (1971), Distribution Management, Hafner.

Erdos, P., J. Spencer, (1974), Probabilistic Methods in Combinatorics, Academic Press.

Fejes Toth, L., (1959), Sur La Representation D'Une Population Infinie Par Un Nombre Fini D'Elements, Acta Math. Acad. Sci. Hung., 9, 299--304.

Fejes Toth, L., (1960), On Shortest Nets With Equal Area, Acta Math. Acad. Sci. Hung., 11, 363--370.

Few, L., (1955), The Shortest Path and the Shortest Road Through N Points, Mathematika, 2, 141--144.

Fisher, M., (1980), Worst-Case Analysis of Heuristic Algorithms, Man. Science, 26, 1--17.

Fisher, M., D. Hochbaum, (1980), Probabilistic Analysis of the Planar K-Median Problem, Math. Oper. Res., 5, 27--34.

Fisher, M., R. Jaikumar, (1981), A Generalized Assignment Heuristic for Vehicle Routing, Networks, 11, 109--124.

Fox, B., (1978), Data Structures and Computer Science Techniques in Operations Research, Oper. Res., 26, 686--717.

Frank, H., (1969), Shortest Paths in Probabilistic Graphs, Oper. Res., 17, 583--599.

Garey, M., D. Johnson, (1979), Computers and Intractability: A Guide to the Theory of NP-Completeness, Will Freeman.

Garfinkel, R., K. Gilbert, (1978), The Bottleneck Traveling Salesman Problem: Algorithms and Probabilistic Analysis, J. ACM, 25, 435--448.

Gilbert, E., (1965), Random Minimal Trees, J. Soc. Ind. Appl. Math., 13, 376--387.

Golden, B., L. Bodin, T. Doyle, W. Stewart, (1980), Approximate Traveling Salesman Algorithms, Oper. Res., 28, 694--711.

Golden, B., T. Magnanti, (1980), Network Optimization, Manuscript.

Haimovich, M., (1984), Large Scale Geometric Location Problems, Ph.D. Dissertation, M.I.T.

Haimovich, M., A. Rinnooy Kan, (1983), Bounds and Heuristics for Capacitated Routing Problems, Manuscript.

Halton, J., R. Terada, (1982), A Fast Algorithm for the Euclidean Traveling Salesman Problem, Optimal with Probability One, SIAM J. Comp., 11, 28--46.

Hammersley, J.M., (1972), A Few Seedlings of Research, Proc. 6th Berkeley Symp. Math. Statist. Prob., 1.

Hansotia, B., (1979), Stochastic Linear Programming with Recourse: A Tutorial, Deci. Sci., 11, 151-168.

Held, M., R. Karp, (1962), A Dynamic Programming Approach to Sequencing Problems, J. SIAM, 10, 1, 196--210.

Held, M., R. Karp, (1970), The Traveling Salesman Problem and Minimum Spanning Trees, Oper. Res., 18, 1138--1162.

Held, M., R. Karp, (1971), The Traveling Salesman Problem and Minimum Spanning Trees: Part II, Math. Prog., 1, 6--25.

Hochbaum, D., J. Steele, (1982), Steinhaus's Geometric Location Problem for Random Samples in the Plane, Adv. Appl. Prob., 14, 56--67.

Kao, E., (1978), A Preference Order Dynamic Program for a Stochastic Traveling Salesman Problem, Oper. Res., 26, 1033--1045.

Karlin, S., H. Taylor, (1975), A First Course in Stochastic Processes, 2nd edition, Academic Press.

Karp, R., (1976), The Probabilistic Analysis of Some Combinatorial Search Algorithms, in J.F. Traub (ed.), Algorithms and Complexity, 1--19.

Karp, R., (1977), Probabilistic Analysis of Partitioning Algorithms for the Traveling Salesman Problem in the Plane, Math. Oper. Res., 2, 209--224.

Karp, R., R. Tarjan, (1980), Linear Expected-Time Algorithms for Connectivity Problems, J. of Alg., 1, 374--393.

Kendall, M., P. Moran, (1962), Geometrical Probability, Charles Griffin and Company Limited.

Kirkpatrick, S., C. Gelatt, M. Vecchi, (1983), Optimization by Simulated Annealing, Science, 220, 671--680.

Klee, V., (1980), Combinatorial Optimization: What is the State of the Art, Math. Oper. Res., 5, 1--26.

Lardinois, C., (1980), Problemes D'Optimisation Combinatoire De Grande Taille Et Inference Statistique, Ph.D. Dissertation, Universite de Montreal.

Larson, R., A. Odoni, (1981), Urban Operation Research, Prentice Hall.

Lawler, E., (1963), The Quadratic Assignment Problem, Man. Science, 7, 586--599.

Lawler, E., (1976), Combinatorial Optimization: Networks and Matroids, Holt, Rinehart, and Winston.

Leipala, T., (1978), On the Solutions of Stochastic Traveling Salesman Problems, Eur. J. Oper. Res., 2, 291--297.

Lenstra, J., A. Rinnooy Kan, (1975), Some Simple Applications of the Traveling Salesman Problem, Op. Res. Quart., 24, 717--733.

Lenstra, J., A. Rinnooy Kan, (1981), Complexity of Vehicle Routing and Scheduling Problems, Networks, 11, 221--227.

Lin, S., B. Kernighan, (1973), An Effective Heuristic Algorithm for the Traveling Salesman Problem, Oper. Res., 21, 498--516.

Lukacs, E., (1968), Stochastic Convergence, Heath Mathematical Monographs.

Magnanti, T., (1981), Combinatorial Optimization and Vehicle Fleet Planning: Perspectives and Prospects, Networks, 11, 179--213.

Marchetti Spaccamela, A., A. Rinnooy Kan, L. Stougie, (1984), Hierarchical Vehicle Routing Problems, Networks, 14, 571--586.

Milton, J., C. Tsokos, (1976), Probability Theory with the Essential Analysis, Addison-Wesley, Reading, MA.

Neveu, J., (1965), Mathematical Foundations of the Calculus of Probability, Holden-Day, Inc.

Panaysotopoulos, J., (1982), Probabilistic Analysis of Solving the Assignment Problem for the Traveling Salesman Problem, Eur. J. Oper. Res., 9, 76--83.

Papadimitriou, C. (1978), The Probabilistic Analysis of Matching Heuristics, Proc. 15th Annual Conf. Comm. Contr. Computing, 368--378.

Papadimitriou, C., (1981), Worst Case and Probabilistic Analysis of a Geometric Location Problem, SIAM J. Comp., 10, 542--557.

Papadimitriou, C.; K. Steiglitz, (1982), Combinatorial Optimization, Prentice Hall.

Parker, R., R. Rardin, (1983), The Traveling Salesman Problem: An Update of Research, Naval Res. Log. Quart., 30, 69--96.

Platzman, L., J. Bartholdi, (1983), Spacefilling Curves and Routing Problems in the Plane, submitted to Math. of Oper. Res.

Psaraftis, H., (1984), On the Practical Importance of Asymptotic Optimality in Certain Heuristic Algorithms, Networks, 14, 587--596.

Renyi, A., R. Sulanke, (1963), Uber die Konvexe Hulle von n zufallig gewahlten Punkten, Z. Wahrsch. Verw. Gebiete, 2, 75--85.

Rosenkrantz, D., R. Stearns, P. Lewis, (1977), An Analysis of Several Heuristics for the Traveling Salesman Problem, SIAM J. Comp. 6, 563--581.

Sahni, S., E. Horowitz, (1978), Combinatorial Problems: Reducibility and Approximation, Oper. Res., 26, 718--759.

Shamos, M., (1975), Geometric Complexity, Proc. 7th ACM Symp. on Theory of Computing, 224--233.

Shamos, M., D. Hoey, (1975), Closest-Point Problems, Proc. 16th Annual IEEE Symp. on Foundations of Computer Science, 151--162.

Solomon, H., (1978), Geometric Probability, Society for Industrial and Applied Mathematics.

Steele, J., (1981a), Subadditive Euclidean Functionals and Nonlinear Growth in Geometric Probability, Ann. Prob., 9, 365--376.

Steele, J., (1981b), Complete Convergence of Short Paths and Karp's Algorithm for the TSP, Math. Oper. Res., 6, 374-378.

Steele, J., (1982), Optimal Triangulation of Random Samples in the Plane, Ann. Prob., 10, 548--553.

Stein, D., (1977), Scheduling Dial-A-Ride Transportation Systems, Ph.D. Dissertation, Harvard University.

Stein, D., (1978), An Asymptotic Probabilistic Analysis of a Routing Problem, Math. Oper. Res., 3, 89--101.

Stewart, W., B. Golden, (1983), Stochastic Vehicle Routing: A Comprehensive Approach, Eur. J. Oper. Res., 14, 371--385.

Tarjan, R., (1983), Data Structures and Network Algorithms, Society for Industrial and Applied Mathematics.

Tillman, F., (1969), The Multiple Terminal Delivery Problem with Probabilistic Demands, Transp. Science, 3, 192-204.

Tomescu, I., (1983), An Upper Bound for the Shortest Hamiltonian Path in the Symmetric Euclidean Case, RAIRO Rech. Oper./Oper. Res., 17, 297--306.

Vajda, S., (1972), Probabilistic Programming, Academic Press.

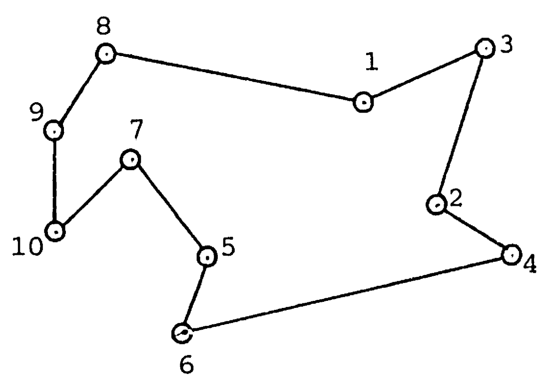
Verblunsky, S., (1951), On the Shortest Path Through a Number of Points, Proc. Amer. Math. Soc., 2, 904--913.

von Neuman, J., (1950), Functional Operators, Vol. 1. Annals of Math. Studies, 21 (Princeton).

# APPENDIX A

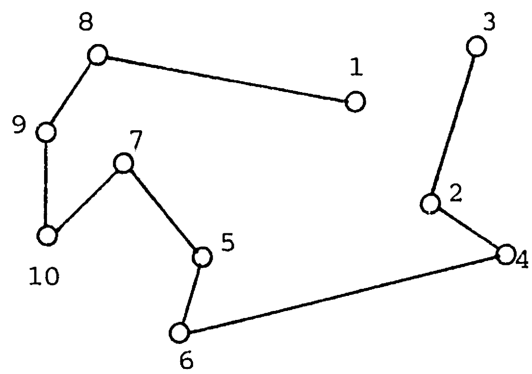
## Graphical Illustrations of the $L_{m,t}^{(r)}$

The purpose of this appendix is to provide a graphical aid for understanding how the $L_{m,t}^{(1)}$ $\left(L_{m,h}^{(r)}\right)$ are obtained from a given tour t (path h) of a graph G. Our examples cover the cases m=0, m=1, and m=2 for tours and m=0, m=1 for paths. To facilitate comparisons, the illustrations start, in each case, with the same initial tour (path) through a set of 10 points.




$L_{m,t}^{(o)}$ m = 0,1, or 2

(The initial tour)



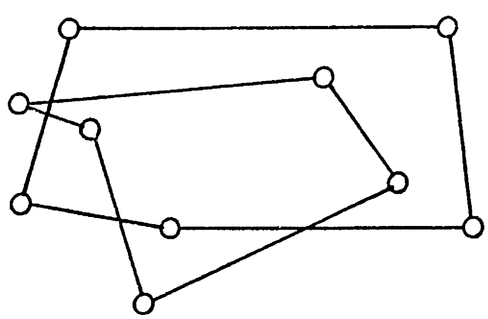

$L^{(o)}_{m,h}$ $\qquad$ m = 0, or 1

(The initial path)

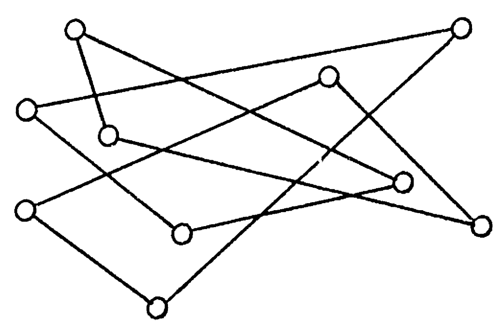
The different cases considered correspond to various assumptions for the color of the nodes.

A.1 Illustrations for tours:

(i) Every node is white (i.e., m=0)

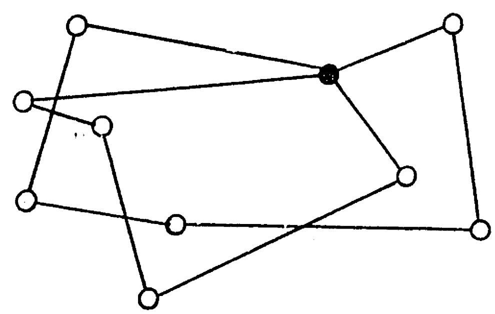


$L^{(1)}_{0,t}$

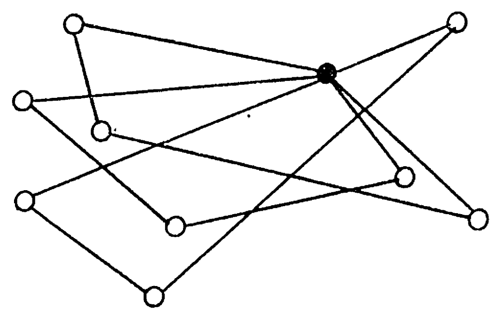


$L^{(2)}_{0,t}$

(ii) <u>Node 1 is black, the other nodes are white (i.e., m=1)</u>

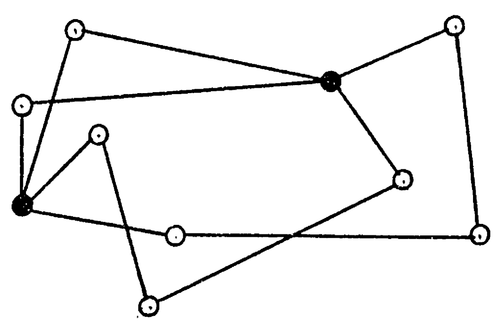


$L_{1,t}^{(1)}$

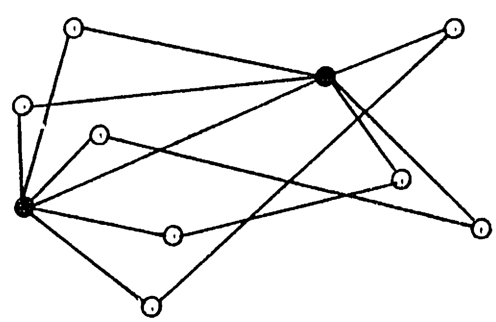


$L_{1,t}^{(2)}$

(iii) <u>Nodes 1 and 10 are black, the other nodes are white (i.e., m=2)</u>



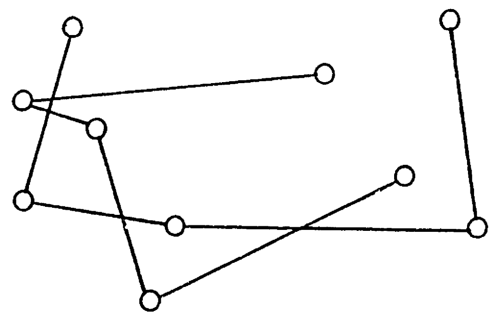
$L_{2,t}^{(1)}$



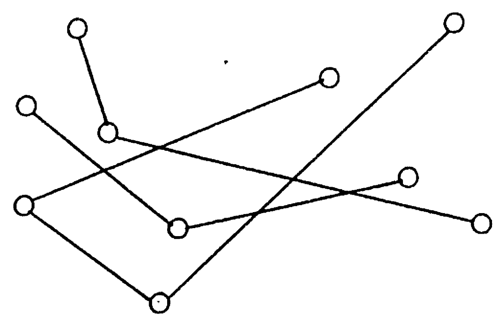
$L_{2,t}^{(2)}$

A.2 Illustrations for paths:

(i) All nodes are white (i.e., m=0)

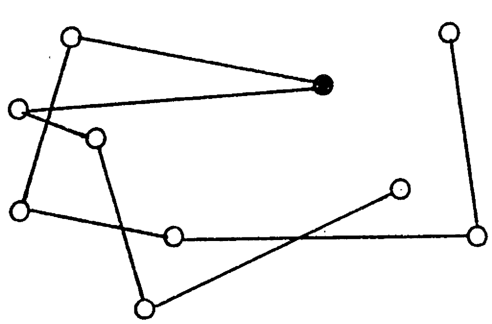


$L^{(1)}_{0,h}$

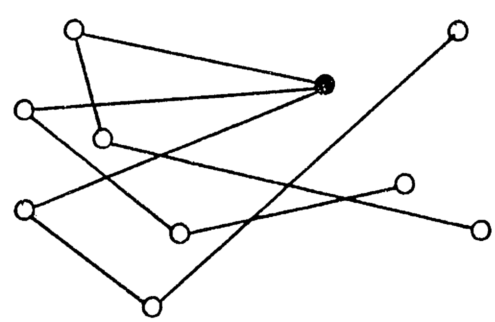


$L^{(2)}_{0,h}$

(ii) 1 is black, the other nodes are white (i.e., m=1)



$L^{(1)}_{1,h}$



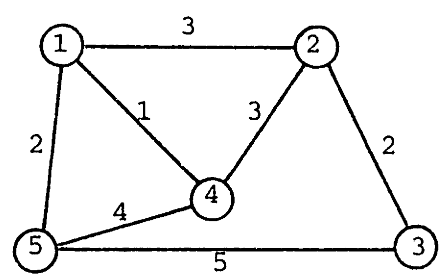
$L^{(2)}_{1,h}$

# APPENDIX B

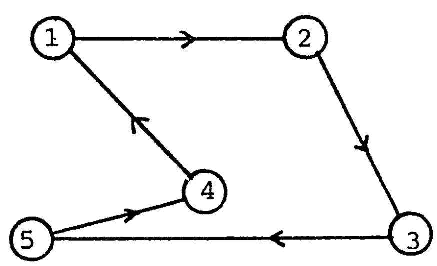
## Case of a Non-Complete Graph G

This appendix is supplementary to the discussion of section 3.1 (3.1.2A.).

Consider the following non-complete, (but connected) symmetric, weighted graph G:

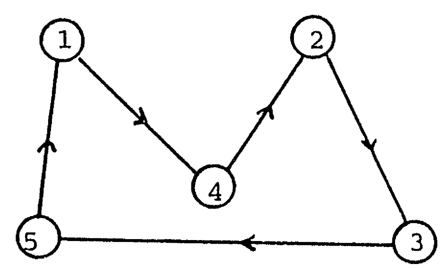



Consider the following two tours of G




tour 1

(4,1,2,3,5,4)




tour 2

(4,2,3,5,1,4)

To be able to compute their expected length by using results of Chapter 2 we need to define the strategy to follow when we wish to go from a node i to a node j for which arc (i,j) does not exist (in other words, we need to construct a complete graph).

Strategy (1):

Go from i to j using intermediary nodes that exist along the tour t (choosing the shortest path along the tour if several possibilities are offered).

For our example: tour 1 => d(4,3) = 3+2=5

d(3,4) = 5+4=9

tour 2 => d(4,3) = 3+2=5

d(3,4) = 5+2+1=8

It is then apparent that this strategy is not very convenient since it is tour-dependent; moreover it usually leads to an asymmetric complete graph, even if the original non-complete graph is symmetric.

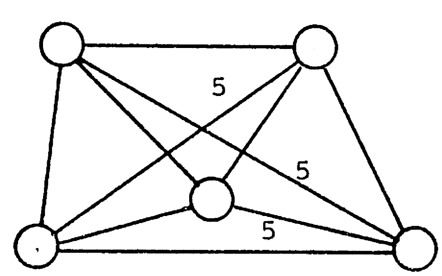
Strategy (2):

For all non-existing arcs (i,j), compute the shortest path to go from i to j using existing arcs. (Note: the presence or not of the nodes has no influence on the a-priori construction of the complete graph; if one wished to use only present nodes as intermediary nodes for finding the shortest path, the constructed complete graph would be instance-dependent, a "non-desirable" property).

For our example:

d(4,3) = d(3,4) = 3+2=5 (through node 2)

Following strategy 2 we can obtain a complete graph; for our example, the final complete graph is as follows:



# APPENDIX C

## The Optimal Euclidean PTSP Tour May Intersect Itself: An Example

In this appendix we present the full set of calculations associated with the counterexample (see Figures 3.1 and 3.2) provided in the proof of Lemma 3.7.

### C.1 Derivations Corresponding to Figures 3.1 and 3.2

The euclidean distances between points are as follows:

$d(1,2) = d(1,3) = \sqrt{260}$

$d(1,4) = \sqrt{160}$; $d(1,5) = \sqrt{388}$

$d(2,3) = 4$; $d(2,4) = \sqrt{52}$; $d(2,5) = \sqrt{464}$

$d(3,4) = \sqrt{20}$; $d(3,5) = \sqrt{320}$

$d(4,5) = \sqrt{212}$

In Figure 3.2 we presented two tours through the set of five points; in fact we have $(5-1)!/2 = 12$ possible tours. The 10 remaining tours are displayed at the end of this Appendix.

Using Theorem 2.3 with m=2 one can compute the expected length of the 12 tours as a in function of p. Let us give $L_j^{(r)}$'s for the 12 tours $j \in [1..12]$. Note that from Facts 3.2, 3.4, $L_j^{(2)}$ and $L_j^{(3)}$ are tour-independent (i.e. independent of j) so that we do not need them to compare the tours j.

- $L_1^{(o)} \simeq 58.854587$ $\qquad L_1^{(1)} \simeq 98.166568$
- $L_2^{(o)} \simeq 61.593553$ $\qquad L_2^{(1)} \simeq 94.514453$
- $L_3^{(o)} \simeq 65.22239$ $\qquad L_3^{(1)} \simeq 104.53437$
- $L_4^{(o)} \simeq 61.446473$ $\qquad L_4^{(1)} \simeq 98.019488$
- $L_5^{(o)} \simeq 68.874505$ $\qquad L_5^{(1)} \simeq 104.53437$
- $L_6^{(o)} \simeq 72.822046$ $\qquad L_6^{(1)} \simeq 91.480061$ $\qquad$ (C.1)
- $L_7^{(o)} \simeq 65.394013$ $\qquad L_7^{(1)} \simeq 98.019488$
- $L_8^{(o)} \simeq 62.359621$ $\qquad L_8^{(1)} \simeq 98.019488$
- $L_9^{(o)} \simeq 69.046129$ $\qquad L_9^{(1)} \simeq 98.019488$
- $L_{10}^{(o)} \simeq 72.674965$ $\qquad L_{10}^{(1)} \simeq 98.019488$
- $L_{11}^{(o)} \simeq 71.908897$ $\qquad L_{11}^{(1)} \simeq 90.566912$
- $L_{12}^{(o)} \simeq 75.413932$ $\qquad L_{12}^{(1)} \simeq 98.019488$

We have:

$$E[L_j] = p^2\left(L_j^{(o)} + (1-p)\ L_j^{(1)}\right) + K \qquad \text{for } j \in [1..12]$$

$$\text{where } K = p(1-p)^2\ L_j^{(2)} + (1-p)^3\ L_j^{(3)}$$

Elimation rule:

If $L_j^{(o)} > L_i^{(o)}$ and $L_j^{(1)} > L_i^{(1)}$ then tour i has a shorter expected length than tour j for any value of p which, in turn, implies that tour j can be discarded.

Following this rule, one can see that only tours 1,2,4 and 11 have to be considered as potential optimal PTSP tours (1 eliminates 3,5; 2 eliminates 7,8,9,10,12; 11 elimates 6).

From (C.1) we have the proof that for p=1 tour 1 is the optimal TSP tour ($L_1^{(o)}$ is the smallest among $L_j^{(o)}$).

For $0 < p < 1$ we have:

$$E[L_1] - E[L_2] = p^2[-2.738966 + 3.652115(1-p)] \tag{C.2}$$

$$E[L_1] - E[L_4] = p^2[-2.591886 + 0.14708(1-p)] \tag{C.3}$$

$$E[L_1] - E[L_{11}] = p^2[-13.05431 + 7.599656(1-p)] \tag{C.4}$$

Now, since $0 < (1-p) < 1$, (C.3) and (C.4) are always negative (i.e. for any p), which implies that 1 is always "better" than 4 and 11.

From C.2 one can see that:

if $p > 1 - \frac{2.738966}{3.652115} \simeq 0.25$, tour 1 is better than tour 2,

if $p < 0.25$, tour 2 is better than tour 1.

Q.E.D.

C.2 Case with one or no black nodes:

Following the discussion provided in the text, one can replace each black nodes of Figure 3.1 by a large number of superimposed white nodes so that tour 2 can still be the optimal PTSP tour for some values of p.

One can verify that tours 1 and 2, with nodes 4 and 5 both replaced by 12 white nodes, are such that:

$$L_1^{(o)} - L_2^{(o)} = -2.738966$$

$$L_1^{(1)} - L_2^{(1)} = 3.652115$$

$$L_1^{(r)} - L_2^{(r)} = 0 \quad \text{for } r \in [2..11]$$

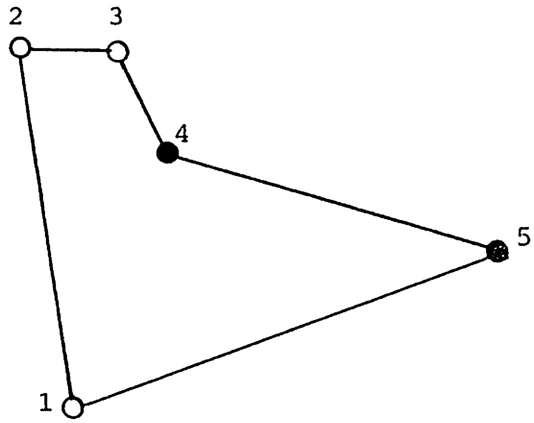
$$L_1^{(12)} - L_2^{(12)} = -0.913149$$

Hence, using Theorem 2.1 and Fact 3.4, we have:

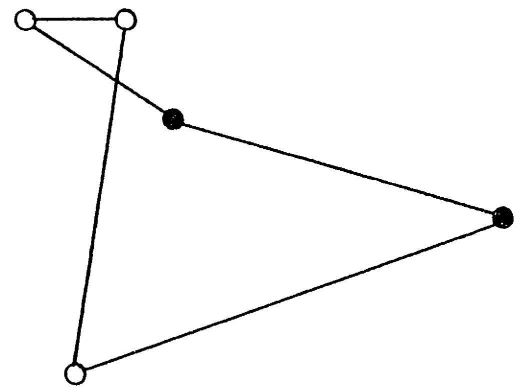
$$E[L_1]-E[L_2] = p^2\Big[\big(1+(1-p)^{25}\big)(-2.738966) + \big((1-p)+(1-p)^{24}\big)(3.652115) + \big((1-p)^{12}+(1-p)^{13}\big)(-0.913149)\Big]$$

For $p = 0.1$ $E[L_1]-E[L_2] = 0.1526134$

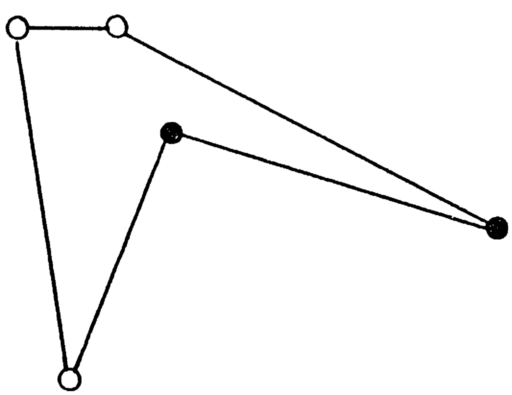
Q.E.D.



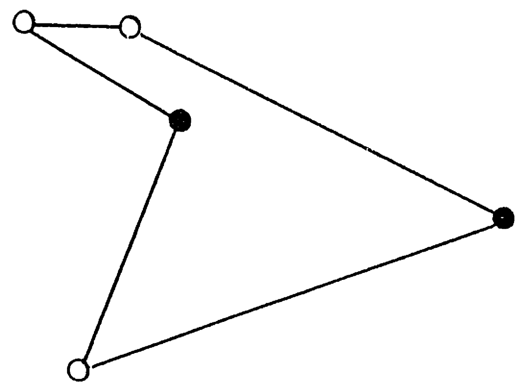

tour 1



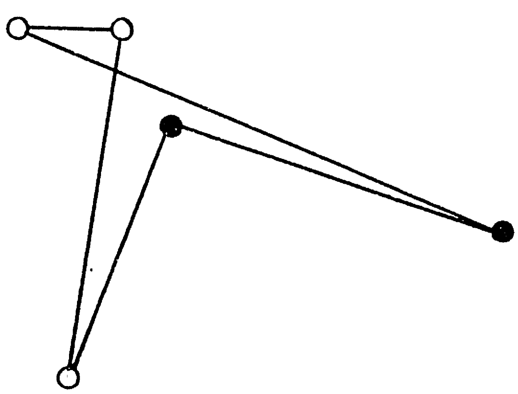
tour 2



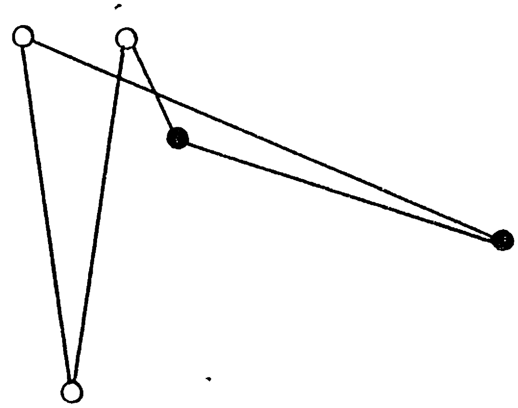
tour 3



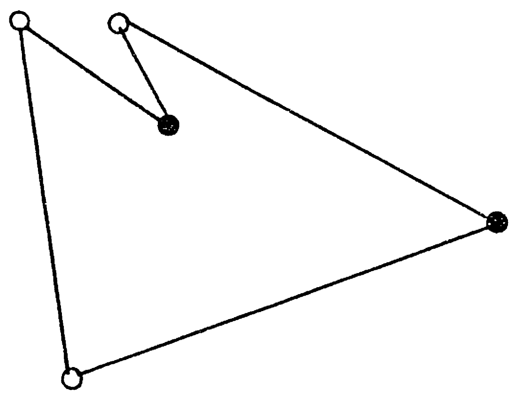
tour 4



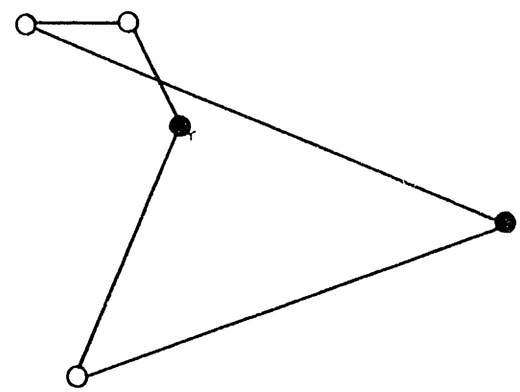
tour 5



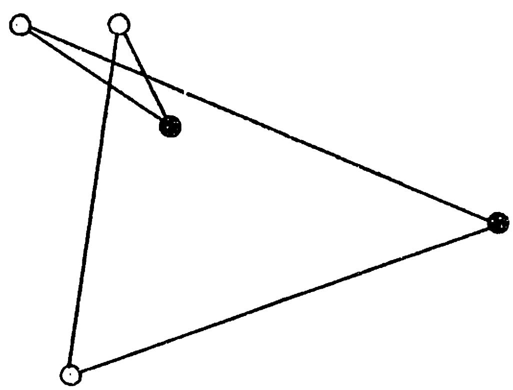
tour 6



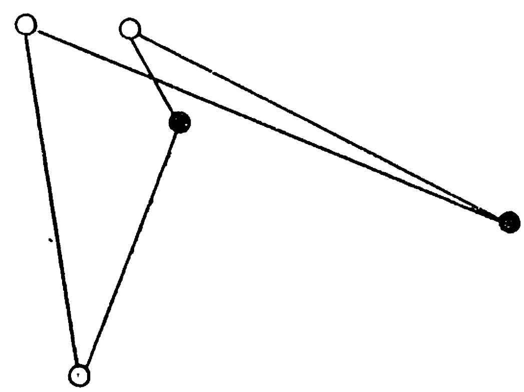
tour 7



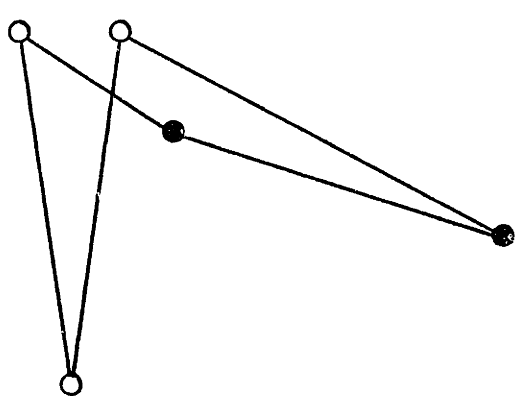
tour 8



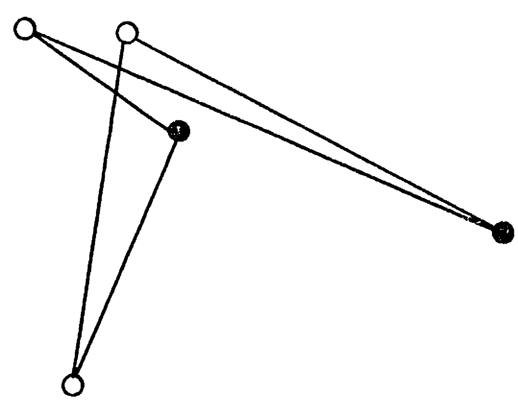
tour 9



tour 10



tour 11



tour 12

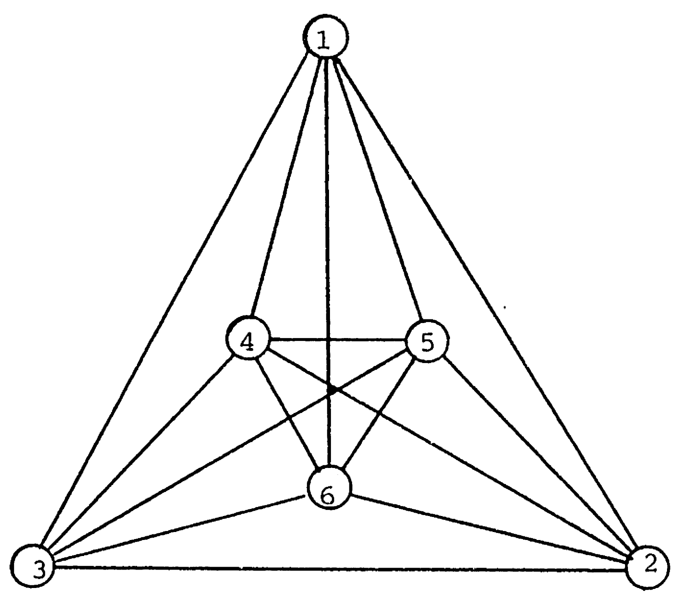
# APPENDIX D

## Exact Calculations for Lemma 3.9

This appendix provides the exact calculations corresponding to the two counterexamples involved in the proof of Lemma 3.9(ii):

D.1 m=0, n=6, coverage probability p:

Consider the following complete graph (it corresponds to Figure 3.4):




The distance matrix D is as follows; (symmetric D).

d(1,3) = d(1,2) = d(2,3) = 4

d(4,5) = d(4,6) = d(5,6) = 0.4

d(1,6) = d(3,5) = d(2,4) = 2.54

d(1,4) = d(1,5) = d(2,5) = d(2,6) = d(3,4) = d(3,6) = 2.203

The tour 3 (see Figure 3.4) is then such that:

$$L_3^{(o)} = 13.206$$

$$L_3^{(1)} = 13.886$$

$$L_3^{(2)} = 13.892$$

The tour 4 (see Figure 3.4) is such that:

$$L_4^{(o)} = 13.206$$

$$L_3^{(1)} = 13.212$$

$$L_4^{(2)} = 15.24$$

Thus, using the expression developed in Chapter 2, we obtain:

$$E[L_3]-E[L_4] = p^2\Big[\big((1-p)+(1-p)^3\big)0.674-(1-p)^2 1.348\Big]$$

$$= p^4(1-p)[0.674]$$

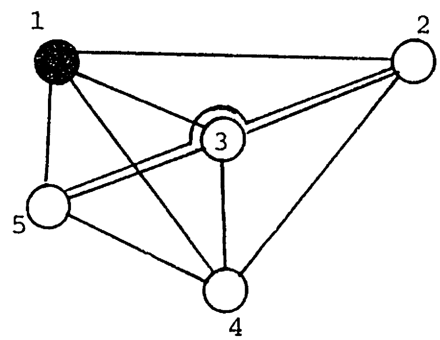
which shows that:

- if p=1, 3 and 4 are of equal length (in fact they are optimal TSP tours, since it it easy to verify that all possible tours have a length at least as large as 13.206).
- if $p < 1$, the optimal TSP tour 3 is not the optimal PTSP tour since then tour 4 has a smaller expected length.

Q.E.D.

D.2 <u>m=1, n=4, coverage probability p</u>:

Consider the following complete graph:




The distance matrix D (euclidean) is as follows:

$$d(1,2) = 4$$

$$d(1,3) = d(2,3) = d(3,5) = d(4,5) = \sqrt{5}$$

$$d(1,4) = d(2,4) = \sqrt{13}$$

$$d(1,5) = d(3,4) = 2$$

$$d(2,5) = 2\sqrt{5}$$

Tours 1 and 2 are then such that:

$$L_1^{(o)} - L_2^{(o)} = \sqrt{13} + \sqrt{5} - 6 < 0$$

$$L_1^{(1)} - L_2^{(1)} = \sqrt{5} + 2 - \sqrt{13}$$

$$L_1^{(2)} - L_2^{(2)} = 4 - 2\sqrt{5}$$

$$L_1^{(3)} - L_2^{(3)} = 0$$

From the expressions of Chapter 2 we conclude, after some calculus, that:

$$E[L_1] - E[L_2] = p^3\left[\left(3\sqrt{5} + \sqrt{13} - 10\right) - p\left(2\sqrt{5} - 4\right)\right]$$

Hence:

- for $p > \dfrac{3\sqrt{5} + \sqrt{13} - 10}{2\sqrt{5} - 4} \approx 0.66$

$$E[L_1] - E[L_2] < 0$$

implying 1 is "better" than 2 (in fact, 1 is the TSP tour)

- if $p < 0.66$,

  2 is "better" than 1

Q.E.D.

# APPENDIX E

## Calculus Associated with the Asymptotic Behavior of the Star-Shaped Construction

In this appendix, we provide the detailed calculations associated with the proof of Lemma 3.11 (refer to the text for definition).

### E.1 Choice of $d_2$ so that $L_a^{(o)} = L_b^{(o)}$ for all n

[Comment: $L_a^{(o)} = L_b^{(o)}$ means that both a and b are optimal TSP tours; however Lemma 3.11 shows that tour b is much better than tour a with respect to the PTSP (end of comment)]

In the text we showed that:

$$L_a^{(o)} = (n-1)\left(\frac{1}{n} + d_2\right) + 2\alpha \qquad (E.1)$$

$$L_b^{(o)} = 2n\alpha \qquad (E.2)$$

$$\alpha = \frac{1}{2\sin\pi/n}\left[\sqrt{(1/n)^2+d_2^2-(2/n)d_2\cos\pi/n}\right] \qquad (E.3)$$

By replacing (E.3) in (E.1) and (E.2), $L_a^{(o)} = L_b^{(o)}$ can be expressed as:

$$\left(\frac{1}{n} + d_2\right)\sin\pi/n = \sqrt{(1/n)^2+d_2^2-(2/n)d_2\cos\pi/n} \qquad (E.4)$$

By taking the square of both sides of (E.4) we can then solve for $d_2$ and we obtain:

$$d_2 = \frac{1}{n}\left[Kn - \sqrt{Kn^2 - 1}\right] \tag{E.5}$$

where $\quad Kn = \dfrac{1+\cos\pi/n-\cos^2\pi/n}{\cos^2\pi/n}$

E.2 Limiting behavior of $L_b^{(r)}$, $L_a^{(r)}$

(i) $L_b^{(r)}$:

We have (from the construction of an n-gon)

$$L_b^{(r)} = 2n\alpha^{(r)} \tag{E.6}$$

where $\quad \alpha^{(r)} = \dfrac{1}{2\sin\pi/n}\left[\sqrt{(1/n)^2 + d_2^2 - (2/n)d_2\cos(2r-1)\pi/n}\right]$ (E.7)

let $a_n \equiv \cos\dfrac{\pi}{n} \qquad b(r,n) \equiv \cos\big((2r-1)\pi/n\big)$ (E.8)

Then after some manipulations (E.5), (E.6), (E.7), and (E.8) imply that:

$$L_b^{(r)} = \sqrt{A+B} \tag{E.9}$$

where

$$A = \frac{a_n^4 + (1+a_n-a_n^2)^2 + (1-a_n^2)(1+2a_n) - 2b(r,n)a_n^2(1+a_n-a_n^2)}{a_n^4(1-a_n)(1+a_n)}$$

(E.10)

$$B = \frac{2}{a_n^4}\sqrt{\frac{1+2a_n}{(1-a_n)(1+a_n)}}\left[b(r,n)a_n^2 - (1+a_n-a_n^2)\right]$$

When $n \to \infty$: $b(r,n) \to 1$ (for any finite r)

$a_n \to 1$

- In fact when $n \to \infty$

$$b(r,n) \sim 1 - \frac{(2r-1)^2\pi^2}{2n^2}$$

$$a_n \sim 1 - \frac{\pi^2}{2n^2}$$

which implies:

$$b(r,n)\, a_n^2 - (1+a_n-a_n^2) = 0\left(\frac{1}{n^2}\right)$$

$$\sqrt{1-a_n} = 0\left(\frac{1}{n}\right)$$

Hence $\lim_{n\to\infty} B = 0$ (E.10)

- Also when $n \to \infty$

$$b(r,n)\, a_n^2(1+a_n-a_n^2) \sim 1 - \left((2r-1)^2+1\right)\frac{\pi^2}{2n^2}$$

$$a_n^4 \sim 1 - \frac{2\pi^2}{n^2}$$

$$(1+a_n-a_n^2) \sim 1 + \frac{\pi^2}{n^2}$$

$$(1-a_n^2)(1+2a_n) \sim \frac{3\pi^2}{n^2}$$

which implies:

$$A \sim \frac{\frac{\pi^2}{n^2}\left((2r-1)^2+3\right)}{\frac{\pi^2}{2n^2}\left(2 - \frac{\pi^2}{2n^2}\right)\left(1 - \frac{2\pi^2}{n^2}\right)}$$

Hence $\lim_{n\to\infty} A = (2r-1)^2 + 3 = 4(r^2-r+1)$ (E.11)

Finally (E.9), (E.10), and (E.11) imply:

$$\lim_{n\to\infty} L_b^{(r)} = 2\sqrt{r^2-r+1} \qquad \text{for any finite } r$$

(as claimed in the text)

(ii) $L_a^{(r)}$

This case is much easier than the previous one: for finite r and $n \to \infty$ $L_a^{(r)}$ behaves as twice $L_{0,t_1}^{(r)}$ of Lemma 3.10

hence $\lim_{n\to\infty} L_a^{(r)} = 2(r+1)$ for any finite r

## E.3 The final step

We have $E[L_t] = \sum_{r=0}^{n-2} p^2(1-p)^r L_t^{(r)}$ for t = a or b

Since for any $0 < p < 1$ we have

$$\lim_{r\to\infty} 2(1-p)^r (r+1) = 0$$

and $\lim_{r\to\infty} 2(1-p)^r \sqrt{r^2-r+1} = 0$

then

$$\lim_{n\to\infty} E[L_a] = \sum_{r=0}^{\infty} 2p^2(1-p)^r(r+1) = 2 \qquad (E.12)$$

$$\lim_{n\to\infty} E[L_b] = \sum_{r=0}^{\infty} 2p^2(1-p)^r\sqrt{r^2-r+1} \qquad (E.13)$$

Since $\qquad r-1 \leq \sqrt{r^2-r+1} \leq r$

(E.13) becomes:

$$2\left(\frac{1}{2-p} - \frac{p(1-p)^2}{2-p}\right) \leq \lim_{n\to\infty} E[L_b] \leq \frac{2}{2-p} \qquad (E.14)$$

Q.E.D.

# APPENDIX F

## The Three Mathematical Programming Formulations of the PTSP

In this appendix we present the three mathematical programming formulations of the PTSP in the context of a graph G with n white nodes and no black node. This is followed by a numerical table comparing the number of variables and constraints in those formulations for some chosen values of n. The definition of the variables is contained in section 5.2.

### $F_1$: Integer Nonlinear Programming Formulation:

Minimize $$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=0}^{n-2} \alpha_r \, d(i,j) \, x_{ij}^{(r)}$$

subject to:

$$x_{ij}^{(o)} \equiv x_{ij} \qquad (i,j \in [1..n])$$

$$x_{ij}^{(r)} \equiv \sum_{k_1,\ldots,k_r}^{n} x_{ik_1} x_{k_1k_2} \cdots x_{k_rj} \qquad (r \in [1..n-2]) \; (i,j \in [1..n])$$

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad (j \in [1..n])$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad (i \in [1..n])$$

$$X = (x_{ij}) \in S$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i,j \in [1..n])$$

$F_2$: Mixed Integer Linear Programming Formulation:

$$\text{Minimize} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=0}^{n-2} \alpha_r \, d(i,j) \, x_{ij}^{(r)}$$

subject to:

$$x_{ij}^{(o)} \equiv x_{ij}$$

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad (j \in [1..n])$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad (i \in [1..n])$$

$$x_{ik}^{(r-1)} + x_{kj} - 1 \leq x_{ij}^{(r)} \leq 1 + x_{ik}^{(r-1)} - x_{kj} \qquad 1 \leq k \leq n \quad \begin{pmatrix} r \in [1..n-2] \\ i,j \in [1..n] \end{pmatrix}$$

$$X = (x_{ij}) \in S$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i,j \in [1..n])$$

$$x_{ij}^{(r)} \text{ continuous} \qquad (r \in [1..n-2],\ i,j \in [1..n])$$

$F_3$: Pure Integer Linear Programming Formulation:

$$\text{Minimize} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=0}^{n-2} \alpha_r \, d(i,j) \, x_{ij}^{(r)}$$

subject to:

$$x_{ij}^{(o)} \equiv x_{ij} \qquad (i,j \in [1..n])$$

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad (j \varepsilon [1..n])$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad (i \varepsilon [1..n])$$

$$x_{ij}^{(r)} \leqslant 1 - z_{ij}^{(r-1)}$$

$$x_{ij}^{(r)} \leqslant 1 + z_{ij}^{(r-1)}$$

$$x_{ij}^{(r)} \geqslant \alpha_{ij}^{(r)} + \beta_{ij}^{(r)} - 1 \qquad \begin{pmatrix} r \ \varepsilon \ [1..n-2] \\ i,j \ \varepsilon \ [1..n] \end{pmatrix}$$

$$\alpha_{ij}^{(r)} \geqslant \frac{1}{n} \left[ 1 + (n-1) \ z_{ij}^{(r-1)} \right]$$

$$\beta_{ij}^{(r)} \geqslant \frac{1}{n} \left[ 1 - (n-1) \ z_{ij}^{(r-1)} \right]$$

$$z_{ij}^{(r-1)} \equiv \left( \sum_{k=1}^{n} k \left( x_{ik}^{(r-1)} - x_{kj} \right) \right) / (n-1)$$

$$X = (x_{ij}) \ \varepsilon \ S$$

$$x_{ij} = 0 \text{ or } 1 \qquad (i,j \ \varepsilon \ [1..n])$$

$$x_{ij}^{(r)}, \ \alpha_{ij}^{(r)}, \ \beta_{ij}^{(r)} = 0 \text{ or } 1 \qquad \begin{pmatrix} r \ \varepsilon \ [1..n-2] \\ i,j \ \varepsilon \ [1..n] \end{pmatrix}$$

$F_4$: Numerical Table:(*)

| | $F_1$ | | $F_2$ | | $F_3$ | |
|---|---|---|---|---|---|---|
| n | #0-1 variables | # constraints(**) | # variables / #0-1 variables | # constraints(**) | #0-1 variables | # constraints(**) |
| 5 | 25 | 10 | 100/25 | 760 | 250 | 385 |
| 10 | 100 | 20 | 900/100 | 18,020 | 2,500 | 4,020 |
| 15 | 225 | 30 | 3,150/225 | 87,780 | 9,000 | 18,655 |
| 20 | 400 | 40 | 7,600/400 | 288,040 | 22,000 | 36,040 |
| 30 | 900 | 60 | 26,100/900 | 1,512,060 | 76,500 | 126,060 |
| 50 | 2,500 | 100 | 122,500/2,500 | 12,000,100 | 362,500 | 600,100 |
| 100 | 10,000 | 200 | X | X | X | X |
| 200 | 40,000 | 400 | X | X | X | X |
| 250 | 62,500 | 500 | X | X | X | X |
| 300 | 90,000 | 600 | X | X | X | X |

(*)this is given without using Fact 5.1 which would reduce the numbers for $F_2$ and $F_3$ approximately by half.

(**)without including the number of constraints associated with the set S of subtour eliminations (depending on the way we express the set S this number of contraints will either be $2^n$ or $n^2-3n+2$; see Golden and Magnanti [1980]).

# APPENDIX G

## A General Branch-And-Bound Scheme for the PTSP

In this appendix we provide a general branch-and-bound scheme for the PTSP based on subsection 5.2.3.B. We express this scheme in terms of "pidgin algol", an informal language, as introduced in Papadimitriou and Steiglitz [1982]. For the reader familiar with Pascal, Algol, or PL/I, this language should be easy to understand.

This scheme is adapted from the basic branch-and-bound algorithm given in Papadimitriou and Steiglitz [1982].

The additional notation is:

$\alpha_o$: weight of $L_{1,t}^{(o)}$ as in Chapter 3.

Ei: expected length in the PTSP sense of child i

h: lower bound on $\sum_{r=1}^{n-1} \alpha_r L_{1,t}^{(r)}$

```
begin
   activeset: = {0}; (comment: "0" is the original problem)
   U: = ∞
   currentbest: = anything;
   while activeset is not empty do
     begin
       choose a branching node, node k ε activeset;
       remove node k from activeset;
       generate the children of node k, child i, i=1,..,n_k,
       and the corresponding lower bounds, z_i;
       for i=1,..,n_k do
           begin
             if z_i > (U-h)/α_o then kill child i
               else if child i is a tour then
                     compute E_i
                     if E_i < U then
                           U: = E_i, currentbest: = child i
                           else add child i to activeset
                     else add child i to activeset
           end
     end
end
```

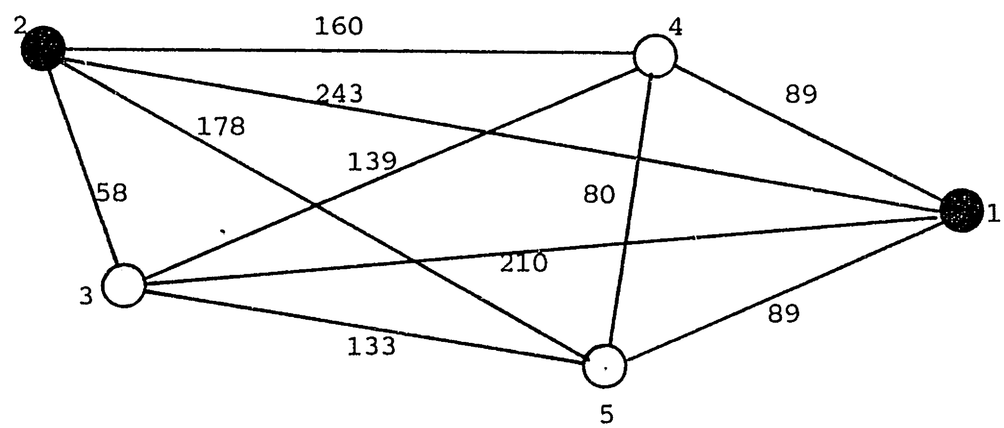
Note: to obtain $z_i$ one can use any method developed for the TSP (for example Held and Karp [1970], [1971])

# APPENDIX H

## The Inadequacy of a Dynamic Programming Approach for the PTSP: A Simple Example

This appendix contains a numerical and graphical example corresponding to the discussion contained in 5.2.3.C.

Consider the following complete, undirected, weighted graph:




Claim:

Assume the "coverage probability" p=0.5 for nodes 3,4,5 (node 1 is the depot). Then using the terminology developed in subsection 5.2.3.C and applying the recursive formula given in (5.30) we obtain:

$$E(\{3,4,5\},2) = 287.5$$

and this corresponds to the "optimal" path (2,3,5,4,1). However the expected length (in the PTSP sense) of the path (2,3,4,5,1) is 286 and this contradicts the validity of (5.30).

<u>Proof</u>: Let us first compute $E(\{3,4,5\},2)$ by using (5.30)

(1) $Q = \phi$:

$E(\phi,2) = 243$; $E(\phi,3) = 210$; $E(\phi,4) = 89$; $E(\phi,5) = 89$

(2) $Q = \{3\}$

$E(\{3\},2) = 255.5$; $E(\{3\},4) = 219$; $E(\{3\},5) = 216$; $E(\{3\},1) = 210$

(3) $Q = \{4\}$

$E(\{4\},2) = 246$; $E(\{4\},3) = 219$; $E(\{4\},5) = 129$; $E(\{4\},1) = 89$

(4) $Q = \{5\}$

$E(\{5\},2) = 255$; $E(\{5\},3) = 216$; $E(\{5\},4) = 129$; $E(\{5\},1) = 89$

(5) $Q = \{3,4\}$

$E(\{3,4\},5) = \min\{240.5,257.5\} = 240.5$ <-> "opt." path (5,3,4,1)

(6) $Q = \{3,5\}$

$E(\{3,5\},4) = \min\{242;257.5\} = 242$ <-> "opt." path (4,3,5,1)

(7) $Q = \{4,5\}$

$E(\{4,5\},3) = \min\{242,240.5\} = 240.5$ <-> "opt." path (3,5,4,1)

(8) $Q = \{3,4,5\}$

$E(\{3,4,5\},3) = \min\{287.5;333.25;340\} = 287.5$

and the corresponding "opt" path is $h^*=(2,3,5,4,1)$

To compute the expected length of the path $h = (2,3,4,5,1)$ one can use results from Chapter 2 (with two black nodes)

We have $L_{h_1}^{(o)} = 366$

$$L_{h_1}^{(1)} = 529$$

$$L_{h_1}^{(2)} = 784$$

$$L_{h_1}^{(3)} = 243$$

Hence $E[L_{h_1}] = (0.5)^2 \; [366+(0.5)529] + 0.5^3(784+243)$

$= 286$

Q.E.D.

Note:

It is not valid to use (5.30) to derive the optimal path but it is valid to use it to obtain the expected length of a path. For example one can check that $E[L_{h*}]$ is indeed 287.5 by using techniques from Chapter 2.