AUC Undergraduate Journal of Liberal Arts & Sciences Capstone Issue Vol. 5 2014

AUC amsterdam university college

V LUSTRUM 2014

AUC Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 5 2014

Special Lustrum Edition

www.auc.nl

University of Amsterdam

VU University Amsterdam

AUC Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 5 2014

AUC Undergraduate Journal of Liberal Arts & Sciences
Capstone Issue 2014

Capstone Coordinator
Dr. Mariette Willemsen

2014 Capstone Awards Committee
Prof. Louis Akkermans
Dr. Forrest Bradbury
Dr. Emma Cohen de Lara
Dr. Allard den Dulk
Dr. Maxim Kupovykh
Dr. Rebecca Lindner
Dr. Roland Luttens

Series Editing
Dr. Rebecca Lindner

Typesetting & Design
Barlock

Images
1. Network Map - http://www.cecgpr.com/networking.html
2. Collagen density – author's own image
3. Kunsthek Oosterpark – http://maartenbrante.com
4. Illustration by Brian Gable, in Goodhart, D. (2004). Too Diverse? Prospect Magazine, 95, 31.
5. Embedding of de Sitter space in Minkowski space – author's own image

The AUC Undergraduate Journal of Liberal Arts and Sciences is a biannual, interdisciplinary publication showcasing outstanding undergraduate academic papers. The Journal aims to demonstrate the strength of undergraduate scholarship at AUC, to reflect the intellectual diversity of its academic programme, to encourage best research and writing practices, to facilitate collaboration between students and faculty across the curriculum, and to provide students with opportunities to gain experience in academic reviewing, editing and publishing.

FOREWORD

This year's Capstone issue of the *AUC Undergraduate Journal of Liberal Arts and Sciences* coincides with the college's five-year Lustrum celebration – an opportunity to reflect on our many accomplishments since we opened our doors in 2009, as well as to look ahead to future challenges and achievements.
To mark the Lustrum, we are including five Capstone theses selected by the AUC Capstone Awards Committee for the awards of 'Thesis of Distinction' and 'Thesis of Highest Distinction'. The students whose work is included in this issue have graduated in the three majors offered by our institution – Humanities, Sciences and Social Sciences – though their research, as shown in the work that follows, frequently cuts across disciplinary borders.
Each of these 5 theses represents a significant individual achievement - in the student's preparation, writing, and presentation of the Capstone thesis. Read together they demonstrate the breadth and ambition of our students' work, the guidance offered by faculty and Capstone supervisors, and the interest and appreciation of the AUC community.

Dr. Rebecca Lindner,

Series Editor
Head of Studies, Humanities

CONTENTS

Faces in the Cloud: Discovering Motifs in Social Networks

Lotte Romijn

Supervisors: Dr. B.S. Ó Nualláin (AUC, UvA) and Dr. L. Torenvliet (UvA)

ABSTRACT

In social networks, substructures can be identified to study interactions among social actors. Such substructures, called motifs, can provide information about the function of an actor in a social network. However, detecting and counting motifs is computationally hard. In this project, approximation algorithms are implemented that approximate the number of motifs in networks using random colorings of nodes (elements of the network). These approximation algorithms can count paths in large networks significantly faster than exact methods. Networks can be decomposed into graphs. Graphs are mathematical abstractions that model relationships (edges or links) between the components of a network (vertices or nodes). In this project, a real-world Irish forum data set was converted into graphs, in which nodes represent users on the forum and nodes are linked to each other if they posted in the same thread on the forum. An approximation algorithm to count "simple paths" (paths with no repeating nodes) was programmed in Python and used to determine the distribution of paths in the Irish forum graphs. By analyzing the distribution of paths in these graphs, it was observed that the Irish forum contains a few highly active users. Moreover, the path counts indicated that new users on the forum are highly likely to post in the same threads as prolific posters on the forum. In the final stages of this project, the approximation method to count paths was used to develop a novel algorithm for "betweenness centrality." Betweenness centrality is a measure of how much information traverses through a given node in a network. This algorithm is faster than the current fastest known algorithm.

1. INTRODUCTION

Many real-world systems are complex networks that consist of a large number of highly connected interacting components. Examples of complex networks are the World Wide Web, Internet, neural and social networks. Complex networks can be represented as graphs. Graphs are mathematical abstractions composed of nodes (or vertices) connected by edges. Such abstractions are models that show the interconnections between elements of a network. Complex networks contain characteristic patterns and substructures such as cycles or triangles. Such patterns are called network motifs, subgraphs or templates. Counting and detecting motifs is a method of identifying functional properties of network nodes. The distribution of certain motifs indicates how nodes behave in the network. For instance, in Web graphs, a Web host (node) is likely to be a source of spam if it is contained in a large number of triangles (Becchetti et al. 2008). The term "motif" was coined by Milo et al, who found recurrent motifs in biochemical, neurobiological, ecological and engineering networks (Milo et al. 2002).

However, counting and detecting network motifs is computationally hard. In Complexity Theory, it is known that counting the exact number of a given motif cannot be done efficiently (i.e., in time proportional to a polynomial function of the input), unless $P=\#P$. (Counting motifs is a #P-complete problem. The definition is included in Section 2.) The main problem in counting motifs is algorithmic efficiency, which poses a major research challenge. Counting motifs has not been attempted on large real-world networks (more than a few thousand nodes) (Zhao et al. 2010).

The research problem of this project: For a given graph G, how can the number of motifs containing a particular vertex be found efficiently? Because there is no known polynomial method to count motifs, approximation methods need to be explored. Approximation algorithms avoid exhaustive searches with exponential complexity and aim to find an approximation of the exact count in polynomial time (or less). Recently, Gonen and Shavitt developed approximation algorithms for counting motifs, presented in their paper "Approximating Network Motifs" (2009). Their counting method is based on color coding techniques by Alon et al. (1995). Gonen and Shavitt (2009) described novel algorithms for counting the following motifs: k-length cycles, k-length cycles with a chord, (k-1)-length simple paths, and all motifs of size at most four. Taking a graph and particular vertex as input, each algorithm computes the number of occurrences of the given motif containing the input vertex. Here, $k \in O(\log n)$ where n is the number of vertices in the graph. These algorithms approximate the number of motifs in polynomial time. If these algorithms work effectively (i.e., approximate the number of network motifs close to the exact solution), these manageable computation times allow counting motifs in large networks within small margins of error.

In this capstone project, the "simple path" algorithm by Gonen and Shavitt (2009) is developed in Python for the first time. Simple paths are paths with no repeated vertices. The simple path algorithm finds an estimate of the number of paths containing a given vertex. Gonen and Shavitt (2009) proved that this algorithm has an efficient computation time. In this project, the simple path algorithm is tested on small, artificially generated graphs in order to assess its accuracy and computation time. After preliminary analysis, the algorithm is used to count paths in real-world social network data. This social network data is derived from an Irish forum data set containing ten years of Irish online life from the website "boards.ie," Ireland's most prominent forum website. In order to get a structural representation of the data, multiple graphs are created in Python. These graphs represent connections in the data set. For example, nodes can represent users with accounts on the website, and connections are between users if they have posted in the same threads. The main goal of this project is to explore color coding techniques to approximate the number of "simple paths" in randomly generated graphs, assess its accuracy and complexity, and count the number of "simple paths" in the forum data set.

This capstone thesis is structured as follows: Section 2 discusses previous research related to counting motifs. In Sections 3 to 6, the steps for counting simple paths in the forum data set and the results are presented. In Section 7, a novel algorithm for betweenness centrality is described, which forms the main theoretical contribution of this capstone project. In Section 8, the simple path counts of the forum data and theoretical findings are discussed, together with suggestions for future research.

2. LITERATURE REVIEW

2.1 SEARCHING AND DETECTING MOTIFS

Two networks with similar global topology can have varying local structures. In fact, local motifs are increasingly considered to be the small building blocks that are responsible for local functions in a network. Milo et al. (2002) found network motifs in biochemical, neurobiological, ecological and engineering networks. An example of functional properties of motifs was illustrated by Becchetti et al. (2008). They showed that the local number of triangles in large scale Web graphs is an indication of spamming activity. Web hosts that are a source of spam are connected to each other via more triangles than normal, non-spam Web hosts (Becchetti et al. 2008). Moreover, Wu, Harrigan and Cunningham (2011) made a selection of motifs that are most predictive of assessing the quality of Wikipedia articles. They analyzed (bipartite)[1] edit networks in which articles and editors are nodes and edges represent edit activities. It was found that motifs in which editors are linked to multiple articles are suggestive of high article quality, from which it could be inferred that experienced editors produce better quality articles (Wu, Harrigan and Cunningham 2011).

Przulj et al. used the term graphlet to denote a connected network with a small number of nodes. All such 3-5-node graphlets are presented below (Przulj, Corneil and Jurisica 2005).

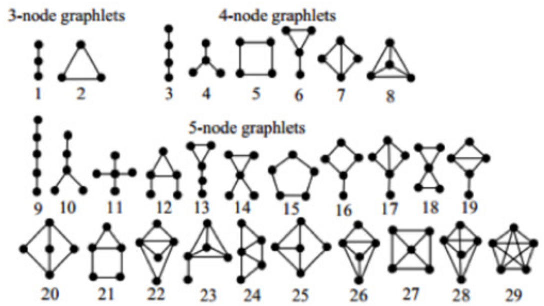


Figure 1. *All 3-5 node graphlets, illustrated by Przulj, Corneil and Jurisica (2005).*

The search for motifs in networks focuses on either induced or non-induced motifs (Gonen and Shavitt 2009). Induced motifs have an additional restriction: an induced motif is a subset of vertices that contains all the edges between those vertices that are present in the original network. In general, searching non-induced motifs is more informative because a vertex in a network could have functions not associated with all of its adjacent edges (Gonen and Shavitt 2009).

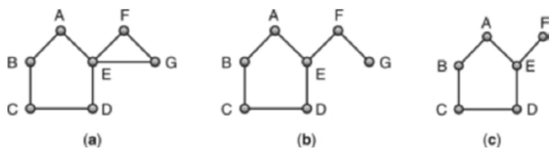


Figure 2. *Graph* b *is a non-induced motif of graph a. Graph* c *is an induced motif of both graph* a *and* b. *(Retrieved from http://openi.nlm.nih.gov/)*

Motif detection is equivalent to the subgraph isomorphism problem, a well-known problem in Graph Theory. Given two graphs G and H, does G contain a subgraph isomorphic to H? In other words, is there a bijection f between a subset of vertices in G and all vertices in H such that any two vertices u and v are adjacent in G if and only if $f(u)$ and $f(v)$ are adjacent in H? The subgraph isomorphism problem is not to be confused with the graph isomorphism problem: the computational problem of determining whether two graphs are isomorphic. It is not known whether the graph isomorphism problem is solvable in polynomial time or whether the problem is NP-complete (Johnson 2005). However, the subgraph isomorphism problem is an NP-complete decision problem (the proof is based on a reduction of the NP-complete clique problem (Cook 1971)). This means that the exact solution cannot be located in polynomial time. By brute-force search, a solution can be found by enumerating all possible combinations of vertices that together form the size of the motif, and checking whether the edges present correspond to the edges in the subgraph. Ullmann described an algorithm in 1976 that can solve the subgraph isomorphism problem more efficiently than brute-force search. The algorithm is based on continuously reducing possibilities while backtracking in a search tree (Ullmann, 1976). It has worst case runtime $O(V!V^2)$, which becomes intractable for large networks.

Counting the number of motifs of a particular vertex builds upon the subgraph isomorphism problem. The subgraph isomorphism problem poses the question whether a graph contains a certain subgraph, while counting motifs amounts to enumerating *how many* subgraphs can be found in which a particular vertex is included. Finding non-induced motifs grows rapidly in computation time with input size, and has not been attempted on large scale real-world networks (Zhao et al. 2010). Reducing this computation time represents a major research challenge.

A brute-force search to count all occurrences of a particular motif requires the enumeration of all possibilities. For instance, finding all triangles in a graph requires finding every pair of two edges with the same vertex as one of their end-vertices, and checking whether there is an edge connecting the other end-vertices of these two edges. Without any form of approximation, the most tractable way to solve this problem uses matrix multiplication, which is of order $O(n^3)$ if schoolbook method is used. Without matrix multiplication, a naive algorithm takes computation time of $O(n^5)$ (Tsourakakis et al. 2009). (There are $\binom{n}{3} < n^3$ ways to choose three distinct vertices. The edges that connect every combination of three vertices can be naively determined by checking the adjacency list. This takes time of $O(n^2)$ (since there are $m < n(n-1)$ edges, and an adjacency list is of

[1] A bipartite graph G can be decomposed into two disjoint sets of vertices U and W such every edge in G connects a vertex in U with a vertex in W. For instance, in the edit network described in the main text, editors are only linked to articles and vice versa.

$O(m+n) \simeq O(n^2)$. Hence, the total computation time is $O(n^3 \cdot n^2) = O(n^5)$.) For larger networks testing such naive algorithms is problematic.

Moreover, finding all paths in a graph could be even harder. Although finding paths is comparable to finding triangles, the complication is that there could be exponentially many of them. A breadth first search tree of a complete graph with *n* nodes has depth *n*, and exponentially many leaves.

A simple backtracking algorithm from the Python library can find all the paths in a graph. A motif in the form of a *k*-length path for a particular vertex can then be found by selecting paths with length *k* of which that vertex is a member. However, the inefficiency of this algorithm is obvious; enumerating all paths in the graph without any restrictions is costly. Moreover, the most simple backtracking procedure is an exponential search, which results in a large computation time for large graphs.

```
graph = {'A': ['B', 'C'],
         'B': ['C', 'D'],
         'C': ['D'],
         'D': ['C'],
         'E': ['F'],
         'F': ['C']}

def find_all_paths(graph, start, end, path=[]):
    path = path + [start]
    if start == end:
        return [path]
    if not graph.has_key(start):
        return [ ]
    paths = [ ]
    for node in graph[start]:
        if node not in path:
            newpaths = find_all_paths(graph, node, end,
            path)
            for newpath in newpaths:
                paths.append(newpath)
    return paths
```

In general, deciding whether a motif is contained in a network is an NP-complete decision problem, equivalent to the subgraph isomorphism problem in Graph Theory. Finding the vertex with maximum number of motifs is NP-hard. In Appendix B, it is explained that NP-complete problems consist of problems in NP to which all other NP-problems can be reduced in polynomial time. NP-complete problems are "decision" problems: there is either a solution (i.e., it is satisfiable), or there is no solution.

Counting the number of a certain motif is #P-complete (Sharp-P). #P-problems is the class of function problems that correspond to their decision problems in NP. Formally, #P-problems are of the form "compute the value of a function *f(x)*," where f is the number of possible solutions that can be accepted by the corresponding NP decision problem (Stockmeyer, 1983). Intuitively, they are at least as hard as the NP-problem, because if it is easy to count the number of solutions, then it must be easy to find one particular solution.

There are a few existing algorithms for counting and detecting non-induced motifs. These algorithms do not find the exact number of motifs, but give an approximation close to the exact answer. These techniques go back to Larry Stockmeyer's (1983) theorem of approximate counting. He proved that for every #P-problem there is a randomized approximate algorithm that determines the count in polynomial time, provided that there is an oracle (a 'black box' capable of solving a problem in NP or harder). This means that for a particular instance *a* of problem *P* and $\epsilon > 0$, the algorithm returns the count *C* with a high probability such that $(1 - \epsilon)P(a) \leq C \leq (1 + \epsilon)P(a)$, where P(a) is the exact count for instance a (Stockmeyer, 1983).
The randomized algorithm is in principle an (ϵ, δ)-approximation method. (Also, Gonen and Shavitt's algorithms are randomized approximate counting algorithms. The details are included in Section 3.)

2.2 COLOR CODING

Many algorithms to count and detect motifs make use of the color coding technique. Alon et al. (1995) proposed the color coding technique to detect simple paths, cycles and bounded treewidth subgraphs. (Bounded treewidth: the width among all possible tree decompositions is bounded (Bodlaender and Koster 2008).) Recently, the color coding technique has been used to detect signaling pathways in PPI-networks (Alon et al. 2008).

The color coding technique is based on random assignments of colors to the vertices of an input graph. It can detect specific subgraphs tractably by only considering specific color assignments, in time proportional to a polynomial function of the input $n = |V|$. If the assignment of colors is repeated sufficiently many times (polynomial with *n*), the method will find a specific occurrence of the motif of size $O(\log(n))$ with a high probability. A simple example of color coding to count cycles is outlined in Appendix A. Multiple algorithms use elements of or are entirely based on the color coding technique (Arvind and Raman (2002), Alon et al. (2008)). Arvind and Raman (2002) used color coding for counting the number of subgraphs isomorphic to a bounded treewidth graph. Alon et al. (2008) described a polynomial time algorithm for approximating the number of non-induced occurrences of trees and bounded treewidth subgraphs with $k \in O(\log n)$ vertices. In 2007, Hüffner, Wernicke and Zichner presented various algorithmic improvements for color coding that lead to savings in time and memory consumption. Zhao et al. (2010) have recently shown that using color coding in addition to parallel programming can find motifs in networks with millions of nodes. They have combined parallelization of color coding with stream-based partitioning. In other words, sequential color coding was shifted to the use of multiple processors that use color coding in parallel on partitions of the original graph. Their 'ParSE' algorithm was tested on large scale synthetically generated social contact networks for urban regions.

In this capstone project, the color coding is used to count simple paths. Gonen and Shavitt's algorithm to count simple paths will be explained in detail, together with its implementation in Python and performance on the forum data set. By testing the algorithm on the forum data, Gonen and Shavitt's simple path algorithm is implemented for the first time on a real-world network.

3. GONEN AND SHAVITT'S SIMPLE PATH ALGORITHM

Gonen and Shavitt's algorithm to find simple paths uses the color coding technique by Alon, Yuster and Zwick (1995). It approximates the number of paths of length $k - 1$, where *k* is the number of colors in the color set and length refers to the number of edges in the path. The input is the graph *G*, a vertex $v \in V$, the path length $k - 1$, fault tolerance ϵ, and error probability δ.

$t = \log(\frac{1}{\delta})$ and let $s = \frac{4k^k}{\epsilon^2 k!}$

1. *For j = 1 to t*
   For i = 1 to s
   i *Color each vertex of G independently and uniformly at random with one of the k colors.*
   ii *For all* $u \in V$ $Ci(u, \emptyset) = 1$
   iii *For all* $l \in [k]$ $Ci(v,\{l\}) = \begin{cases} 1 & \text{if } col(v)=l \\ 0 & \text{otherwise} \end{cases}$
   iv *For all* $S \subseteq [k]$ *s.t.* $|S| > 1$, $C_i(v,S) = \sum_{u \in N(v)} C_i(u, S \setminus \{col(v)\})$
   v. $P_i(v,[k]) = \sum \binom{k}{l=1} \sum_{(u \in N(v)} \sum_{(S_1,S_2) \in A_{(l,v)}} C_i(v,S_1) C_i(u,S_2)$
   where $A_{(l,v)} = \{(S_i,S_j) \mid S_i \subseteq [k], S_j \subseteq [k], S_i \cap S_j = \emptyset$
   $|S_i| = l, |S_j| = k - l\}$
   vi. *Let* $X_i^v = P_i(v, [k])$
   b. *Let* $Y_j^v = \frac{\sum_{i=1}^{s} X_i^v}{s}$

∩

2. *Let Zv be the median of* $Y_1^v, Y_2^v, \ldots, Y_t^v$
3. *Return Zv. kk/k!*

This algorithm is an (ϵ, δ)-approximation for counting simple paths of length $k - 1$ containing vertex v.[2] "Simple" indicates that there are no repeated vertices in the path. $P_i(v,S)$ is the number of colorful paths containing v using colors in *S* at the *i*th coloring. $C_i(v,S)$ is the number of colorful paths for which one of the endpoints is *v* using colors in *S* at the *i*th coloring. The algorithm finds an approximation of the number of paths within $[(1 - \epsilon)r, (1 + \epsilon)r]$, where *r* is the actual number of paths in the graph, with a probability of at least $1 - 2\delta$.

3.1 EXPLANATION OF APPROXIMATION IN STEPS

The steps necessary to approximate the number of simple paths of length $k - 1$ can be divided into two steps: the color coding and the counting step. The two steps will be discussed separately.

3.1.1 COLOR CODING STEP

As in Alon et al. (2008), let *r* represent the total number of simple paths that are present in the graph *G*. The result of the color coding step is to get an estimate of *r*. The color coding step colors every vertex of *G* uniformly at random with one of the colors in $[k] = \{1,\ldots,k\}$. A simple path of length $k - 1$ consists of *k* vertices and becomes "colorful" if all the vertices have a distinct color. The probability that a simple path becomes colorful is $p = k! / k^k$. (The path counting step described below

[2] Gonen and Shavitt (2009) call a path 'adjacent to' a vertex v if that vertex is contained in the path. I have left this terminology out for clarity.

explains how the actual number of colorful paths are counted for each coloring).

Call $\mathcal{F}$ the set of all simple paths in G. For every element $F\in\mathcal{F}$, let x_F be the random variable that has value 1 iff the simple path is colorful and 0 otherwise. For one coloring, let $X = \sum_{F\in\mathcal{F}} x_F$ denote the random variable that counts the number of colorful paths in G. The expected value of X is $E(X) = rp$. (X has a binomial distribution with expectation rp and variance $rp(1 - p)$.)

Let F, F' be two distinct paths $\in\mathcal{F}$. The probability that they are both colorful is strictly smaller than p, since it can only be p if they cover the exact same vertex set. Therefore, the covariance of the occurrence of both paths in a given coloring can be described as follows: $Cov(x_F, x_{F'}) = E(x_F x_{F'}) - E(x_F)E(x_{F'}) \leq p$. Hence: $Var(X) = \sum_{F\in\mathcal{F}} Var(x_F) + \sum_{F\neq F'\in\mathcal{F}} Cov(x_F, x_{F'}) \leq rp(1-p) + r(r-1)p = r^2p - rp^2 \leq r^2p$. If p is small, p^2 is very small and the upper bound is tight. When taking Y as the average of s times X (s colorings of G), then $E(Y) = E(X) = rp$ and $Var(Y) = Var(X) / s \leq \frac{r^2p}{s}$.

Next, Chebyshev's Inequality is used, which states that for a random variable Y, some constant k and standard deviation σ: $P(|Y-E(Y)| \geq k\sigma) \leq \frac{1}{k^2}$. This inequality says that in any probability distribution nearly all values are close to the mean. In the color coding, the probability that Y is $\epsilon\, rp$ smaller or bigger than the expected value is $P(|Y-rp| \geq k\sigma)$ where $k\sigma$ is replaced by $\epsilon\, rp$.

Since $\sigma \leq \sqrt{\frac{r^2p}{s}}$,by Chebyshev's inequality we find that

$$P\left(|Y-rp| \geq k\sqrt{\frac{r^2p}{s}}\right) \leq \frac{1}{k^2}.$$

Setting $k\sigma = \epsilon\, rp$, so $k = \frac{\epsilon rp}{\sqrt{\frac{r^2p}{s}}}$ and $\geq \epsilon rp \leq \frac{1}{\left(\frac{\epsilon rp}{\sqrt{\frac{r^2p}{s}}}\right)^2} =$

$$\frac{r^2p}{\epsilon^2 r^2 p^2 s} = \frac{1}{\epsilon^2 ps}$$

For example, if $s = 4 / \epsilon^2 p$, this probability is at most $\frac{1}{4}$. Call this probability p_Y. Let's stick with this value for p^Y in order to make s not too large (s is the number of colorings of G) and the probability $P(|Y-rp| \geq \epsilon rp)$ reasonably small.

Y is computed t times independently, after which the median Z is taken of the t Y values. Hence, Z is less than $(1 - \epsilon)rp$ (or more than $(1 + \epsilon)rp$) if at least half of the values are less than $(1 - \epsilon)rp$ (or more than $(1 + \epsilon)rp$). The calculations below consider the case $(1 - \epsilon\, rp$.

Define T_i to be $T_i = \begin{cases} 1 \text{ if } Y_1 \leq (1 - \epsilon)rp \text{ with } p_y \leq \frac{1}{4} \\ 0 \text{ otherwise} \end{cases}$, *f or i in* $[1,t]$.

All T_i's are independent.
Let $T = \sum_{i=1}^{t} T_i$.
Then the probability that T is less than $(1 - \epsilon)rp$ is $P(T \geq \frac{t+1}{2})$ for uneven t and $P(T \geq |\frac{t}{2}| + 1)$ for even t. Let's denote this by $P(T \geq \frac{t}{2} + 1)$ for both even and uneven t. Since each $T_i \sim Bern(p_Y)$, $T \sim Bin(t, p_Y)$.

For all t, $P(T \geq \left|\frac{t}{2}\right| + 1) = \sum_{i=|\frac{t}{2}|+1}^{t} \binom{t}{i} p_y^{\,i} (1 - p)_y^{t-i}$ where $p_Y \leq \frac{1}{4}$. A property of the binomial distribution is that it is monotone decreasing for $i > (t + 1)p_Y$ (The proof is included in Appendix B). Since $\left|\frac{t}{2}\right| > (t + 1)\frac{1}{4}$, we know that $P(T = \left|\frac{t}{2}\right| + 1) > P(T = \left|\frac{t}{2}\right| + 2) > P(T = \left|\frac{t}{2}\right| + 3)$.

We want to give a bound on $P(T \geq \left|\frac{t}{2}\right| + 1)$. Alon and Spencer (2004) use Chernoff bounds to show that this probability is bounded by 2^{-t} (Alon and Spencer 2004, Appendix A).[3]

The same result holds for $(1 + \epsilon)rp$. The largest probability for the median of the Yi's to have a value above $(1 + \epsilon)rp$ is smaller than 2^{-t}.

Hence, if $t = \log(\frac{1}{\delta})$ (base 2), then with probability at least $1 - 2 \cdot 2^{-t} = 1 - 2 \cdot 2^{-\log(\frac{1}{\delta})} = 1 - 2\delta$ the median Z of $Y_1,...,Y_t$ is in $[(1 - \epsilon\, rp, (1 + \epsilon\,)rp]$.

The total number of colorings is: $O\left(\frac{\log(\frac{1}{\delta})}{\epsilon^2 p}\right) = O\left(\frac{e^k \log(\frac{1}{\delta})}{\epsilon^2}\right)$, since there are $s \cdot t$ loops and $p = k! / k^k > 1/e^k$. After the calculation of Z, it is divided by p to get an estimate of r (see the expected value of Y above): $r = Z / p = Z \cdot k^k / k!$.

3.1.2 PATH COUNTING STEP

The function of the counting step is to count the number of colorful simple paths for each coloring. Define [k] to be a set consisting of k colors represented by distinct integers. For each vertex v and subset S of the color set $[k]$, $C_i(v,S)$ is defined as the number of colorful paths for which one of the endpoints is v using colors in S at the ith coloring. First of all, $C_i(v,\{\})=1$.

[3] In their book A Probabilistic Method, Alon and Spencer (2004) justify this bound in Appendix A. However, the reader is left unsure which theorem(s) are used for this specific bound. Together with my probability and statistics teacher at AUC, M. de Gunst, I have attempted to prove the bound 2-t, using Chernoff bounds and the theorems stated by Alon and Spencer (2004). These efforts have been inconclusive so far.

Given a specific color l, for all $v \in V(G)$: $C_i(v,\{l\}) = \begin{cases} 1 \text{ if } col(v)=l \\ 0 \text{ otherwise} \end{cases}$

In this case, since the color set consists of one single color l, only vertex v is considered and is colorful if it has the same color as l.

For each vertex v and color set S such that $|S| > 1$: $C_i(v,S) = \sum_{(u\in N(v)} C_i(u, S \setminus \{col(v)\})$. This counts the number of colorful paths of length $k - 1$ with v as an end-vertex by iterating through the neighbors of v, and their neighbors etc. to find vertices with colors that are still in the color set.

Then $P_i(v,[k])$ (step 1.a.v in the simple path algorithm) is computed. This counts the total number of colorful paths that contain v using all colors in the color set. It considers all the neighbors u of v, and all partitions of the color set $[k]$ into S_1 and S_2. (It is important to note that the algorithm considers one particular partition twice, where the second partition is similar to the first but S_1 and S_2 changed around.)

3.2 EXAMPLES AND MODIFICATIONS FOR THIS PROJECT

For example, let P_1 be a path of length 3 consisting of 4 vertices u, v, w, x where u is blue, v is yellow, w is green and x is red. We are approximating all paths of length 3 containing vertex v. The complete color set is $[k] = \{0,1,2,3\}$ where *0* stands for blue, 1 yellow, 2 green and 3 red. All the partitions of the color set considered in the algorithm are *[{0,1},{2,3}], [{2,3},{0,1}], [{0,2},{1,3}], [{1,3},{0,2}], [{0,3},{1,2}], [{1,2},{0,3}], [{0},{1,2,3}], [{1,2,3},{0}], [{1},{0,2,3}], [{0,2,3},{1}], [{2},{0,1,3}], [{0,1,3},{2}], [{3},{0,1,2}], [{0,1,2},{3}], [{0,1,2,3},{}]*. Here, the first element of each partition is S_1 and the second S_2. It considers only S_1's with at least one element up to k elements. In this particular coloring of the vertices, it is inferred that when iterating over all partitions, only the partitions *[{0,1},{2,3}]* and [*{1,2,3}*,*{0}*] return non-zero answers in the computation of $\sum_{l=1}^{k}\sum_{u\in N(v)}\sum_{(S_1,S_2)\in A_{l,v}} C_i(v,S_1)\cdot C_i(u,S_2)$. Intuitively, one can see that the color of v (i.e., 1) needs to be in S_1 and that S_1 should have at least two elements (since vertex v has two vertices to its left including itself (v and u), and three vertices to the right including itself (v, w, x)). The colors in S_1 should match the colors of either the two vertices on the left, or the three vertices on the right. The colors in S_2 are the remaining colors in the complete color set. Then, $C_i(v,S_1) \cdot Ci(w,S_2) = 1$ for [*{0,1}*,{2,3}] and $C_i(v,S_1) \cdot C_i(u,S_2) = 1$ for [*{1,2,3}*,*{0}*]. Hence, $P_i(v,[k]) = 2$. This colorful path is counted twice and, consequently, the total count of simple paths containing vertex v should be divided by 2. Gonen and Shavitt do not state this factor 2 in their algorithm, but in this project it has been implemented.

When iterating over all neighbors u, it is important not to over-count paths. If v is an end-vertex of one single path $P_2 = (v, a, b, c)$ with colors 0,1,2,3 respectively, looking at the partitioning [*{0,1,2,3}*,{}] will return 1 for every neighbor u of v in the summation $\sum_{l=1}^{k}\sum_{u\in N(v)}\sum_{(S_1,S_2)\in A_{l,v}} C_i(v,S_1)\cdot C_i(u,S_2)$. It does not matter what color u has since S_2 is empty and so $C_i(u,\{\})=1$ for all u. Hence, path P_2 is counted u times instead of once. In the implementation of the algorithm in this project, the partition where S_1 contains the full color set $[k]$ is separated from the other color sets. This results in the following formula (including the division by 2 stated above):

$$P_i(v,[k]) = \frac{\left(\sum_{l=2}^{k-1}\sum_{u\in N(v)}\sum_{(S_1,S_2)\in A_{l,v}} C_i(v,S_1)\cdot C_i(u,S_2)\right)}{2} + C_i(v,[k])$$

However, this is still incorrect. One can see that $P_i(v,[k])$ returns 1 in the summation for the color partitioning [*{0}*,*{1,2,3}*]. In path P_2, v is an end-vertex with color 0. If $|S_1| = 1$ and S_1 contains color 0, it will count path P_2 since $C_i(v,S_1)$ and $C_i(u,S_2)$ are both 1. Therefore, path P_2 is counted twice; once by the summation for $S_1 = \{0\}$ and once by $C_i(v,[k])$.

In order not to double-count paths with v as end-vertex, either the color sets with $|S_1| = 1$ or $C_i(v,[k])$ should be disregarded. See the revised counting $P_1(v,[k])$ below, where the sum only considers color partitions where S_1 has size 2 up to $k - 1$.

The algorithm used in this project implements Gonen and Shavitt's algorithm as follows:

$$P_i(v,[k]) = \frac{\left(\sum_{l=2}^{k-1}\sum_{u\in N(v)}\sum_{(S_1,S_2)\in A_{l,v}} C_i(v,S_1)\cdot C_i(u,S_2)\right)}{2} + C_i(v,[k])$$

where $A_{l,v} = \{S_i,S_j\} \mid S_{i\ i} \subseteq [k], S_{ji} \subseteq [k], S_i \cap S_j = \emptyset, |S_i|=l, |S_j| = k-l\}$

3.3 COMPLEXITY

In the color coding step it was shown that the number of colorings is of $O\left(\frac{e^k \log(\frac{1}{\delta})}{\epsilon^2}\right)$

Each path counting step is of $O(2^k |E|)$, where $|E|$ is the number of edges in the graph:
Proof: We want to find the total complexity of $P_i(v,[k])$. This is done in two parts:

Part 1: Assume $C_i(v, S)$ is known for all v.

Assuming $C_i(u, S)$ is known for any vertex u and color set S and taking $P_i(v,[k]) = \sum_{l=1}^{k} \sum_{u \in N(v)} \sum_{(S_1,S_2) \in A_{l,v}} C_i(v,S_1) \cdot C_i(u,S_2)$ in the original simple path algorithm by Gonen and Shavitt (2009), the time of computing $P_i(v,[k])$ for all v is of order $\sum_{l=1}^{k} \sum_{v \in V} \binom{k}{l} \deg(v)$. The sum over $\binom{k}{l}$ represents all combinations of color sets of length 1 to k. The sum over the degrees of v represents all u's considered in the computation of $C_i(u, S)$. This sum is equal to $\sum_{l=1}^{k} \binom{k}{l} \sum_{u \in N(v)} \deg(u) \leq 2^{k-1} \cdot 2|E| = 2^k |E|$, by the Handshaking Theorem (the sum of all degrees equals twice the total number of edges).

Part 2: Find complexity of $C_i(v,S)$ for all v.

Also, the complexity of finding $C_i(v,S)$ for all v (the number of colorful paths of length $k - 1$ with v as end node) is of $O(2^k |E|)$. It is equal to $\sum_{v \in V} \sum_{u \in N(v)} O(2^k) + \sum_{v \in V} O(k) = O(\sum_{v \in V} \deg(v) \cdot 2^k)$ (*since $O(k)$ smaller than* $O(2^k)$) $= O(2k|E|)$ by the Handshaking Theorem; maximally all combinations of color sets are considered for all neighbors of all vertices in the graph (first part) and each vertex is given a color out of k colors (second part).

By part 1 and part 2, the total complexity of $P_i(v,[k])$ is $O(2^k|E|) + O(2^k|E|) = O(2^k|E|)$. (Once all $C_i(v, S)$ are known ($O(2^k |E|)$), compute $P_i(v,[k])$ ($O(2^k|E|)$)).

The path counting step is performed $O\left(\frac{e^k \log(\frac{1}{\delta})}{\epsilon^2}\right)$ times.

Hence, the product of the color coding step and the path counting step gives the total complexity of the simple path algorithm:

$$O\left(\frac{e^k \log(\frac{1}{\delta})}{\epsilon^2} \cdot 2^k|E|\right) = O\left(\frac{(2e)^k \log(\frac{1}{\delta})}{\epsilon^2}|E|\right)$$

4. DATA AND METHODS

The simple path algorithm discussed in the previous section is written in Python code. First, it is tested on random graphs which can easily be generated by the Python package NetworkX. NetworkX has many built-in functions to create artificial graphs, such as random, small-world and preferential attachment graphs. In Section 5, "Implementation and preliminary results," the initial performance of the algorithm is discussed for complete graphs, since it is easy to mentally compute the exact amount of paths in complete graphs. Moreover, the accuracy and complexity of the algorithm are tested and analyzed. In Section 6, the motif counts of the real-world Irish forum data are presented. Here, the Irish forum data is used as input to the simple path algorithm, where subsets of the data have been extracted and converted to graphs.

4.1 FORUM DATA

The real-world complex networks for which the number of simple paths is counted is generated from an Irish forum data set. This data set was put online in 2008 for the "boards.ie SIOC Data Competition." (SIOC stands for Semantically-Interlinked Online Communities Project.) The complete data set contains ten years of Irish online life from Ireland's largest community website "boards.ie" over the years 1998-2007. Since the foundation of the website in 1998, over 36 million posts have been made and the current posting rate is around 17.000 a day (retrieved from http://www.boards.ie/content/about-us). The data set is a large collection of RDF-files (Resource Description Framework), in which each file contains a post in a thread on a forum. The RDF-files have a tree-like structure, in which the boards website contains multiple threads and each thread contains board posts that are chronological replies to each other.

In this project, useful information from the data set is extracted by parsing the RDF-files. Such information involves the topic of the post (title), the username of the person who posted, his profile, and the thread which contains the posts. The Python package that is used for this purpose is rdflib, which can query and extract certain elements from an RDF-file. After that, graphs are created that represent the structure of the data set by using Python's package NetworkX. For the main analysis of this project, graphs are constructed in which nodes represent users with accounts on the boards website, and edges the connections between users if they posted in the same threads. Initially, the data of the years 1998-2000 is used for analysis, since these are the first years of the boards website and contain the fewest nodes (176, 621, 1033 respectively). The distribution of simple paths in these graphs is compared to artificial data, such as randomly generated graphs, preferential attachment graphs, and small-world graphs.

While testing the algorithm on the data sets, the algorithm is revised and further optimized. The results of the motif counts are analyzed and compared, and further analysis of these motif counts yielded conclusions about the Irish forum data set. These conclusions are presented in Section 8.

5. IMPLEMENTATION AND PRELIMINARY RESULTS

The implementation of Gonen and Shavitt's simple path algorithm in Python is done by using the packages NetworkX and Numpy. NetworkX can generate graphs as objects to which weights and labels are attributed. Hence, in each coloring, a node can be labelled with its color. Numpy provides a library of mathematical functions to perform computations (such as computing the average or median) on large arrays.

The code of the algorithm contains a few functions. $C_i(v,S)$ has been separated from the main function $P_i(v,S)$. Moreover, the color partitions are represented by integers, where each S_1 of every partition is represented by an integer. This is done by an additional script that converts color sets into bitsets and vice versa. A bitset is a compact representation of a set of non-negative integers as a single integer. By encoding sets as integers, operations on small sets can be done very efficiently.

For example, the set {6,4,3,1} is represented by the integer $2^6 + 2^4 + 2^3 + 2^1 = 90$. It can be shown that each finite subset of non-negative integers can be uniquely represented as an integer and each non-negative integer can be uniquely decomposed into a finite subset of non-negative integers. The complete bitset of n elements ({0,1,2,3,..., $n - 2$, $n - 1$}) is represented by the number $2^n - 1$. As an example for the color algorithm, for the color partition [{*0,2*},{*1,3*}] we have that S_1 is 5 and S_2 is 10. In the bitsets script, S_2 is computed as the complement of S_1: take the complete bitset of the total color set – i.e. $2^3 + 2^2 + 2^1 + 2^0 = 15$ – and calculate S_2 by the difference $15 - 5 = 10$.

5.1 PRELIMINARY FINDINGS AND CALCULATIONS

First, the simple path algorithm is compared to an exact algorithm for counting paths, using complete and random graphs with less than 40 nodes. The exact algorithm uses depth first search and has exponential complexity. It makes use of the built-in NetworkX function all_simple_paths and extracts the paths with a specific length containing vertex v. The exact algorithm can find the exact number of paths for any graph, but as the graph size grows linearly the computation time grows exponentially. For complete graphs the exact count can also be arrived at by mental computation (See Section 5.1.1).

5.1.1 VERIFICATION BY CALCULATIONS FROM COMPLETE GRAPHS

Say we have a complete graph of 10 nodes and want to find paths of length 2 with a specific vertex in it. The length refers to the number of edges in the path. Hence, it contains three vertices, of which one is v. There are $\binom{9}{2}$ ways to choose two other vertices. Then we need to multiply by 3, because 3 vertices can be ordered in $3! = 6$ ways and, because of symmetry, there are $6/2 = 3$ distinct paths. This results in 108 paths.
The same holds for paths of length 3, containing 4 vertices. There are $\binom{9}{3}$ ways to choose three other vertices. Then we need to multiply by 12, because 4 vertices can be ordered in $4! = 24$ ways and, because of symmetry, there are $24/2 = 12$ distinct paths. The end result is 1008 paths.(For cycles, the calculations are easier. The number of cycles of length 3 that contain vertex v in a complete graph of 10 nodes is $\binom{9}{2}$, since any two other nodes in the graph can be picked.)

5.1.2 VERIFICATION BY EXPERIMENTS

Figure 3 below illustrates how much the estimate of the simple path algorithm deviates from the count by the exact algorithm for random graphs, with $\epsilon = 0.1$ and $\delta = 0.05$. These random graphs all contain 2/3 of the total number of possible edges. In the figure, the path length is the number of edges in the path. In order to compare the exact count with the color algorithm, the error margin is $[(1 - \epsilon)r, (1 + \epsilon)r]$, where r is the exact number of paths in the graph. By the color coding step of the simple path algorithm, the probability that the approximation is within this region is at least $1\text{-}2\delta = 0.9$. Figure 4 shows the same results for path length 3. For every number of nodes a new random graph is created.

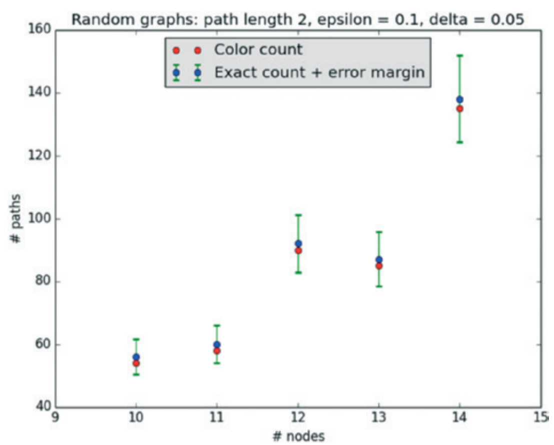


Figure 3. *Simple path counts of length 2 for random graphs with 2/3 possible edges. For each number of nodes a new graph is created.*

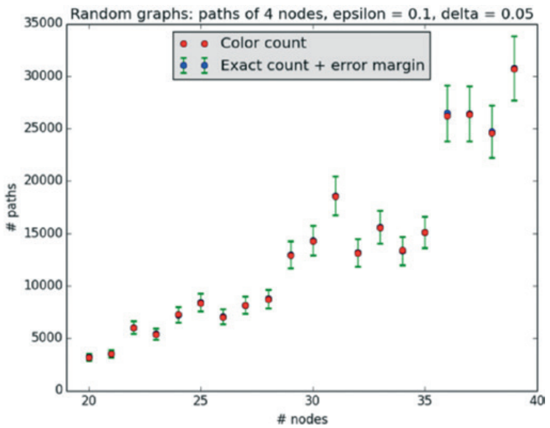


Figure 4. *Simple path counts of length 3 for random graphs with 2/3 possible edges.*

5.2 QUESTIONS AND OBSERVATIONS AFTER THE FIRST IMPLEMENTATIONS

5.2.1 REDUCING THE NUMBER OF ITERATIONS

By comparing the exact exponential algorithm with the simple path approximation, it was found that the number of iterations can be reduced in the approximation algorithm. For instance, we know that a vertex in a complete graph of 10 nodes is part of 108 paths of length 2. Table 1 shows what the effect is of bringing *s* down in the original algorithm for counting paths of length 2. In the last row of the table, *s* is computed as stated in the original algorithm and Ð was taken to be 0.1 (i.e., $s = 4k^k / \epsilon^2 k! = 1800$ ($k = 3$ for path length 2)). *t* is chosen to have a value 5, and so δ is equal to e^{-5}. Performing 200 experiments for each *s* and *t*, Table 1 shows the mean and standard deviation of the number of paths containing a random node in the 10-complete graph. The standard deviation increases as the number of iterations decrease by choosing a lower value of *s*.

t	s	Number of iterations	Mean	Standard deviation
5	20	100	107.166375	6.722599292
5	90	450	107.85075	3.584042681
5	180	900	108.01275	2.277940394
5	270	1350	107.99725	1.998800875
5	360	1800	108.150938	1.535474752
5	540	2700	108.042375	1.245173181
5	720	3600	108.0345	1.199928865
5	900	4500	107.92755	1.023959715
5	1080	5400	108.06	0.980441546
5	1260	6300	107.981786	0.87689016
5	1440	7200	108.090234	0.878974341
5	1620	8100	108.073917	0.747840553
5	1800	9000	107.9613	0.64225233

Table 1. *Counts for length 2 paths in a 10-complete graph with varying s and t*

If we assume that for a large number of experiments the distribution becomes normal, we can construct approximate 95% confidence intervals for the number of paths of length 2 containing a given node (given by the mean ± 2·standard deviation). For *s* = 20 this interval is given by [93.7, 120.6]. This interval traps the true value (108) and only slightly extends outside the interval [(1 − ϵ)108, (1 + ϵ)108] = [97.2,118.8] (i.e., the interval which traps the true number of paths with probability at least 1 - 2δ = 1 − 2e^{-5} = 0.986). Since even *s* = 20 is enough to generate a value around 108 on average and the interval is not too large, it is suggested that the number of iterations of the original algorithm can be reduced significantly.

The same analysis can be performed for larger graphs. For 200 experiments, Table 2 shows the means and standard deviations of the number of calculated paths for a complete graph with 200 nodes. The claim that s can be brought down to 20 still holds in this case.

The true count for 200 nodes is $\frac{\binom{199}{2} 3!}{2} = 59103$, which is close to the calculated mean. Moreover, approximately 95% of the values are well within the (1± ϵ)-factor of the true value: 2 · stdev = 2 · 772.44 = 1544.88 < ϵ · 59103 = 0.1 · 59103 = 5910.3.

From this result it can be concluded that reducing the number of iterations does not yield unacceptable errors in larger graphs.

#nodes	t	s	Number of iterations	Mean	Standard deviation
200	5	20	100	59088.258	772.44

Table 2. *Counts for length 2 paths in 200-complete graph*

However, the probability that a path becomes colorful $p = k! / k^k$ is smaller for larger path lengths. Hence, it is expected that for longer path lengths a higher number of iterations is required to obtain accurate results. The assignment $s = 4k^k / \epsilon^2 k!$ from the original algorithm also means that the number of iterations increases as *k* increases. For now, we stick with *s* = 20 and *t* = 5 for the main analysis of the forum graphs in this project. These graphs have around 200-1000 nodes and mainly short paths are counted. More experiments are needed to assess whether multiplying the number of iterations by $k^k / k!$ as *k* increases produces results with similar accuracy for all path lengths.

5.2.2 TESTING THE COMPLEXITY

The total complexity of the simple path algorithm is

$$O\left(\frac{(2e)^k \log(\frac{1}{\delta})}{\epsilon^2} |E|\right)$$

The number of iterations is of $O\left(\frac{e^k \log(\frac{1}{\delta})}{\epsilon^2}\right)$.

In order to speed up the analysis of the data in the following, *s* is chosen to be 20 and *t* to be 5. By choosing *s* = 20 and *t* = 5, the number of iterations is always 100. This reduces the complexity by a constant factor. The complexity is now: $O(2^k \cdot 100|E|)$.

The following graph shows the time complexity graph of the simple path algorithm for graphs with varying densities.

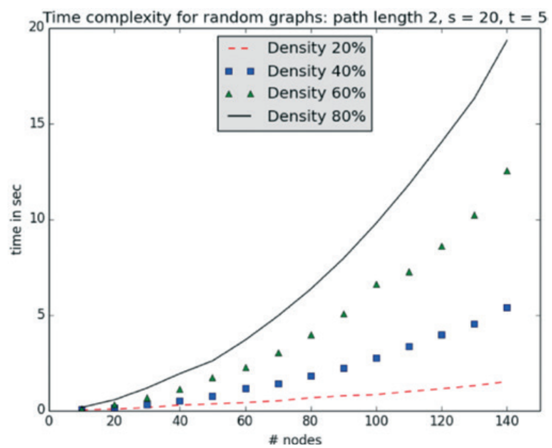


Figure 5. *Time complexity for random graphs with varying densities. For each number of nodes 4 new random graphs were created with different density. For each graph the average time was calculated over all nodes.*

For each *n* number of nodes there are $\frac{n(n-1)}{2} \cdot$density edges. Therefore, when the density is kept constant the number of edges is a quadratic function of the number of nodes. The complexity of the simple path algorithm $O(2^k \cdot 100|E|)$ depends only on the number of edges $|E|$ for *k* constant. Hence, the time complexity should also follow a second degree polynomial relationship. Performing a polynomial regression analysis in the statistical program R does not reject the hypothesis that the time data of Figure 5 could be a second degree polynomial function of the number of nodes. For instance, for density 80% it was found that the data can be modelled by the equation $y = 1.070\text{e-}03\ x^2 + 2.510\text{e-}03\ x + 4.301\text{e-}02$, where y is the independent variable "time" and *x* the dependent variable "number of nodes". The residual standard error is 0.06785. The F-statistic was calculated to test the null hypothesis whether there is a lack of fit. A big F-value (6.933e+04 on 2 and 11 degrees of freedom) and a p-value of 2.2e-16 indicate that the null hypothesis is rejected. However, these results should be interpreted with care; all that a statistically significant F-test says is that the data supports evidence that the best-fitting quadratic model has at least one term with a non-zero coefficient. Below the fitted line is plotted in red, together with the data points, and the residual errors are plotted versus their fitted values. It can be seen that there is no distinct trend in the distribution of points. The plot in the lower left is a Q-Q plot, which suggests that the residual errors are from a normal distribution.

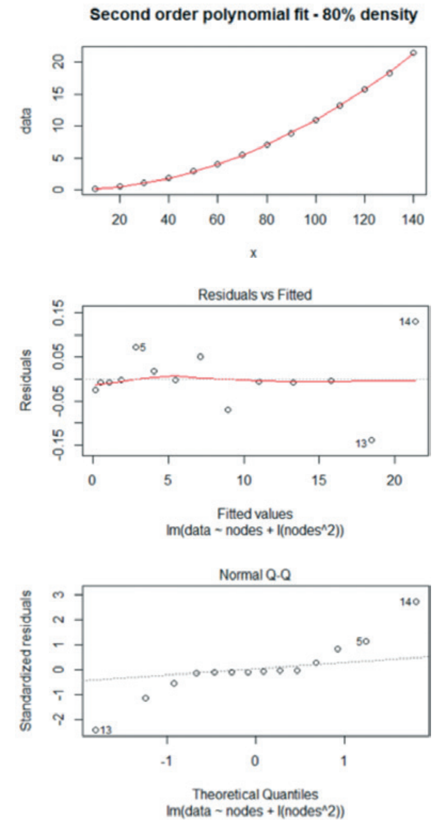


Figure 6. *Regression analysis in R showing plot of the fit and the data, plot of the residuals and fitted errors, and a Q-Q plot of the residuals.*

For lower densities the same results were obtained, with lower coefficients of the two polynomial terms which makes the second degree polynomial is more flat.

5.3 GRAPH STRUCTURE ANALYSIS IN NETWORKX

NetworkX has many built-in functions to perform analysis on networks. It was found that the boards forum data of the years 1998-2000 have the properties shown in Table 3.

NetworkX analysis	1998 boards.ie	1999 boards.ie	2000 boards.ie
Nodes	176	621	1033
Edges	2132	17799	40429
Largest betweenness	0.125	0.05	0.035
Number of components	27	64	109
Number of singletons	26	63	108
Number of nodes in largest component	150	558	925
Diameter	4	4	4
Average degree connected component	28.4	63.7	87.4
Density	0.138441558	0.092457535	0.075847798

Table 3. *NetworkX analysis of 1998-2000 SIOC data*

From this analysis it can already be seen that it is highly likely that these networks can be classified as either small-world or scale-free networks, since the diameters are small. In small-world networks most nodes are not neighbors from each other but each node can be reached in only a few hops. Here, the densities of the graphs are small and it can be concluded that most nodes are not neighbors. However, the diameter is still only four and every node can reach another in just four steps. Many social networks exhibit small-world properties. Small-world networks are characterized by high clustering coefficients and small average path lengths (Watts and Strogatz 1998). A characteristic of scale-free networks is that their degree distribution follows a power law, i.e. the fraction of nodes in the network with degree k is proportional to $k^{-\gamma}$, where $2 < \gamma < 3$ (Barabási 2009). Hence, on a double logarithmic scale the degree distribution is a straight line (Barabási 2009). The real-world forum graphs do not show this characteristic (Figure 7):

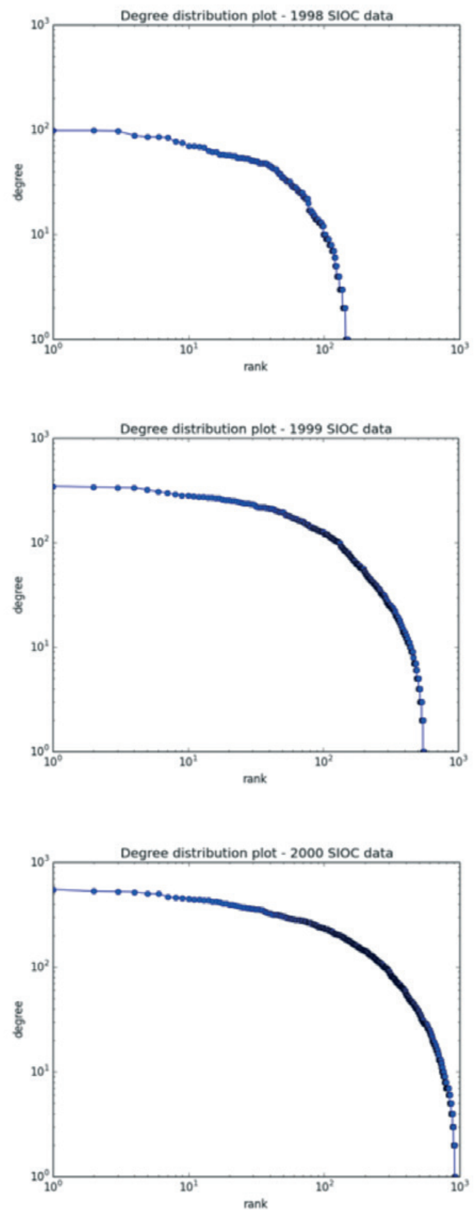


Figure 7. *Degree rank plots for 1998, 1999 and 2000 graphs. Plots the degree distributions in a log-log plot. The rank is the number of nodes with a certain degree.*

5.4 COMPUTATION TIME FOR GRAPHS WITH DIFFERENT DEGREE DISTRIBUTIONS

Graphs can have different degree distributions. It is expected that if a node has many neighbors, the simple path algorithm will take more time to produce an estimate of the number of paths containing that node. Figure 8 shows the sorted times for two graphs: a random and a preferential attachment graph, each having 150 nodes and 2144 edges. The random graph has a degree distribution that is approximately normally distributed (i.e., the tails of the distribution are short). The degree distribution of the preferential attachment graph follows a power law. A preferential attachment graphs typically contains

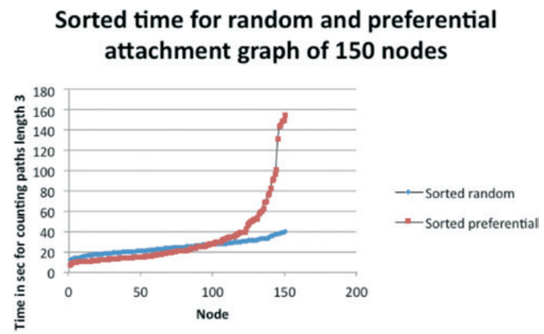


Figure 8. *Sorted computation time for graphs with different degree distributions.*

a few 'hubs' (i.e., nodes with a very high degree), while most nodes have small degrees. This distribution is more skewed than a normal distribution. (In Section 6 it is explained how preferential attachment graphs are dynamically generated.)

It is observed that the computation time is indeed influenced by the degree of a node. For the random graph the computation time for each node does not differ as significantly since the degrees are more evenly distributed. In the statistical program R, the empirical cumulative density functions and estimates of the probability density functions (by kernel density estimation) can be plotted. These plots suggest that in the random graph the computation time is normally distributed. The computation time for the preferential attachment graph is very large for only a few nodes (the hubs), but small for most nodes.

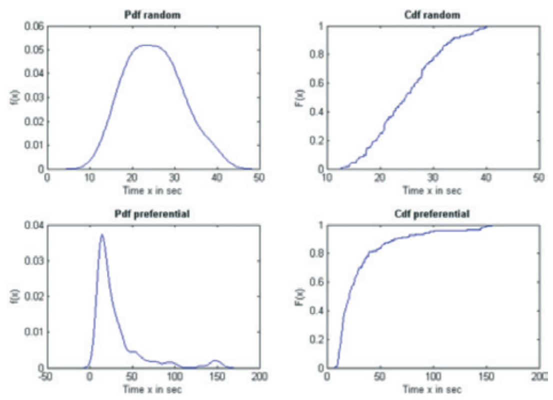


Figure 9. *Estimated pdf and empirical cdf of time for random and preferential attachment graph.*

6. RESULTS FOR SIOC DATA

Figure 10 shows the logarithm of the mean number of paths containing a given node in the forum graph of 1998, as a function of path length. Since this relationship is linear, the mean of the counts grows exponentially in path length. The standard deviation is close to the mean for all path lengths. This shows that the distributions per length have large tails. Similar results were obtained for 1999 and 2000.

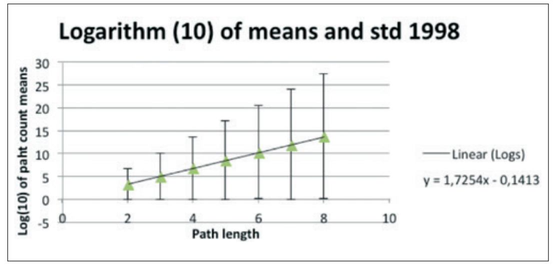


Figure 10. *Analysis of 1998 path counts. Logarithm of means and standard deviations.*

Figure 11 shows the relationship between the number of paths of length 3 vs. length 2 of the same node for the 1998 graph. The relationship is linear and this was found for each pair of lengths up until length 8. Figure 12 shows length 7 vs. length 4 paths.

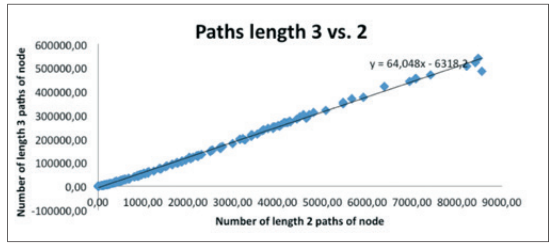


Figure 11. *Paths of length 3 versus 2 for 1998.*

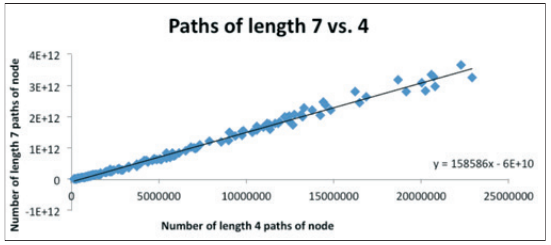


Figure 12. *Paths of length 7 versus 4 for 1998.*

The distribution of paths in the boards forum data is compared against random, preferential attachment and small-world graphs. For the graphs of 1998, 1999 and 2000 the distribution of the number of paths of a specific length can be plotted. In this section, the results for the year 1998 are presented. The plots for the years 1999 and 2000 appeared similar and did not provide additional insights.

In Figures 13-18 the *x*-axis represents the number of paths of specific length that were counted of which a given node is a member. The *y*-axis shows the relative number of occurrences of this number of paths for all nodes in the graph in Figures 13 and 14. In Figures 15-18, the *y*-axis shows the relative occurrence of this number of paths for 150 graphs generated by a graph model in NetworkX. For each of these 150 graphs one random node was picked for which the number of paths was counted. (These graphs have the same number of nodes and edges as the 1998 data.) Based on the minimum and maximum number of paths, the data is segmented into 50 bins. The data was fitted to all valid parametric probability distributions in Matlab by using maximum likelihood estimators in Matlab's function Allfitdist (retrieved from http://www.mathworks.nl/matlabcentral/fileexchange/). The four best fitting probability density functions are displayed in the plot. Since the relationship between different path lengths is linear (Figure 11, 12), the main analysis below is performed with path length 3.

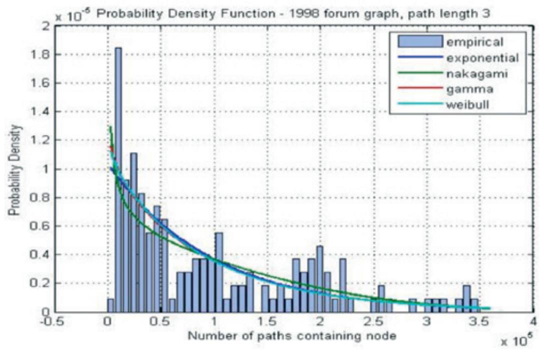


Figure 13. *Histogram of length 3 paths of 1998 and fitted probability density functions.*

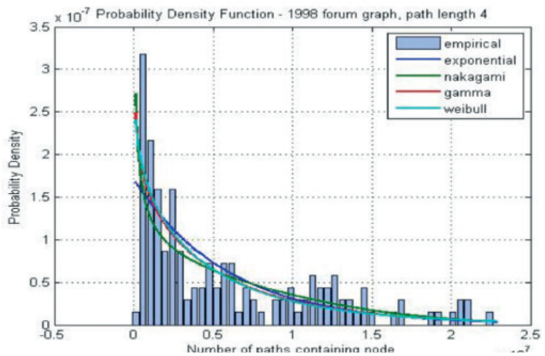


Figure 14. *Histogram of length 4 paths of 1998 and fitted probability density functions.*

Figure 15 was produced by choosing random graphs out of all possible graphs with an equal number of nodes as the 1998 data and counting the number of paths of length 3 that contain randomly chosen nodes.

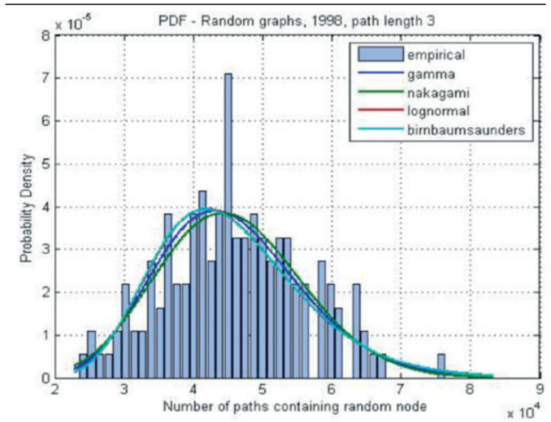


Figure 15. *Histogram of length 3 paths of 150 random graphs, each with 150 connected nodes and 2132 edges (i.e., same number of nodes and edges as 1998 forum graph), and fitted probability density functions.*

NetworkX can also generate scale-free networks with an equal number of nodes as the 1998 data (Figure 16), according to the preferential attachment model by Barabási and Albert (2009). These graphs are dynamically formed by an algorithm that makes new nodes more likely to connect to existing nodes that have more links. Each new node is linked to existing nodes with a fixed number of edges. The probability that a new node connects to existing node *i* is $p_i = k_i / \sum_i$ k_i , where k_i is the degree of node *i* and $\sum_i k_i$ is the sum of the degrees of all nodes in the graph. This algorithm makes it more likely for new nodes to connect to nodes that are more connected in the graph.

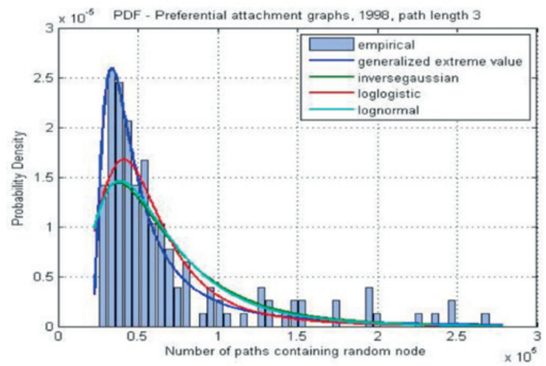


Figure 16. *Histogram of length 3 paths of 150 preferential attachment graphs, each with 150 connected nodes. The number of edges that attach from each new node to existing nodes is 16 (in order to end up with 16(150-16) = 2144 edges (first, there need to be 16 existing nodes before adding 16 from each new node), which is closest to 2132).*

NetworkX has an built-in function to generate small-world graphs according to the model of Watts and Strogatz (1998). These graphs are generated starting from a ring lattice, in which all nodes are in one circle. Depending on the rewiring probability *p*, each edge (*u*,*v*) is rewired with probability *p* to (*u*,*w*), where *w* is a random node in the graph. In Figures 17 and 18 the path counts are shown for random nodes in 150 small-world graphs with rewiring probability 0.03 and 0.5. Other probabilities were analyzed too, but showed a similar result.

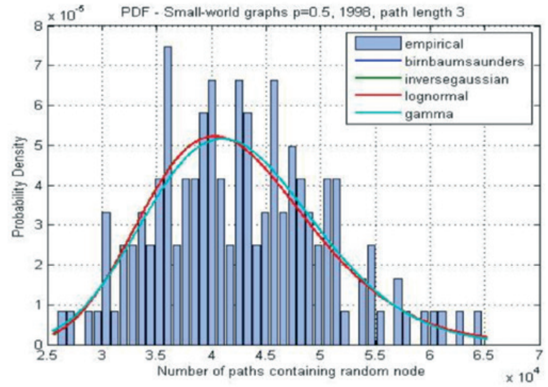


Figure 17. *Histogram of path counts of 150 small-world graphs with same number of nodes and edges as 1998 SIOC data. The rewiring probability is 0.5.*

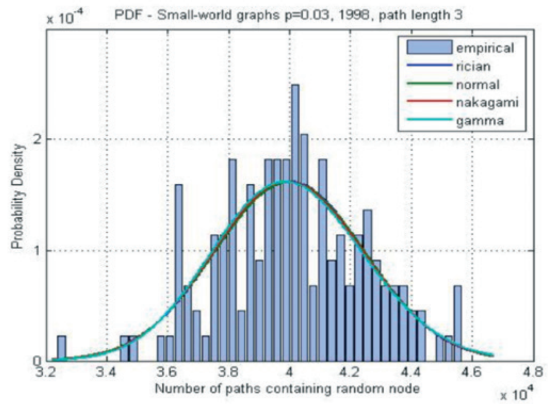


Figure 18. *Histogram of path counts of 150 small-world graphs with same number of nodes and edges as 1998 SIOC data. The rewiring probability is 0.03.*

The number of paths of which a node is a member can be plotted against degree. In Figure 19, the number of paths of length 3 for each node in the 1998 forum graph is plotted against the degree of the node.

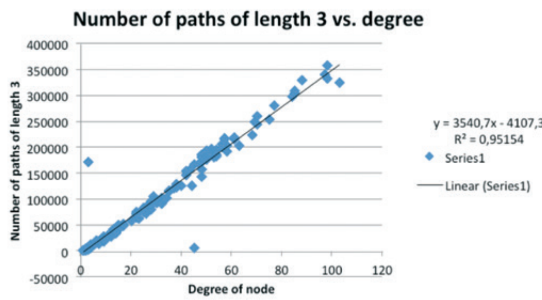


Figure 19. *Number of paths of length 3 against degree of a node in 1998 forum graph.*

7. THEORETICAL RESULT – A NOVEL ALGORITHM FOR BETWEENNESS CENTRALITY

In this section a novel algorithm for betweenness centrality using color coding will be explained. In the final stages of this project, the question arose whether important graph measures could also be determined using color coding. The novel algorithm is a color coding algorithm for the measure "betweenness centrality" and constitutes the main theoretical contribution of this capstone project. The algorithm has been implemented in Python but has not yet been tested on real data.

7.1 DEFINITION OF BETWEENNESS CENTRALITY

Various notions of *centrality* have been proposed by sociologists in order to determine the relative importance of individuals in a social group. On the level of graph representations of social networks, centrality calculations use distances between nodes and are a measure of how information propagates. These measures usually only consider shortest paths and not all paths in a graph. Examples are closeness, graph, stress and betweenness centrality (as cataloged by Brandes 2001).

The term betweenness *centrality* was coined by Linton Freeman in 1977 (Freeman 1977). Betweenness centrality of a vertex *v* in graph *G* is the sum over all pairs of vertices (*s,t*) of the fraction of shortest paths between *s* and *t* that pass through *v*. In formula: $C_B(v)=\sum_{s\neq v\neq t\in V}\frac{\sigma_{st}(v)}{\sigma_{st}}$, where σ_{st} stands for the number of shortest paths from vertex *s* to *t* and $\sigma_{st}(v)$ is the number of shortest paths from *s* to *t* on which *v* lies (Brandes 2001). (For an undirected graph $\sigma_{st} = \sigma_{ts}$.) For instance, in a telecommunications network, betweenness centrality computes the accumulated total number of messages that pass through node *v* when every pair of nodes sends and receives a single message along the shortest path connecting the pair. Hence, it measures the total load that passes through a node.

As a side note, the measure betweenness centrality can be extended to a group level: *group betweenness*. Group betweenness evaluates the importance of a group of nodes in a network and computes the proportion of shortest paths between nodes outside the group that go through the group. Puzis et al. (2009) have shown that a group of Internet users with high group betweenness can highly affect the anonymity of other users.

A well-known algorithm to calculate betweenness centrality is the Floyd-Warshall algorithm, using adjacency matrices. It compares all possible paths between all pairs of nodes and runs in time $O(V^3)$, where *V* is the number of the vertices. This computation time is problematic for large networks. In 2001, Brandes used the Single Source Shortest Path problem (SSSP) results obtained by Fredman and Tarjan (2001). The SSSP problem aims to efficiently determine all shortest paths from one fixed vertex to all others. This can be done by a BFS algorithm in unweighted graphs and Dijkstra's algorithm in weighted graphs. Next, Brandes accumulates pair dependencies of each vertex $s \in V$ recursively – i.e. the proportion of shortest paths from *s* on which *v* is lying. Brandes' algorithm runs in $O(VE)$ on unweighted networks and is currently the fastest known algorithm to compute the betweenness centrality of a given vertex. It requires $O(V+E)$ computer memory space, where *V* is the number of the vertices and *E* is the number of edges. (For weighted networks the time complexity becomes $O(VE+V^2\log(V))$.)

7.2 BETWEENNESS CENTRALITY BY COLOR CODING

Betweenness centrality can also be computed by color coding. Particularly, if the diameter of the graph is smaller than log(*V*) (base 2) the color coding method can be used effectively (with less than *V* colorings) (See Appendix A). In the algorithm below the diameter is given by *d*. This *d* is a constant that is often small, particularly for small-world networks and scale-free networks. (For instance, in the 1998 forum graph, the diameter has value 4 and log(*V*) is 7.23, where *V* is the number of nodes in the connected component. Since the 1999 and 2000 graph have the same diameter but more nodes, it also holds that d < log(*V*).) The color coding method used below is similar to the technique described in Appendix A.

The algorithm aims to discover all paths in the graph of a certain length by color coding, starting from length 1 up until the diameter of the graph. Once a colorful path between two nodes is found, it is certain that their shortest distance is not larger than the current length.

In order to count colorful paths for each length, an extra node *z* is added to the graph. This node is linked to all other nodes. The graph is colored sufficiently many times as to observe many paths of length *k* with *z* as end node. In every coloring the new node is given a different color from all other nodes in the graph. For each path length *k* the graph is colored ek times. As explained in Appendix A, this ensures that the expectation of finding all paths of length *k* is larger than 1. In fact, coloring Ce^k times makes sure that the probability that a given path goes undetected is smaller than e^{-C}. (This random element can also be removed by derandomization techniques, which adds an extra log(*V*) factor to the complexity (See Appendix A).)

On the next page, the algorithm to compute the betweenness centrality (*BC*) of vertex v is described in pseudo-code. *BC* is given as a Python dictionary with tuples as keys and lists as values. An entry in this dictionary is given by (*w,u,k*):[*i,j*], where *w* and *u* are end-vertices with shortest path of length *k*. There are *i* occurrences of shortest paths between w and u on the graph, and *j* of these *i* paths contain vertex *v*. For instance, if you have a dictionary {(1,2):[5,3], ...}, the first entry (1,2):[5,3] contributes 3/5 to the betweenness centrality calculation as it indicates that there are 5 shortest paths between vertex 1 and 2, and 3 of these paths contain the vertex *v*. will be incremented with 3/5. During the computation the following will be saved:

- All pairs of nodes in the graph are saved together with the length of their shortest path in the *BC* dictionary. This uses space $O(V^2)$ maximally.
- For each path length, all paths found are saved in memory and require maximum space $O((eV)^k)$.

The total computation time of the algorithm is $O((2e)^k E)$, expected time. $C_i(z,[k+1])$ is of $O(2^k|E|)$ (Section 3). The e^k factor arises from repeating the colorings. The *Return_paths* function referred to in step 2.c.*ix* is described in Appendix C. It returns the colorful paths found during the computation of $C_i(z,[k+1])$ in a list.

Algorithm: compute betweenness centrality for a vertex *v* by color coding

1. *BC* = {}
2. *For length* = 1 to *d*
   - a. *Paths* = []
   - b. *k = length* + 1
   - c. *For i* = 1 *to* e^k
     - *i.* *Add vertex z and add edge* (z,v_i) *for all* $v_i \in V$
     - *ii* *Color each vertex of G independently and uniformly at random with one of the k colors.*
     - *iii* *Color z with color k* + 1
     - *iv* *listpaths* = []
     - *v* *For all* $u \in V$ $C_i(u,\emptyset)=1$
     - *vi* *For all* $l \in [k]$ $C_i(v,\{l\})$ = (*if col*(*v*)=*l*:*return* 1 *and listpaths.insert*(*0*,[*v*]) *else: return* 0
     - *vii* *For all* $S \subseteq [k]$ *s.t.* $|S| > 1$, $C_i(v,S)=\sum_{u\in N(v)} C_i(u,S \setminus \{col(v)\})$ and if Đ $C_i(v,S) > 0$, *listpaths.insert*(0,*v*)
     - *viii* *Compute* $C_i(z,[k+1])$
     - *ix* *Paths_i = Return_paths*(*listpaths*)
     - x *For j* in Paths_*i*
       1. *Disregard j if j in Paths*.
       2. *Disregard j if for w = j*[1] (*second node*) *and u = j*[-1] (*end node*) *if* (*w, u, k_small*) *in BC for k_small* < *k*
       3. *Else*:
          - *a. Append j to Paths*
          - *b. BC*[(*w,u,k*)] += [1,0]
          - *c. If v≠w≠u in j*, *BC*[(*w,u,k*)] + = [0,1]
3. *Compute centrality*: $\sum_{(w,u,k)\in BC} \frac{\textit{second value in dictionary}}{\textit{first value in dictionary}}$

8. DISCUSSION OF RESULTS AND CONCLUSION

8.1 DISCUSSION OF SIOC DATA RESULTS

The graphs analyzed in this discussion have been created as explained in Section 4; nodes represent users with accounts on the website "boards.ie," and edges the connections between users if they posted in the same threads.

The histograms of the number of paths in the 1998 forum data (Figure 13, 15) show that the distribution of the number of paths is broader than for random graphs with the same number of nodes and edges. The peak for path length 3 of the 1998 data is at $1.5\cdot10^4$ paths (Figure 13), while for random graphs with the same number of nodes as the 1998 graph the peak is at $4.5\cdot10^4$ (Figure 15) (≈3 times larger for forum graph). On the other hand, the tails of path distributions of the forum graph are much larger for all path lengths. Again, for path length 3 the largest number of paths of the forum data is around $3.5\cdot10^5$, while for random graphs it is around $2.3\cdot10^4$ (≈15 times larger for forum graph). Similar results were obtained for larger path lengths. This suggests that there are a few members extremely active on the boards website and are connected with other nodes through many paths.

For instance, when a node is a member of a large number of paths of length 3 (4 nodes), it suggests that the node itself posts prolifically and has many threads in common with other users and/or it could have posted in threads in which other extremely active users have posted (i.e., it has extremely active neighbors). However, from Figure 19 (number of length 3 paths versus degree) it can be concluded that the number of paths a given node is a member of and the degree of the node are highly correlated. This suggests that prolific users are generally in more paths. The two outliers in this plot could be explained by considering their neighbors. There is one node with low degree (degree 3) and a high number of paths (173000). This suggests that it is connected to one or more very prolific users. The other outlier is a node with a higher degree (degree 45) and a relatively low number of paths (8022), suggesting that its neighbors are generally not prolific users.

The distribution of the number of paths in the forum graphs is similar to that of preferential attachment graphs (Figure 13, 16), while the small-world model by Watts and Strogatz with varying rewiring probabilities exhibits a different distribution (Figure 13, 17, 18).

It appears plausible that the forum graphs are built up by a similar mechanism to the preferential attachment process. The preferential attachment model (as explained in Section 6) is a graph generation model. In a preferential attachment graph, every new node connects to already existing nodes with a higher probability of connecting to "popular" nodes in the graph (Barabási 2009). This mechanism could explain how a forum network is dynamically generated. Each new individual on the boards website posting in a thread has a higher probability of posting in the same thread as a person who posts prolifically on the forum website. Hence, the individual has a higher probability of forming an edge with a prolific poster.

However, the preferential attachment model differs from the real-world boards graphs in several aspects. In a preferential attachment graph, a new node has a fixed number of starting edges. Each new node connects to existing nodes with the same amount of edges, while a new individual on a forum could actively engage in many threads, be an observer and/or create new threads. In addition, the degree distribution of a scale-free network that is produced by the preferential attachment model generally follows a power-law distribution. This was not observed for the boards data (Figure 7). In the boards data, the number of nodes with higher degree decreases less rapidly.

8.2 DISCUSSION OF BETWEENNESS CENTRALITY BY COLOR CODING

There exist numerous measures evaluating the relative importance of a node in a network. These measures generally give an indication of the amount of information that passes through a node. One such measure I have devised is as follows: count the number of paths that pass through a node of all lengths up to the diameter of the graph (with a minimum length of 2). Intuitively, this could give an indication of how well a given node is connected in a graph. However, as observed in Figure 19, the degree and number of paths of a node are highly correlated for the forum data. Hence, counting the number of paths containing a node up until the diameter is expected to be correlated with degree centrality. It might be informative to research to what extent these measures are correlated.

In research literature on determining the importance of nodes, there are a few common centrality measures that only consider the shortest paths in a graph, such as closeness and betweenness centrality. The color coding algorithm for betweenness centrality described in Section 7 is faster than any currently known algorithm. More research is needed in order to evaluate its practical value and compare its performance to Brandes' algorithm. In this project the algorithm was programmed in Python, but not yet tested on real data. In addition, other practical improvements can be suggested. The function Return_paths could be removed from the algorithm. By doing more iterations, the proportion of shortest paths on which v is lying could be estimated. In this way it is not necessary to check whether the exact same path has been encountered in a previous coloring for the same path length. This reduces memory usage. Moreover, it is valuable to research whether the algorithm can be effectively parallelized in order to further reduce the computation time.

8.3 FINAL CONCLUSION

In this project it has been shown that color coding is an effective method to approximate the number of paths containing a vertex, with high accuracy and tractable computation time. The differences between the path distributions of the forum data and artificially generated graphs gave insights into the dynamics of a forum network and how users behave in the network. To further explore color coding, the simple path algorithm can be adapted to count other motifs such as cycles, cycles with chords and cliques. Using color coding to compute betweenness centrality also needs further exploration. The theoretical algorithm described in this thesis has the potential to outperform current methods.

REFERENCES

Alon N, Yuster R, Zwick U. 1995. Color-coding. Journal of the ACM. 42(4):844-856.

Alon N, Dao P, Hajirasouliha I, Hormozdiari F, Cenk Sahinalp S. 2008. Biomolecular network motif counting and discovery by color coding. Bioinformatics. 24(13):241-249.

Alon N, Gutner S. 2009. Balanced hashing, color coding and approximate counting. Electronic Colloquium on Computational Complexity, Report No. 12. Retrieved 1 April 2014 from http://eccc.hpi-web.de/report/2009/012/.

Alon N, Spencer JH. 2004. The probabilistic method, 2nd edition. New York, NY: J. Wiley and Sons.

Arvind V, Raman V. 2002. Approximation algorithms for some parameterized counting problems. In: Bose P, Morin P, editors. ISAAC 2002. LNCS 2518, pp. 453-464. Springer Verlag Berlin, Heidelberg 2002.

Barabási, 2009. Scale-free networks: a decade and beyond. Science. 325(5939):412-413.

Becchetti L, Boldi P, Castillo C, Gionis A. 2008. Efficient semi-streaming algorithms for local triangle counting in massive graphs. Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 16-24.

Brandes U. 2001. A faster algorithm for betweenness centrality. Journal of Mathematical Sociology. 25(2):163-177. doi:10.1080/0022250X.2001.9990249

Bodlaender HL, Koster AMCA. 2008. Combinatorial Optimization on Graphs of Bounded Treewidth. The Computer Journal. 51(3):255-269 doi:10.1093/comjnl/bxm037

Cook S. 1971. The Complexity of Theorem-Proving Procedures. Proceedings of the 3rd ACM Symposium on Theory of Computing, pp. 151-158. doi:10.1145/800157.8050047

Freeman L. 1977. A set of measures of centrality based on betweenness. Sociometry. 40:35-41. doi:10.2307/3033543

Gonen M, Shavitt Y. 2009. Approximating the Number of Network Motifs. In: Avrachenkov K, Donato D, Litvak N, editors. WAW 2009, LNCS 5427, pp. 13-24. Springer-Verlag Berlin Heidelberg 2009. doi:10.1.1.146.5217

Hüffner F, Wernicke S, Zichner T. 2007. Algorithm engineering for color-coding with applications to signaling pathway detection. Algorithmica. 52:114-132.

Johnson DS. 2005. The NP-Completeness Column. ACM Transactions on Algorithms. 1(1):160-176. doi:10.1145/1077464.1077476

Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002. Network motifs: Simple building blocks of complex networks. Science. 298(5594):824-827.

Przulj N, Corneil DG, Jurisica I. 2005. Modeling interactome: Scale-free or geometric? Bioinformatics. 150:216-231.

Puzis R, Yagil D, Elovici Y, Braha D. 2009. Collaborative attack on Internet users' anonymity. Internet Research. 19(1):60-77. doi:10.1108/10662240910927821

Stockmeyer L. 1983. The complexity of approximate counting. STOC '83, Proceedings of the 15th annual ACM symposium on Theory of Computing. 118-126.

Tsourakakis CE, Kang U, Miller GL, Faloutsos C. 2009. Doulion: Counting Triangles in Massive Graphs with a Coin. KDD '09, Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 837-846.

Ullmann JR. 1976. An Algorithm for Subgraph Isomorphism. Journal of the Association for Computing Machinery. 22(1):31-42. doi:10.1145/321921.321925

Watts D, Strogatz SH. 1998. Collective dynamics of 'small-world' networks. Nature. 393:440-442. doi:10.1038/30918

Wu G, Harrigan M, Cunningham P. 2011. Characterizing Wikipedia pages using edit network profiles. SMUC '11, Proceedings of the 3rd International Workshop on Search and Mining User-generated Contents. 45-52. doi:10.1145/2065023.2065036

Zhao Z, Khan M, Kumar VSA, Marathe MV. 2010. Subgraph enumeration in large social contact networks using parallel color coding and streaming. ICPP '10, Proceedings of the 2010 39th International Conference on Parallel Processing. 594-603.

APPENDIX A: COLOR CODING TECHNIQUE BY ALON ET AL. (1995)

This appendix outlines the technique described by Alon et al. (1995) to find cycles of length *k*:

Finding cycle of length k in graph $G = (V,E)$. By assigning random colors (k distinct colors), each cycle has probability of $p = k! / k^k > 1/e^k$ to become 'colorful'. 'Colorful' means that all the nodes in the cycle have a distinct color. (There are k^k ways to color k vertices on the cycle, of which $k!$ are colorful.)

In order to have a high probability for a colorful cycle, it is usually sufficient to repeat the experiment Ce^k times. In each experiment the probability that there is no cycle is $1 - p = 1 - k! / k^k < 1 - 1/e^k$. Then the probability not to find a cycle when there actually is one is

$P(\text{not cycle}) < (1 - \frac{1}{e^k})^{Ce^k} = (1 - \frac{C}{Ce^k})^{Ce^k} \approx e^{-c}$, since $(1 + \frac{x}{n})^n$ equals e^x in the limit. Thus, the larger C the smaller is the probability that a cycle is not detected.

Moreover, it is expected that the cycle will become colorful at least once in $k^k / k! < e^k$ experiments (Alon and Gutner 2009). This is the case since each coloring X is a Bernoulli experiment with parameter p. Repeating X r times gives $T \sim Binomial(r,p)$. Since $E(Y) = rp$, this expectation is > 1 if $r = e^k$.

Therefore, a motif of size $O(\log(n))$ is expected to be found in $O(n)$ colorings, where n is the number of nodes in a given graph.

Algorithm (Alon, Yuster and Zwick 1995):

Find all pairs of vertices in V connected by simple path of length k-1 and check whether the two vertices in each pair are connected.

Define a coloring function c: $V \rightarrow \{1,\dots,k\}$ which randomly assigns colors to every vertex. Enumerate all partitions of the color set {1,...,k} into two subsets C_1, C_2 of size $k/2$ each (for simplification floor and ceiling omitted). (So for $\{1,2,3,4\}$ the partitions would be $[\{1,2\},\{3,4\}]$, $[\{3,4\},\{1,2\}]$, $[\{1,3\},\{2,4\}]$, $[\{2,4\},\{1,3\}]$, $[\{1,4\},\{2,3\}]$, $[\{2,3\},\{1,4\}]$). There are only $\binom{k}{k/2} < 2^k$ such partitions. Divide V into V_1 and V_2, and let G_1 and G_2 be the subgraphs induced by V_1 and V_2. Then recursion is used to find colorful paths of length $k/2$-1 in each of G_1,G_2.

Say A_1 and A_2 are the matrices for which an entry A_{ij} is 1 if vertex i is connected to vertex j by a colorful path of length $k/2$-1. The matrix multiplication $A_1 B_{A2}$, where B represents the adjacency relations between $V1$ and V_2, gives all pairs of vertices in V that are connected by a colorful path of length k-1. Here the first $k/2$ vertices have colors in C_1 and the last $k/2$ vertices have colors in C_2. Hence, the time it takes to compute the number of colorful paths is $t(k) \leq 2^k \cdot t(k/2)$, since the algorithm recursively partitions into two equal sized subsets. Because the matrix multiplication satisfies the recurrence, the total algorithm is of order $2^{O(k)} \cdot V^\omega \in O(V^\omega)$, where ω is the exponent of the $O(n\omega)$ matrix multiplication algorithm used. (The lowest exponent currently known is 2.3727 (by Vassilevska Williams)).

Hence, the expected overall computation time to find a simple cycle of length k is $e^k \cdot O(V^\omega)$.

Additional comments

A useful characteristic of this algorithm is that it finds simple paths (with no repeated vertices). However, the algorithm only finds the end-points of the colorful path. Alon and Noar have described an extended approach that uses witnessed matrix multiplication.

Alon et al. (1995) also include derandomization results. These derandomization results put a bound on how many steps need to be taken to remove randomness. Such rigorous checking would result in being absolutely certain whether a motif is in a graph or not. Alon et al. (1995) describe how randomness can be eliminated by requiring that for every subset $V'ĐV$ with size k there exists a coloring that gives each vertex in V' a distinct color. This elimination is similar to the existence of a k-perfect family of hash functions from {1,2,...,|V|} to {1,2,...,k}. This family has size $2^{O(k)} \cdot \log(V)$, but the authors remark that the derandomization process can easily be parallelized whence the complexity is brought back to polylogarithmic time (Alon and Gunter 2009). For more information on perfect hash functions and derandomization please refer to Alon and Gunter (2009).

APPENDIX B: BINOMIAL DISTRIBUTION IS MONOTONE DECREASING - PROOF

The binomial distribution is monotone decreasing.
Consider binomial distribution T with parameters t and p.

Then it holds that $\frac{P(T=x)}{P(T=x-1)} < 1$, then $x > (t + 1)p$.

Proof:

Helpful information: $\frac{\binom{t}{x}}{\binom{t}{x-1}} = \frac{\frac{t!}{x!(t-x)!}}{\frac{t!}{(x-1)!(t-x+1)!}} = \frac{t-x+1}{x}$

Monotone decreasing if:

$$\frac{P(T=x)}{P(T=x-1)} = \frac{\binom{t}{x} p^x (1-p)^{t-x}}{\binom{t}{x-1} p^{x-1} (1-p)^{t-x+1}} = \frac{\binom{t}{x} p^x (1-p)^{t-x}}{\binom{t}{x-1} p^{x-1} (1-p)^{t-x+1}} = \frac{t-x+1}{x} \cdot \frac{p}{1-p} < 1$$

Hence, $(t-x+1) < x\frac{1-p}{p}$,

$t < x-1+x\frac{1-p}{p}$,

$t < -1+x(1+\frac{1-p}{p})$,

$t < -1+\frac{x}{p}$,

$p(t+1) < x$,

Therefore, $x > (t + 1)p$

APPENDIX C: RETURNING COLORFUL PATHS DURING PATH COUNTING

This appendix contains the Return_paths function called in the Betweenness Centrality algorithm in Section 7.

During the computation of $C_i(z,[k+1])$ the vertices are saved recursively. The list *listpaths* contains all colorful paths for one coloring. However, since $C_i(z,[k+1])$ is a depth first search in *k*, the vertices in *listpaths* do not represent the paths yet. Although the ordering is correct, the paths need to be extracted from the list. This is what the function *Return_paths* in step 2.*c.ix* takes care of.

Algorithm for returning paths: *Return_paths(listpaths* = $[v_1, v_2, [v_3] \dots , [v_n]]$)

1. Paths = []
2. First end_vertex is of type list, so $[v_3]$ in this case.
3. The first path is the list up and including end_vertex, so insert $[v_1, v_2, [v_3]]$ in Paths.
4. Delete this path from *listpaths*.
5. While *listpaths* ≠ []
    a. *Part _path* is the list up and including a new end_vertex of type list.
    b. *k – length*(*Part _path*) are the vertices from the previous path which this path has in common.
    c. Insert the *k – length*(*Part_second_path*) vertices of the previous path at the front of *Part_path* to obtain the next path.
    d. Append path to Paths.
6. Return Paths

Example: If *list_paths* is [2, 3, 4, [5], 9, 6, [7], 10, [11], 12, [13]], the paths that will be returned are [[2,3,4,5], [2,9,6,7], [2,9,10,11], [2,9,12,13]

Framing Public Art: Situating the Kunst Hek Project in Community-led Urban Regeneration

Rosa ter Kuile

Supervisor: Dr. A. Brown (AUC)

ABSTRACT

Cultural policy makers increasingly frame public art in the discourse of urban regeneration, resulting in a plethora of community-led arts schemes. One such initiative is the Kunst Hek Project, a low-key arts initiative in the East of Amsterdam, the Netherlands. This paper analyses the three contributors to the discussion around community-led public art projects, involving (but not limited to) policy makers, critical literature, and the participants themselves. In staging a dialogue between these key players, the paper proposes that "democracy", "participation" and "temporality" are terms that act as points of axis into the case study. Ultimately, the paper questions how far the critical literature can understand the Kunst Hek Project, and proposes new ways in which the Kunst Hek Project can operate in the public domain.

INTRODUCTION

In 2009, the Dapperbuurt neighbourhood of Amsterdam, the Netherlands was allocated 150,000 euros to spend on community-led cultural projects (Weber). This was not merely a generous state gift to the art enthusiasts of the neighbourhood, but had wider political connotations. In fact, the money aimed to generate active citizenship that would shake the neighbourhood from the pitfalls of its various infrastructural and social shortcomings concerning issues around housing, employment, education, integration and security (Permentier et al. 111). Consequently, a wide array of events took place that year; Dapperbuurt locals could join the "meet-your-neighbour dinner", participate in Sunday morning football matches held in the vacated market space, or even strut down the "catwalk" with their domestic animal in tow (Weber). Also among the festivities was a new space for locals to exhibit their artwork, named the Kunst Hek Project. This "Fence Art Project," which continues to line the Oosterpark fence well into 2014, is one of the few remaining markers of the attempt to regenerate the neighbourhood through community-led cultural events.

This story of the Dapperbuurt regeneration is situated in a wider multi-million euro government scheme to reconceptualise the forty faltering *probleemwijken* (problem neighbourhoods) in the Netherlands as *krachtwijken* (empowered neighbourhoods) (Permentier et al. 111). Headed by the Minister for Housing and Integration, Ella Vogelaar, the national scheme invited residents to be active in the decision-making and treated them as "field experts" (ibid.). The emphasis put on public empowerment in the Netherlands, is part of a long tradition in 'area-based policy' originating in the 1980s (ibid.).

But this emphasis on culture-led urban regeneration is not a phenomenon unique to the Netherlands. Rather, investing in community art projects is a common strategy employed by local governments to impact the renaissance of struggling neighbourhoods (Hall & Robertson 2001, Kay 2000). These strategies appeal to local governments on multiple levels, as community-led cultural schemes are cost-effective ways to build a sense of place, increase community interaction and create public empowerment (Hall & Robertson 2001, Kay 2000, Myerscough 1988, Petherbridge 1987, Cameron & Coaffee 2005). Yet the increasing policy emphasis on the social benefits of public art initiatives has provoked criticism from academics in the last decade (Hall & Robertson 2001, Sharp et al 2005, Pollock & Sharp 2012). These studies question the essentialist claims of cultural policy since the 1980s, and propose that the move towards current participation-led policy is not to give more political agency to the public, but is rather indicative of more state control.

This paper is concerned with the three primary groups participating in the discussion around community-led public art projects. The first is academic discourse on public art that sprouted in the 1980s, which critically examines the vocabulary imbued in public art. Secondly, the discussion involves policy makers, who in the last years have invested in public art initiatives and have taken extensive measures to relay responsibility to the their creative constituents. In this case in particular, the paper involves the Dutch context, which is little covered elsewhere in the literature. And thirdly, I turn to the public, to those who carry out the task of organising and participating in public art projects. Staging a dialogue between these three actors, this this paper draws out three key terms: "democracy", "participation" and "temporality", providing a framework to analyse the site of the Kunst Hek Project. Each of the three conversations between the participating groups contributes to the formation of the discourse on public art, and it is the objective of this paper to discern how these are related to, part of, or distinct from each other. I will also discern to what extent the critical literature is useful in furthering our understanding of the site. Ultimately, I will propose new ways in which the Kunst Hek Project can operate in the public domain.

This paper triangulates an analysis of critical literature, cultural policy and artistic practice in the Kunst Hek. To do so, I combine literature review with ethnographic research at the site of the Kunst Hek Project. This was in keeping with the underlying qualitative approach that allows for a "level of creativity and expressiveness that is largely absent in the more conventional research methods" (Prasad 291). In incorporating observation and interviews, this paper analyses the Kunst Hek Project through vocabulary that is specific to the site, creating a rich framework that embraces the fine arts, the humanities, and the social sciences.

The ethnographic component of the project included site observation conducted between November 2013 and May 2014, as well as four in depth interviews concerning the location, mission and organisation of the Kunst Hek Project. Concerning the Kunst Hek Project directly, one interview was conducted with the primary project manager, Anke Weber, and the other with Teo Krijgsman, a participant and organiser, conducted via email. The other two interviews took place in the meeting area of the municipality (Stadsdeel Oost) with Gijs Hendrix, a cultural policy adviser for the Dapperbuurt. And the last was held with the "participation initiator" of the Dapperbuurt, Stephany van Veen. The interviews were loosely structured in order to enable the respondents to fully express their views and opinions.

Each phase of the research was posted to the blog (organisingpublicspace.tumblr.com) to facilitate a conversation on the use of public space in the Dapperbuurt. The blog continues to act as a direct and accessible way in which the wider public can see the results of the research. By integrating both the critical literature and ethnographic investigations, the resulting analysis of the Kunst Hek Project is constructed through terms that are specific to the site.

Figure 1. *Blog Homepage: organisingpublicspace.tumblr.com*

THE KUNST HEK PROJECT: CONTEXT, PLACE AND PURPOSE

The Kunst Hek Project is located on the Linneausstraat in the Dapperbuurt, a multicultural area and traditionally low-income neighbourhood (Weber). The project is comprised of seven large frames placed in the middle of the fence that runs parallel to the Oosterpark in the east of Amsterdam (Figure 2). These frames alternate artworks on a monthly basis, of which the content is either produced by Dapperbuurt residents or is reflective of the neighbourhood. The small exhibitions have varied from thin papier-mâché sculptures to black and white holiday snapshots. At its heart, the Kunst Hek Project is geared to initiating participation in the local community. Not only is the production of artwork "local" activity, but the decision-making around the space and maintenance of the exhibition space is all relayed to a group comprised of local volunteers (Weber).

The Oosterpark is the main green public space in the neighbourhood, and is used for recreational or travelling purposes. Early in the morning there are joggers and cyclists, followed later in the day by young mothers with prams and grocery shopping. On sunny weekends the entire East of Amsterdam flocks to the park and the smell of barbeque envelopes the surrounding area. The most consistent demographic of the Oosterpark, however, remains the clusters of older men occupying a significant number of the park benches.

In comparison with the centre of Amsterdam, the Linneausstraat is a wide street that hosts cars, plenty of bikes and two tramlines that stop in front of the

Figure 2. *The Dapperbuurt is shaded in blue; the Kunst Hek Project is marked by the red dot. https://maps.google.nl/maps/*

Kunst Hek Project. On the opposite side of the street there are also a string of cafes and restaurants whose terraces stretch far out onto the pavement. At night this side of the Linneausstraat buzzes with the sound of chatter and drinking.

Yet the intimacy that is so present on this side of the road is little reflected on the pavement of the Kunst Hek Project. With regard to this, two observations can be made. Primarily, the pavement and fence beside it go largely unnoticed, as the pavement fence does not attract immediate attention. The fence is not in itself a marker of cultural intention - in a way that a public square leads the eye to a central feature. Secondly, pedestrians tend to flow through the Oosterpark itself in order to reach other points further up the Linneausstraat, re-joining the main street by means of exits beyond where the artworks are on display. This severely limits the number of potential viewers of the pieces. Further observation shows the extent in which the Dutch climate impacts the visibility of the space. When the sun is out, more people are likely to go to the park and pass the fence. In rain, few will stop and take time to look at what is displayed in the frames. Visibility is further hindered in cold weather when the glass in the frames produces condensation, and the artwork becomes blurred. And, just as the visibility of the Kunst Hek Project depends on the climate, my observation concluded that it is also dependant on the time of day. In the daytime the demographic of the

Figure 3. *Facing the Kunst Hek project and Oosterpark*

pedestrian traffic largely consists of older generations, whereas school children are often to be found passing by in the afternoon. At night, as with the rest of Amsterdam, the streets are quiet and the pavement is largely un-trodden.

1. THE CRITICAL DEBATE

This section briefly establishes and reviews the wider academic debate around public art, as the field has come under great scrutiny since the term was coined some forty years ago (Cartiere & Willis 1). The discussion and practice of public art shifted radically in Europe and America in the 1980s. In this decade, public art became justified not solely on its aesthetic function, but on its supposed contribution to urban regeneration (Hall & Robertson 5). This alignment greatly impacted the reception of public art as it became synonymous with "compromise, dilution, and dependency" - public art never gained credibility as a fine art discipline (Cartiere & Willis 1). The discourse around public art now produces a wide vocabulary of elusive terminology, dominated by words such as "public", "space", "temporality" and "democracy".

A MARGINALISED TERM

Public art does not differ from other contemporary art practices; it encompasses every medium possible – painting, sculpture, architecture, performance, graffiti, poetry and new media ("Public Art", Cartiere & Willis 1). Interestingly, although many artists and critics produce work that defines itself in these terms, they consistently refrain from naming their work "public art". Instead they prefer a lexicon of alternatives, often choosing to talk about public art in terms of "interventions", "socially engaged practice", "site-specific works", "community-produced projects" and "social practice art" (Cartiere & Willis 1). City councils are seemingly the only group comfortable with applying the term (ibid.).

However, the annexing of public art in urban regeneration plans is not the sole reason for its marginalisation. Cartiere & Willis (2008) propose the lack of "criticism, education" and "evaluation" around public art works also contributes to its demise in popularity (2). They argue that although there are countless cultural institutions, programs and funds dedicated to the public arts in Europe and the United States, there remains a "disconnection between the work produced" and the critical analysis of public art (Cartiere & Willis 2). This is echoed by the fact that there are only two magazines dedicated to the field: the *Public Art Review* in the United States and *The Public Art Forum* in the United Kingdom (ibid.). Similarly, unless the artwork causes some controversy, public art receives little attention in the press (ibid.). And when it is featured, it is all too often side-lined to "summer features" or "colour commentaries" (ibid.).

Some even see the marginalisation stemming from the association of public art with art of a lower and non-professional standard, such as community-produced art (Hendrix). Gijs Hendrix, a cultural policy adviser in the Dapperbuurt, draws on the example of the "mosaic bench" the council avidly avoids at all costs:

> *We get a lot of applications for…making mosaics on benches together. But…we don't really find [this activity] stimulating. Because it may be nice for the two or three people who do it. But afterwards nobody knows why, or when or how this was done… We would like to see more inventive projects.*

Though the "mosaic bench" (which manifests itself in many forms – mosaic walls, mosaic tiles etc.) may have taken part in tainting the image of public art, there is enough evidence to show that contemporary public art has evolved from this from this stereotype.

IS THERE A "PUBLIC" ART?

The very existence of public art is also problematic in itself, as according to Hilde Hein (1996), public art engages in a contradictory discourse: it is both private and public. On the one hand, art is understood as the product of individual autonomous expression and is therefore a uniquely personal production that does not engage in the public domain. And, since the meaning of art is created through a private act of contemplation, the reception of a work of art is similarly an individualistic experience (Hein 1). Hein (1996) therefore considers that the construct of a *public art* is at its heart an oxymoron, and is only feasible when art entails the "artist's self-negation and deference to a collective community" (1).

If we look back in western European artistic tradition however, it is evident that public art did embody such a collective model. The celebrated treasures of Greece and Rome, in addition to the Christian works of the Middle Ages and the following "age of the Fresco", do not "exalt the private vision of individual artists so much as they bespeak the shared values and convictions of cultural communities" (ibid.). Found in public places and establishments, this public art was able to resonate with the shared values and beliefs of the public (ibid.).

Modernism has effectively reversed this order through its unhesitant "glorification of the individual" (ibid.). It has shaped artistic practice to invest "personhood with uniqueness", placing the social as a "derivative aggregate" (ibid.). Therefore "modernist representation of art", Hein continues, "gives pride of place to that which is irreducibly personal" (ibid.). However, partisans from Kant to the present defend this turn towards glorified personhood (ibid.). They recognize that within this structure, art can have a "liberatory function" that is "conceptually constructed out of a fusion of artistic independence (the unregulated genius) with political autonomy (the absence of heteronomous coercion)" (ibid.).

But how do we define public art if modernism has created a post-public world? Place is very important (Hein 1996, Phillips 1989). However, just as placing a tiger in a backyard does not make it a domestic animal, neither does art become public when simply placed in a public space (Hein 2). The crux of the argument then is that the artwork does not derive its identity from the "character of the place in which it is found" (Hein 4). Public-ness "has social and political connotations that are untranslatable to public access" (Hein 1).

Hein (1996) also draws upon the definition that public art is "art installed by public agencies in public places and at public expense" (Hein 2). But this is too "narrow" and "pragmatic" a definition (Hein 2). She adds that, conventionally, "the term "public art" has referred to a family of conditions including the object's origin, history, location, and social purpose" (Hein 1). But this term does not sufficiently encompass the non-traditional contemporary nature of public art (Hein 2). Additionally, each of these conditions loses meaning in a world affected by the influence of technology, secularization, cultural migration and economic restructuration (Hein 1). But what then is public art?

Thankfully, we can derive two conclusions from the question. First and foremost, contemporary public art occupies public space, a space which no longer attaches to materiality (Hein 4). Secondly, public art is active in the public sphere and should engage in critical discourse and communicative exchange (Hein 5). Public art is "public because of the kinds of questions it chooses to ask or address", Phillips (1989) adds, "and not because of its accessibility or volume of viewers" (332). Michael North eloquently expresses this in his text *The Public as Sculpture*:

> *It is not the public experience of space but rather public debate that becomes a work of art. They make manifest an important truth about public space, that unless it is embedded in a larger public sphere that values debate… then it will always be decorated by mass ornaments, no matter what sort of art is put into it. (qtd. in Hein 4)*

Though this definition may provide some solace, it remains a challenging, obscure and difficult definition of public art (Phillips 332). For those involved in creating public art, "it requires a commitment to experimentation" and to the conviction that "public art and public life are not fixed" (Phillips 332). And because public art is not fixed, the terminology that we use to discuss it "needs to be as flexible as the material itself" (Cartiere & Willis 3).

2. KEY TERMS

In acknowledgement to the flexibility of terms, this section proposes three terms – related (but not limited) to democracy, participation and temporality - through which we can understand the Kunst Hek Project. These terms are first put in relation to the literature, then to the cultural policy in the Netherlands and Dapperbuurt, and lastly to the Kunst Hek Project itself.

DEMOCRACY AND (COMMUNITY) PARTICIPATION

One cannot posit public art in the public sphere, without discussing democracy (Hein 1996). The emergence of "democracy" in the art world, whether in "a nascent state or in more sophisticated efforts to formulate the terms of democratic aesthetic practice", corresponds to an eruption of struggles around the meaning of the term (Deutsche 35). Democracy has become a political catchword that, Deutsche (1992) argues, is "largely articulated in a conservative direction" (36). Increasingly, the term "democracy" aligns itself with anticrime campaigns and "crusades" for public decency that routinely provide the "democratic justification for the imperatives of surveillance and exclusion in public space" (Deutsche 37). To illustrate this point, Deutsche (1992) draws on the example of Jackson Park, New York, where in the 1990s the local residents decided to lock "their" park at night to evict the homeless. The public space, according to the locals, was being threatened by their presence (Deutsche 38). Such actions are indicative of the larger privatisation of space, which denies the underprivileged, in this case the homeless, "the right to have rights"(Deutsche 38).

In urban regeneration plans, the aesthetic shaping of the city landscape is loaded with questions of democracy (Deutsche 34). When an authoritative body, such as the city council, draws criteria for placing art in public space, they "routinely employ a vocabulary that invokes, albeit loosely, the tenets of both direct and representative democracy" (34). The type of questions they ask concern whether the public art is accessible: "Is it for the people?" Does it encourage "participation?" Does it "serve [the constituency]?" (Deutsche 34). This emphasis on community participation dominates the wider discourse on urban regeneration strategies (Pollock & Sharp 2012, Jones 2003, Fung & Wright 2003).

Pollock and Sharp (2012) discern the origins of participatory policies, arguing that in place of "passive representational democracy", participation can forge active citizenship and reflect ideals of public empowerment (3064). In effect, participation is centred on this redistribution of power - of transferring responsibility from authorities to community participants (Pollock & Sharp 2012, Jones 2003, Fung & Wright 2003).

According to Pollock and Sharp (2012), the emphasis on citizen participation can be traced to growing concerns over citizen apathy, in addition to the belief that such methods have genuine societal impact (3064). The "meteoric" rise of participatory projects is premised, Jones (2003) adds, upon the supposed benefits brought to projects on bases of "efficiency", "sustainability" and even the "empowerment" of participants (582). Ideally, projects that focus on community inclusion should induce some type of "social reform" and "relearning" (Jones 589). Fung and Wright (2003) defend these claims and recognise that such policy reforms:

> *aspire to deepen the ways in which ordinary people can effectively participate in and influence policies which affect their lives. They are participatory because they rely upon the commitment and capabilities of ordinary people to make sensible decisions through reasoned deliberation and empowered because they attempt to tie action to discussion (qtd. in Pollock & Sharp 3064).*

With such value placed on participation, the authors recognise that these "empowerment" strategies are now concretely embedded in the discourse of urban regeneration practices (Pollock & Sharp 3064).

While participatory methods may appear justifiable at first glance, critical literature has emerged revealing how these policies are in fact little more than tokenism (Jones 2003, Pollock & Sharp 2012), with some even suggesting that participation has become a new "tyranny" (Cooke & Kothari 2001). Leading the contemporary criticism are Pollock and Sharp (2012), who identify the three main shortfalls of participatory policies and bring into question the extent to which participation is real and realised (3064). Primarily, they express concern over the citizen's capacity to negotiate on the same level as other stakeholders (Fraser 1996). Secondly, they question the very institutional structures and modes of discourse through which decisions are made, which are inherently paternalistic and top-down (Dargan 2009). And thirdly, the authors express concern about the motivation behind participatory policies, which they believe disguise a wider political agenda (Pollock & Sharp 3064). This third criticism is echoed by Jones (2003), who suggests that participation is solely implemented on the basis that it slots tidily into "the pre-determined and externally defined aims and objectives of a programme or project" (588).

> *The presence of "the aesthetic" - whether embodied in art-works…helps give redevelopment democratic legitimacy since, like "the public," "art" often connotes universality, openness, inclusion. "Public art, "combining the two terms, comes doubly burdened as a figure of universal accessibility. (Deutsche 37)*

Despite the democratic ideals upon which they are based, such redevelopment programs, Deutsche (1992) states, "are profoundly authoritarian, technocratic mechanisms, transforming cities to facilitate capital accumulation and state control" (37).

In the same way that participation is criticised, so too has the term "community" inspired a chorus of critique (Pollock & Sharp 2012, Cochrane 1986). This comes as projects designed to regenerate areas are, after being branded under the vague banner of "community", seem to instantly assume an air of legitimacy (Pollock & Sharp 2012). The term community, as Cochrane (1986) argues, is used "as if it were an aerosol can, to be sprayed on to any social programme, giving it a more progressive and sympathetic cachet" (qtd. in Pollock & Sharp 3064). Further concerns centre on the wider implications of dubbing urban generation plans in relation to communities as MacLeavy (2009) explains, "the positioning of the community at the forefront of urban regeneration policy constitutes the medium through which the state continues to represent and intervene in the lives of citizens" (871).

Despite their clear criticisms of community empowerment, Pollock & Sharp (2012) do recognise that the term "community" is not uniquely indicative of false intentions, since the term is representative of the coming together of people through "common interests and compromise" (3064). However, their arguments - and the wider critical literature on community participation - maintain that the terms participation and community have predominantly become the means for an authoritative body to legitimise policy (3064, MacLeavy 2009, Cochrane 1986).

PARTICIPATION: THE DUTCH CONTEXT

Because the majority of literature on participation policies and culture concerns the United Kingdom (Pollock & Sharp 2012, MacLeavy 2009, Jones 2003) and the United States (Deutsche 1992, Grodach 2010), it is important for this research to take the Dutch context into account. The contemporary position of participation in the Netherlands is clearly laid out in the report on civic participation released by *The Netherlands Institute for Social Research* in March 2014. According to the report, the Dutch government is keen to promote the idea of a "participation society", a society in which people take responsibility for their own surroundings (Houwelingen et al. 242). The benefit of a participation society, according to the interview with Gijs Hendrix, is the increasing establishment of "social cohesion". Encouraging participation "should result in less alienation, less trouble in the street. If everybody knows each other then they will be more like friends instead of making trouble for each other" (Hendrix).

The idea of participation it is not a new phenomenon in the Netherlands. The report proposes that the present participation-based practices are representative of past civic duties (Houwelingen et al. 245). Due to the "lack of strong central administration" in early modern Dutch towns, residents were left with no choice but to take on tasks themselves (ibid.). These tasks varied from scale, but ranged from carrying out nightwatch patrols, to keeping the canals ice-free in freezing temperatures (ibid.). The report also documents how the wealthy participated by establishing hospitals and almshouses for the poor - though these acts were mostly for those belonging to the same religious community (ibid.).

The report is even able to give a percentage of participation levels in Dutch society. It suggests that in recent decades, the level of civic participation has remained stable: "two out of five Dutch citizens say they do voluntary work, spending an average of just under one hour per week doing so" (Houwelingen et al. 242). And, although the percentage of citizens reporting they had contributed to an issue of "national or international importance " had declined, the percentage of those who are active in their own community or municipality had grown (ibid.).

Most importantly however, the report exposes a critical flaw in labelling a country a "participation society", questioning who exactly these "participants" are. The authors assert that some citizens are more likely to participate than others, and that not all Dutch citizens participate to the same degree. The typical participant is "indigenous", "older", "well-educated" and in "paid employment" (Houwelingen et al. 244). Gijs Hendrix, the policy adviser for arts and culture in the East of Amsterdam, echoed this conclusion. Hendrix holds a weekly meeting open to the public ("Spreekuur") that encourages residents to come and propose cultural projects. He noted that from "the people who come to the Friday morning Spreekuur, it is mostly white 40 year olds, and mostly artists. So they are probably not representative of the local population" (Hendrix).

Although the state promotes the idea of a "participation society", and although there is evidence to situate participation in the Dutch context, it is evident that the current roles of participation are only assumed by a privileged demographic. This signals that there are underlying structures in participation policies that accommodate to certain participants.

PARTICIPATION: THE KUNST HEK PROJECT

Community participation is central to the survival of the Kunst Hek Project, and two types of participation are demanded within the project. Firstly, the public is required to make content for the frames, and secondly, a dedicated group of volunteers is needed to organise the project. Though the project has been successful in involving many Dapperbuurt residents along the way, the bureaucracy that was involved in the initial process did hinder those who could participate. During her time as head organiser, Anke Weber sent more than 20 letters to the local council. The process, she explains, "was quite complicated. You [had] to write letters and go to the city council and explain everything and [only] then they [decided]". Though the Kunst Hek Project did invite locals to exhibit their work, Weber's principal involvement also shadows how participatory the project actually was. This shows that the structure needed to organise the Kunst Hek Project merely mimicked the top-down agenda it wanted to resist.

However, due to the layered levels of participation at work in the Kunst Hek Project, those who are not able to participate on an organisational level are able to participate by exhibiting their work in the frames. This created an opportunity for those who do not fall into the typical participant (older, educated, employed) to participate. Still, this meant that participation was limited to the few aspiring artists in the neighbourhood.

Secondly, the realised participation of the Kunst Hek can be put into question. Participation is supposed to promote active citizenship and inspire social cohesion, but the Kunst Hek requires little interaction between the exhibitors, the organisers and the site. When asked if Weber was ever present at the site, she replied: "No, only when [the artists] ask." Weber only comes into contact with the participants via email, "And a lot of [people] who make an exhibition, I don't even meet them. No. They write a mail to me and usually they send an example [of their work]." Therefore, contact between the primary participants does little to echo the ideals of active citizenship advocated by the Dutch government.

Thirdly, roles of participation are limited to the fitting demographic. In order to have their ideas heard, participants of the Vogelaar scheme attended pitch meetings in which their ideas were either approved or rejected by the council (Weber). Though the Kunst Hek proposal went through, another of Weber's idea was rejected:

> *I also suggested to make a nice meal for the alcoholic people in the Oosterpark. I wanted to make a nice table...[of] Italian food...And [the authorities] said no, that is not a good idea because many people don't like the alcoholics.*

Though Weber perhaps simplifies the council's reasoning, the rejection of the proposal was ultimately up to whether the project aligned to the wider plans, and is indicative of the top down structure in which the Kunst Hek Project was orchestrated. The rejection of the dinner-for-the-alcoholics-of-Oosterpark plan also resonates with Deutsche's (1992) example of the eviction of the homeless in Jackson Park, New York. Both occupy public space, and both are denied the right to be included.

Though the Kunst Hek Project is built on ideals of community participation, the realised participation of the project is hugely limited in practice. As much as local participation is spurred through the changing nature of exhibitions, the project is shaped through top-down bureaucratic practices that resemble the institution that put it in place. Secondly, the scope of participation is greatly limited, as the participants do not come in contact with each other. And thirdly, the establishment of Weber's project (and rejection of her other) indicates which groups the participation roles aim to include. Consequently, the Kunst Hek Project merely reiterates the critiques forwarded by the literature, showing that these roles are not tokens of community empowerment, but in fact disguise a wider political agenda.

TEMPORALITY

As much as the vocabulary of democracy is imbued in public art, it is also important to take into account the weight of temporality in public art discourse (Hein 1996, Phillips 1989). For Phillips (1989) and Hein (1996), the two are synonymous. Temporality operates uniquely in the public domain as "[t]he reality of ephemerality" is "most persuasively and unmistakably felt in the vast public landscape" (Phillips 331). Unlike the "quiet refuge" and predictable routine that the private can provide, the public is composed of "shifting differences that compose and enrich it" (ibid.).

According to Phillips (1989), public art has also been exploited as a "modest antidote or grand solution, rather than perceived as a forum for investigation, articulation, and constructive reappraisal" (331). The one way that all those involved in creating public art can continue to remain critical to the medium is, according to Phillips (1989), to "support more short-lived experiments in which variables can be changed and results intelligently and sensitively examined" (331). Hein (1996) too sees the importance of temporality, "today's public art works may be impermanent and discontinuous... They may subsist only momentarily or in multiple instantiations, immaterially suspended" (Hein 2). Ephemeral public art also:

> *provides a continuity for analysis of the conditions and changing configurations of public life, without mandating the stasis required to express eternal values to a broad audience with different backgrounds and often different verbal and visual imaginations. (Phillips 335)*

This begs the question as to whether this can be found in other spaces dedicated to culture and visual aesthetics.

Here we may to turn to the role of the museum or gallery, which have been viewed as the very antithesis of public art (Deutsche 1992, Hein 1996). Though these cultural institutions remain open to the public, they are paradigmatically designed for "private aesthetic viewings" and often discourage entry by charging the public (Hein 4). Following the previous conclusion that public art is not tied to the material and can exist in space, could not their walls provide space for public art and discussion? Hein is adamant in her answer: "were [museums] to guarantee universal access to their contents, the items contained in museums would not be public art" (4). Though the founding function of the museum was to "liberate" work that previously lay confined in private hands and to place these in the public sphere by claiming them as "national property", these "objects became "privatized" and extracted from the public sphere by virtue of the very aesthetic appropriation that made them "museum pieces" (5). Unlike the museum, public art that embraces the ephemeral, "provides the flexible, adjustable, and critical vehicle to explore the relationship of lasting values and current events" (335). It is able to engage the more abstract interpretations of memory, site and meaning (Hein 5). Temporality is thus viewed as essential in rendering a work "public".

TEMPORALITY: THE KUNST HEK PROJECT

Temporality is a key factor to the format of the Kunst Hek Project. The frames have been placed there to provide a momentary insight into an artist of the neighbourhood, who may not have had such an opportunity through other cultural institutions. Although the exhibits are never intended to be permanent, in the last years the variation in programming has ground to a near halt. Weber describes how in the first years of the project, there was high demand from their neighbourhood to showcase work in the frames:

> *In the beginning we were very strict. The [time between each exhibition was] no longer than four weeks. You know, there were people asking [to showcase their work]. Last year we see...its more quiet. And now maybe [the exhibition changes] every two months.*

Such lack of demand shows how important temporality is in the public domain, before too long, a temporary project slips into its environment and becomes part of the unquestioned visual domain of the city. The demand to showcase work in the Kunst Hek Project has decreased to the extent that in periods when there are no submissions, the artists who help to organise the project, step in themselves: thus leaving behind any democratic ideals of participation. Teo Krijgsman, one of the photographers in the team, showcased the same series from November 2013 to April 2014:

> *Teo is one of the people in the Kunst Hek Project, so when there's no one asking for room, he always has an exhibition. Yes and also you know, when there's nothing, I don't want it to be empty. It has been one night empty, and then there was graffiti. (Weber)*

For Weber, showing the same thing is better than showing nothing at all. Furthermore, Weber sees no issue in using the work of a professional photographer who is part of the organising body. This raises an important question around the maintenance of the project, namely, if the aim was to give visibility to the people in the neighbourhood, has the project now unsubscribed to the very ideals it stood for? Krijgsman, the photographer, reflects on his decision to display his work and says, "In this case...it was pure a matter of photographs being 'at hand', since I was asked to publish on very short notice and I had to go abroad. So it was a quick action" (Krijgsman). Clearly for Weber and Krijgsman, the necessity to have the frames filled and to have art work visible trumps the ethical questions around who is showing what, and whether the "public" is being represented.

The large frames themselves also do not propose ideals of temporality; each frame is secured by a visible lock and protective glass. These are not frames that can be simply hung and taken away, they have been placed there with the idea to stay and have done so for five years now. But this was never the plan, "Actually, we got a permit for one year – a temporary permit – and then after a year we asked for a permanent permit and [the council] said no we don't do that" (Weber). The status of the Kunst Hek Project has therefore shifted from legal public art project, to illegal yet "tolerated" (Weber). Although the frames have lined the same fence for five years, they have occupied this space without authority to do so for the last four years.

This fluctuating state of legality poses a new dimension to the discourse of public art and urban regeneration that is little discussed in the literature. The space of quasi-legality in which the Kunst Hek Project is situated in neither commits the project fully in line with urban regeneration, nor as an illegal intervention to the Dapperbuurt landscape. Had the project been founded independently from the state, it would no doubt escape the labelling of "public art" by critics, and would rather be treated as an "intervention" as Cartier & Willis (2008) discuss.

3. IS THE KUNST HEK PROJECT PUBLIC ART?

In considering the presence/absence of ephemerality, the discussion raises some important questions on whether the Kunst Hek Project can be labelled as public art. Firstly, it raises the issue that as much as the site embraces ephemerality, the Kunst Hek merely echoes the structure of a traditional museum exhibition: the artwork is framed and rendered separate from the public, the work stays for a predetermined period of time and then silently makes way for the next display. Not only does the Kunst Hek Project mimic the museum structure, but it has even become the extension of an actual museum: in early May 2014 the Tropenmuseum ("Tropics museum"), a nearby museum, exhibited a photo series they had initiated with local school children. This annexing of the frames for institutional purposes, severely limits the autonomy of the site and its position in public art discourse.

Secondly, where academics have stressed the importance of temporality in public art, within Weber's own definition, this does not pose an issue. According to her, public art is "art made by the public. [It] can be made by anyone, on any place." This harks back to the "pragmatic" definition, which paints public art as "art installed by public agencies in public places and at public expense" (Hein 2). Weber does however linger on the question of public art in addressing graffiti. She notes that, "Places with graffiti…that's very public. There is a difference [to the Kunst Hek Project]… There has never been an exhibition from a graffiti [artist]. So for them, it's not public." By situating the Kunst Hek in opposition to graffiti in this example, Weber herself identifies that the project partakes in a wider – and less "public" - hierarchical framework.

Thirdly, the location is also incongruous to the public art promoted by the literature. The Kunst Hek Project is to be used by "either… people who live in the neighbourhood… Or [as] an exhibition about the neighbourhood" (Weber). The key organisers are Dapperbuurt locals (Weber). However, although the project aspires to be representative of the Dapperbuurt, the Kunst Hek Project in fact lies geographically outside of the neighbourhood. Weber explains, "actually the [fence] of the Oosterpark is not in the Dapperbuurt. So that was a kind of problem, but it is so connected to the Dapperbuurt, [that] we said, well, it's O.K." Phillips (1989) puts forward that the volume of viewers of public art does not matter, as long as the art asks the right questions (332). But in situating the Kunst Hek Project outside of the Dapperbuurt, the site is extremely limited in doing so. Drawing upon Hein's critical terminology, we might say that it is not able to project abstract interpretations of memory, site and meaning because the content is not relevant to the audience that comes in contact with it (Hein 5).

In consideration of these three arguments, the Kunst Hek Project is not public art: firstly, the site mimics the framework of a museum (to the extent where it is even appropriated by a museum), secondly the founder recognises the site is incongruous with the idea of public art, and lastly the very space it inhabits lies outside of the intended audience.

4. DOES IT MATTER?

Having identified that many aspects of the Kunst Hek Project are incongruous when viewed against a notion of "public art", it may be useful to view the project under a different term. Perhaps it should be categorised as part of a practice that values placing images in public space. In this way we can honestly appreciate the images by local school children, dog-walkers and other art enthusiasts when passing by. After all, the Kunst Hek Project was never initiated to challenge the public's perception, but rather was conceived to celebrate the talents of the ordinary and sometimes extravagant people in the neighbourhood:

> *It was funny, in the beginning of last year there was someone who painted some of the dogs in Oosterpark, and it was [shown] on TV. And that was really nice. So there were these frames, with the paintings of the dogs, and the dogs in front of their own paintings. And then they were talking about it – it was quite funny. That's like how people meet each other. (Weber)*

Instead of problematizing the Kunst Hek Project, we should be able to find value in what may simply be colourful ornamentation of the city. However, if the frames still posit some value, we need to ask what the future of the Kunst Hek Project is.

REIMAGINING THE KUNST HEK PROJECT

In light of the depleted interest in continuing the series, it is important to consider alternative ways in which the project can continue. Stephany van Veen, the "Participation Initiator" (Participatiemakelaar) of the Dapperbuurt, expressed that a change in location would increase the scope of the project:

> *[The organisers] want it in that place, on that fence. In my opinion it's better if it's more in neighbourhood. But I'm not forcing the change. They know the neighbourhood. But I can imagine that if it is placed more centrally, you will reach more people. At the current location, there are a lot of pedestrians who do not actually stand still.*

Van Veen indicated that the council had allocated some small funds recently for the site's upkeep. With this donation was agreement that the Kunst Hek management would oblige artists to host openings at new exhibitions. This, van Veen believes, would spark more active participation within the community:

> *[At these small exhibition openings, the artist] can invite their own network, because you are proud and want to expose their work, but you also invite people from the neighbourhood who are in a network that want to keep informed of the Kunst Hek. So new connections are made from one exhibition to the next; the network keeps getting bigger, and those people will also get to know each other and there will be new exchanges and more social cohesion. So in that way new things will arise because each network will inspire each other…. Then it really becomes something… Because if it is only one fence hanging there, then it's too passive.*

Another option Van Veen suggests is for the site to have more interaction with the restaurants and cafes on the other side of the road. So, when artists have an opening, they could make use of the inside space. The restaurants could also write a (perhaps digital) pointer to the Kunst Hek Project, in order to bring more people in touch with the site.

A different and more radical alternative I propose is to relocate the series to space with increased visibility and so encourage wider engagement with the public. For example, by placing the frames in spaces traditionally reserved for advertising (at tram stops, electricity cabinets, temporary walls etc.) the frame concept has far more agency to engage in public sphere. In this setting, the frames no longer replicate a gallery setting but open up space to critically examine the visual domination of advertising in cities. Of course, such an initiative would take a considerable amount effort to surpass the boundaries of bureaucracy, which, in turn, would bring about the same issues around participation and privilege as previously discussed. However, if such a project did not partake in the discourse of "participatory community-led projects", this would no longer be of issue.

Another alternative is to re-examine and make clear the medium through which the Kunst Hek Project operates. This is poignantly shown in the addition of a smaller frame to the collection in early May 2014. This frame did not belong to the official series and was seemingly un-authored. Although baring some resemblance to the format of the larger frames, the content and its size did not fit into the aesthetic presentation of the series. To any passerby, the smaller frame seems to sit somewhat awkwardly at the edge, hanging in a state of belonging and not-belonging. This new frame draws attention to the museum-like medium which the Kunst Hek Project currently subscribes to. And, in its amateur approach it delegitimises that format. By drawing attention to this, the new frame inevitably exposes the question that has concerned critics; does placing art outside

Figure 4. *The added frame in the Kunst Hek Series*

deem it *public art*? Or more specifically to this case, does placing pictures in a frame on a fence render something public art? Hopefully this paper has been able to argue that this is not the case.

CONCLUSION

In staging a dialogue between the literature on public art and urban regeneration, the cultural policy makers involved (in the Netherlands) and of the public participants, this paper analyses the Kunst Hek Project through vocabulary that is specific to the site. As Cartier and Willis (2008) identify, because public art assumes such a wide array of forms, it requires a flexible terminology in order to understand it. In this case, the terms "democracy", "participation" and "temporality" were pivotal points of entry into the contested and problematic discourse of the Kunst Hek Project.

In its current form, the Kunst Hek Project will remain a site that provokes little reaction to those who see it, and will continue to go by unnoticed by the majority who do not. The seven frames on the fence sit all-too-neatly into the planned surroundings. In their fixed form, composed of locks, frames and glass, the site neglects the importance of ephemerality, a component of public art so valued by critics. As a consequence, the frames are not read as "public art" but rather become akin to their distant cousin, the information board. Though the site might act as a platform for local art enthusiasts, ultimately, the Kunst Hek Project does not disrupt public space and is not capable of invoking abstract interpretations of memory, site or meaning.

In aligning the Kunst Hek with the wider themes of urban regeneration practices, the discussion around the site enters further vast grounds of contestation. Created as a result of a national drive towards participatory policies, the site sits too firmly imbedded in the popular and tainted discourse of democracy to ever be removed from this context. After all, Weber's project was funded because it reflected the wider values the local council wanted to imbue. Unlike her other proposal to host a park dinner for the alcoholics, the Kunst Hek Project projected just the right type of participation and set out to include the right people. As is echoed throughout the literature, an increasing emphasis on civic participation is evidence of increased state intervention. Furthermore, the plans to devolve the burden of action to "communities" can be seen as a way of displacing responsibility and obligation to those same "communities".

However, the current semi-legal position that the Kunst Hek Project now finds itself in reconfigures the conversation around public art and urban regeneration. In its illegal-yet-tolerated form, the Kunst Hek Project continues to embody both the authoritative structures that put it in place and the more challenging aspects that comprise great public art works. This suggests that there are limitations to understanding the site in terms of the existing literature on public art, to which this paper has aimed to provide a contribution.

Whether the Kunst Hek Project takes on a new and more critical form or stays put in its quasi-legal state, the project should aspire to value temporality in order to distinguish itself from the permanence of its surroundings. And, more importantly, by embracing ephemerality, the Kunst Hek Project will be able to continue to delight and spark conversations in the neighbourhood with its presence.

BIBLIOGRAPHY

Cameron, Stuart, and Jon Coaffee. "Art, gentrification and regeneration–from artist as pioneer to public arts." *European Journal of Housing Policy* 5.1 (2005): 39-58.

Cartiere, Cameron, and Shelly Willis, eds. *The Practice of Public Art.* Routledge, 2008.

Cochrane, Allan. "Community politics and democracy." Cited in: Held, David, and Christopher Pollitt. *New forms of democracy*. Sage Publishing, 1986.

Cooke, Bill, and Uma Kothari, eds. *Participation: The new tyranny?*. Zed Books, 2001.
"Dapperbuurt, Amsterdam." Map. Google Maps. Google, 10 May. Web. 10 May.

Dargan, Lorna. "Participation and local urban regeneration: the case of the New Deal for Communities (NDC) in the UK." *Regional studies* 43.2 (2009): 305-317.

Denzin, Norman K., and Yvonna S. Lincoln, eds. *The SAGE handbook of qualitative research*. Sage, 2011.

Deutsche, Rosalyn. "Art and public space: Questions of democracy." Social Text (1992): 34-53.

Fraser, Nancy. *Justice Interrupts*. Routledge, 1996.

Fraser, Nancy. *Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy.* Milwaukee, WI: Univ. of Wisconsin-Milwaukee, Center for Twentieth Century Studies, 1990. Print. Cited in Deutsche, Rosalyn. "Art and public space: Questions of democracy." Social Text (1992): 34-53.

Fung, Archon, and Erik Olin Wright. "Deepening democracy: innovations in empowered participatory governance." *Politics and society* 29.1 (2001): 5-42.

Grodach, Carl. "Art spaces, public space, and the link to community development." *Community Development Journal* 45.4 (2010): 474-493.

Hall, Tim and Robertson, Iain. "Public Art and Urban Regeneration: advocacy, claims and critical debates." *Landscape Research* 26.1 (2001): 5–26

Hein, Hilde. "What is public art?: Time, place, and meaning." *Journal of Aesthetics and Art Criticism* (1996): 1-7.

Houwelingen, P. Van, Anita Boele, Paul Dekker, and Julian Ross. *Burgermacht Op Eigen Kracht?: Een Brede Verkenning Van Ontwikkelingen in Burgerparticipatie.* Den Haag: Sociaal En Cultureel Planbureau, 2014.

Jones, Peris S. "Urban regeneration's poisoned chalice: is there an impasse in (community) participation-based policy?." *Urban Studies* 40.3 (2003): 581-601.

Kay, Alan. "Art and community development: the role the arts have in regenerating communities." *Community Development Journal* 35.4 (2000): 414-424.

Krijgsman, Teo. Personal Interview. 3 Dec. 2013.

ter Kuile, Rosa. "Blog Homepage." *organisingpublicspace.tumblr.com*. Tumblr. Screenshot. 10 May 2014.

ter Kuile, Rosa. "Kunst Hek Project in Autumn." 2013. JPEG file.

ter Kuile, Rosa. "A new frame added to the Kunst Hek, Spring." 2014. JPEG file.

MacLeavy, Julie. "(Re) Analysing community empowerment: Rationalities and technologies of government in Bristol's new deal for communities." *Urban Studies* 46.4 (2009): 849-875.

Myerscough, John. *The economic importance of the arts in Britain*. Policy Studies Institute, 1988.

Permentier, Matthieu, J. Kullberg, and Lonneke Van Noije. Werk Aan De Wijk: *Een Quasi-experimentele Evaluatie Van Het Krachtwijkenbeleid*. Den Haag: Sociaal En Cultureel Planbureau, 2013. Print.

Petherbridge, Deanna. *Art for Architecture: a handbook on commissioning.* Department of the Environment/ HMSO. 1987.

Phillips, Patricia C. "Temporality and public art." *Art Journal* (1989): 331-335.
Pollock, Venda Louise, and Sharp, Joanne. "Real participation or the tyranny of participatory practice? Public art and community involvement in the regeneration of the Raploch, Scotland." *Urban Studies* 49.14 (2012): 3063-3079.

Prasad, Pushkala. *Crafting qualitative research: Working in the postpositivist traditions*. ME Sharpe, 2005.

"Public Art". Tate. N.d. Web. 18 May 2014. <www.tate.org.uk/learn/online-resources/glossary/p/public-art>.

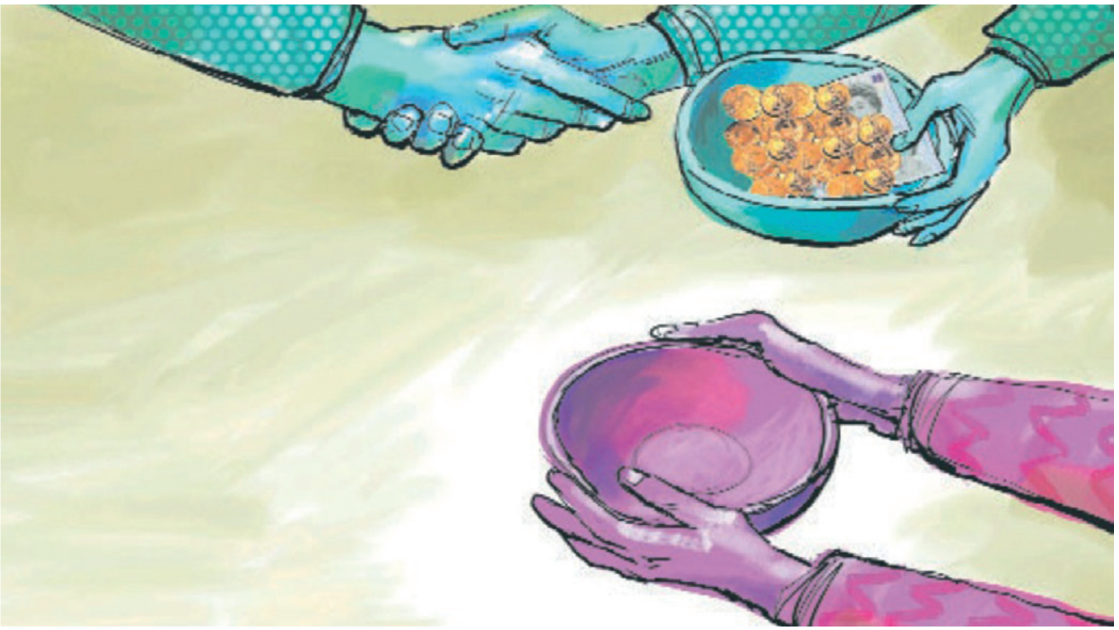
Sharp, Joanne, Venda Pollock, and Ronan Paddison. "Just art for a just city: public art and social inclusion in urban regeneration." *Urban Studies* 42.5-6 (2005): 1001-1023.

Weber, Anke. Personal Interview. 10 Feb. 2014.

The Identity Crisis of the Welfare State: An Intergroup Contact Analysis Of Welfare Chauvinism

Joep Jansen

Supervisor: Dr. M. Eijberts (AUC)



ABSTRACT

The emergence of a progressive dilemma has had a profound impact on Western welfare states in recent years. According to the advocates of the progressive dilemma, ethnic diversity may be inherently damaging for the welfare state. Alternatively, it has been argued that multicultural policies may have created an environment that is unsustainable for the welfare state. While the literature has established a weak negative correlation between ethnic heterogeneity and income redistribution, multiculturalism as such, on the other hand, does not seem damaging for the welfare state. However, there is a possibility that a particular interpretation of multiculturalism does contribute to a progressive dilemma. To examine this possibility, and arrive at a possible solution for the issue of welfare chauvinism, Allport's (1954/1979) intergroup contact theory, in combination with Van Oorschot's (2006) conditions for deservingness, is proposed as a theoretical framework to analyze this problem. By applying this theoretical framework to the Dutch case, it is found that the separatist multicultural policies of the 1980s contributed to an environment where positive, prejudice-reducing intergroup contact was unlikely. On the other hand, strains on Dutch national identity made the inclusion of immigrants into the native Dutch in-group improbable. Together, separatist multiculturalism and pressures on Dutch identity made for a situation where particularly non-Western or Muslim immigrants could easily be perceived as a threat to Dutch in-group welfare and undeserving of support. Current trends in integration policy towards assimilation and culturalization are found to offer no solution to the issue. Instead, potential solutions are found in Canadian 'integrative multiculturalism' and Parekh's (2001) "community of communities."

INTRODUCTION

The welfare state is in a crisis, and one that has been long coming. That is, the underpinnings of income redistribution in Western societies, as some see it, have come under immense pressure from increasing ethnic diversity. To understand the problem, one must go back to the founding myth of the welfare state. As the British politician David Willetts (in Goodhart, 2004) asserts, the welfare state was founded on the basis of societal homogeneity; contributors were willing to pay into the system, because it supported people very similar to themselves who were confronted with 'normal' problems. Nowadays, the situation is decidedly different; the times of ethnic homogeneity in Western societies is long gone, and a contributor to the welfare state can no longer assume that his money will go on to support his own spitting image. According to Willetts, this has made the welfare state lose its legitimacy for many: "Why should I pay for them when they are doing things I wouldn't do?" (cited in Goodhart, 2004, p. 30). Suddenly, the old progressive ideals of lenient immigration policies and a generous welfare state seem irreconcilable; people are faced with a 'progressive dilemma' to choose one or the other.

The progressive dilemma is part of a wider surge of welfare chauvinism, which refers to the phenomenon of a native group desiring or attempting to exclude immigrants from welfare provisions (Freeman, 2009). Throughout Europe and North America, calls have been made to limit access to the welfare state to native groups only, in response to increasing ethnic diversity in Western societies (Banting, 2010). Proponents of the progressive dilemma (Goodhart, 2004; Klausen, 2000) essentially make two claims: some argue that ethnic heterogeneity in itself is damaging for the welfare state; others have criticized the multicultural approach many Western countries have taken to accommodate for this diversity.

While a growing literature has documented this development, theoretical explanations of the issue are still lacking (Koning, 2013). In response to this gap, this thesis will explore how Allport's (1954/1979) intergroup contact theory can help understand the phenomenon of welfare chauvinism. A theoretical framework will be proposed which combines intergroup contact theory with Van Oorschot's (2006) conditions for deservingness, as a valuable framework for studying the phenomenon of welfare chauvinism. To this end, these theories will be applied to the Dutch case, in an effort to gain an understanding of the dynamics underlying the exclusion of immigrants from welfare provisions. Moreover, the case study will serve to critique current Dutch integration policies and to arrive at new solutions for this issue.

With regard to the larger structure of the thesis, it will first provide a brief overview of the way in which ethnic diversity came to be framed as a problem. Subsequently, the available literature will be reviewed to determine what the evidence is for the hypotheses of the progressive dilemma. The thesis will then move on to discuss the development of intergroup contact theory and its potential for explaining welfare chauvinism. Finally, this framework will be applied to the Dutch case, after which a few potential solutions will be considered.

FRAMING DIVERSITY AS A PROBLEM

Although neither the welfare state nor migration from the Global South to the Global North are particularly new phenomena, a potential relationship between the two has become the subject of considerable scholarship only fairly recently (Koning, 2013). Moreover, as Careja and Emmenegger (2013) point out, this relationship has not always been considered in terms of an inherent tension; nonetheless, the initial optimism about the expansion of access to benefits for non-nationals has since overwhelmingly been replaced by the notion that ethnic heterogeneity poses a threat for Western welfare states. This section will outline how diversity, and political support for diversity in the shape of multiculturalism, came to be framed as a problem.

THE PROBLEM WITH DIVERSITY

Pointing at growing global integration and increasingly dense, transnational interconnections across the globe, many globalization theorists predicted the demise of the nation-state in favor of less localized forms of belonging (Carnoy and Castells, 2001). Castells (2000, p. 696), for instance, envisioned the advent of a network society, a product of the modern technologies allowing for globe-spanning communication, in which the past focus on concrete localities would gradually be replaced by a global 'space of flows'. This argument echoes Sassen's (1991) concept of the Global City, a space composed of geographically distant but (economically) integrated places characterized by a discontinuity from the national contexts the component parts are in individually. Similarly, Lechner and Boli (2005/2010, p. 421) have argued that an expanding 'world culture' may come to substitute nationality for new, global dimensions of identity. Their views on globalization can be seen as part of a larger trend which emphasizes the unifying potential of globalization, engendering cosmopolitan attitudes across the globe (Mewes and Mau, 2013, p. 229)

Unsurprisingly, this hypothesis produced a fair amount of optimism among scholars with regard to the broadening of social rights beyond the realm of citizenship (Koning, 2013, p. 1). In an extension of T. H. Marshall's (1950/2010) famous model of the evolution of citizenship to include first civil, then political, and ultimately social rights[1)], Soysal (1994) proposed a 'postnational' model in which national citizenship is no longer a requirement to benefit from these rights previously reserved for citizens of national communities. Interestingly, she draws on the example of postwar labor migration to countries such as Germany and the Netherlands to illustrate that, despite their initial lack of formal citizenship, guest workers were still able to attain a permanent position in their host countries and eventually gained access to similar, if not the same rights as native citizens (ibid., p. 2). In this way, Soysal (2000, p. 5) suggests a "decoupling of rights and [national] identity" as a result of increased global migration and transnationalism.

Yet such optimism was soon called into question, both by scholars who criticized the predicted demise of the national citizenship as premature (Carnoy & Castells, 2001; Mann, 1997), as well as by actual developments in Western welfare states in the 1990s and 2000s (Careja & Emmenegger, 2013, p. 140). Mann (1997, p. 494), for example, disputes the assumption that the forces of globalization necessarily threaten the nation, finding instead that the same forces may also strengthen particular aspects of the nation-state. More importantly, several Western countries started intentionally shutting non-nationals and even new citizens out from certain social welfare provisions (Careja & Emmenegger, 2013, p. 151). In Europe, numerous countries engaged in strategies of 'differential exclusion', including immigrants in certain capacities (as cheap labor) but excluding them from the social rights of 'real' citizens (Castles, 1997, p. 7). Rather than weakening national identity and broadening access to welfare, mass global migration moved several European states, including the Netherlands, to implement stricter integration policies, which in fact raised the requirements for immigrants to become eligible for citizenship and social rights (ibid., p. 8). By the mid-2000s, Banting and Kymlicka (2006, p. 40) observe, policies had been implemented in virtually every Western nation to instill national sentiments in its population and reinforce the significance of citizenship.

Yet for the Dutch case in particular, the 1990s were a period of transition from one extreme – the 'lenient' multicultural approach – to another – policies of (implicit) assimilation (Vasta, 2007). For much of the 1960s and 1970s, the Dutch government had treated the growing groups of labor migrants as temporary, and had therefore encouraged immigrants to preserve their native culture (Snel & Scholten, 2005). Although it soon became clear that the guest workers were settling in the Netherlands permanently, the multicultural policies of the 1980s still emphasized the maintenance of minority culture as a means to achieve greater societal participation (Eijberts, 2013, p. 119). However, the restructuring of the Dutch economy towards a greater focus on the service sector led to rising unemployment levels among immigrants, most of whom had been employed in low-skilled, industrial jobs (Entzinger, 2006, p. 181). To account for the growing immigrant unemployment, the multicultural project of the 1980s became widely perceived as having insufficiently encouraged the cultural integration of immigrants, and was thus branded a failure (Snel & Scholten, 2005). In response, the Dutch government came to demand increasing cultural integration from immigrants, bringing Dutch immigration policies closer and closer to the model of assimilation (Vasta, 2007). The harshening of the official immigration and integration discourse reflected growing anxiety about the pressures immigrants supposedly placed on the Dutch state (Koning, 2013, p. 108). Observations of the bad position of immigrants on the labor market soon translated to a fear that they would make a generous welfare state unsustainable; despite the general consensus in the 1990s that to voice such concerns would be politically incorrect, the Dutch government's disquiet about this presumed tension continued to grow (Entzinger, 2006, p. 182). By the early 2000s, a report by the Scientific Council for Government Policy (WRR, 2001, p. 121) called the overrepresentation of immigrants in welfare provisions "one of the primary sources of concern," not just in terms of its supposed effect on social spending, but also in reaffirming societal prejudice against this group. Before long, Dutch politicians were denouncing the multicultural policies of the past as having been a fateful mistake (Entzinger, 2006, p. 184).

THE PROGRESSIVE DILEMMA

The retreat of multiculturalism in the Netherlands was part of a wider backlash in Western welfare states; as the persistence of socio-economic equalities along ethnic lines and the overrepresentation of ethnic minorities in welfare programs inevitably raised questions about the efficacy of past policies, the multicultural approach was often singled out as having been too lenient with and insufficiently demanding of immigrant groups (Banting & Kymlicka, 2006). By failing to encourage immigrants to integrate into society, multiculturalism was condemned for contributing to their segregation and marginalization, leading to an overrepresentation of immigrants

1) According to Marshall (1950/2010, p. 32), citizenship in the era of the welfare state entitles citizens not just to individual liberties and the right to political participation, but also "the right to a modicum of economic welfare and security to the right to share to the full in the social heritage and to live the life of a civilized being according to the standards prevailing in the society," as guaranteed by social welfare and educational programs.

in welfare provisions (ibid., p.2). If growing ethnic heterogeneity was already in tension with the maintenance of generous welfare policies, complicating the image that by contributing one was simply supporting one's 'brother' (Crepaz, 2006) – someone highly akin to oneself – then multiculturalism was only accentuating the supposed discrepancy between the average welfare beneficiary and this traditionally assumed mirror image (Goodhart, 2004, p. 30-1). Or rather, if the welfare state made increasing demands of its benefactors to trust those on the other side of racial and ethnic divides, then these ethnic divisions were allegedly made all the more pronounced by multicultural policies, which served to acknowledge rather than to unify different ethnic groups under one (national) identity (Miller, 2006, p. 325-6). Suddenly, those (generally on the political left) who had previously supported 'relaxed' immigration and integration policies as well as substantial income redistribution, now felt that they had upheld conflicting causes – they were faced with a "progressive dilemma"[2] (Goodhart, 2004, 2013).

According to Goodhart (2013, p. 262), the foundation of the welfare state relied on the assumption that contributors part with a portion of their income to support people similar to them who undergo difficulties which could possibly one day also arise for themselves. Yet due to increasing ethnic heterogeneity in Western societies, he argues, the main contributor to the welfare state – the ordinary, hard-working citizen – can no longer rightly imagine his charity to benefit a member of his own ethnic group (Goodhart, 2013, 2004, p. 33). This tension between solidarity and diversity thus erodes the basis on which the modern welfare state is built (ibid., 2013, p. 261). Contrary to the early optimism from the part of globalization theorists such as Soysal (2000) who anticipated a future where rights were no longer exclusively tied to (recognized) national identities, proponents of the progressive dilemma perceive an inherent conflict between the broadening of such identities and the maintenance of social rights.

As Banting and Kymlicka (2006, p.3-4) explain, the progressive dilemma essentially revolves around two distinct hypotheses: firstly, advocates of the hypothesis imagine themselves as having to make a difficult choice between supporting redistributive policies or ethnic diversity, as the latter is seen as intrinsically undermining feelings of unity and solidarity necessary for the former; secondly, they claim this fundamental contradiction of interests is exacerbated by the multicultural policies of many Western welfare states which actively recognize and promote diversity rather than reducing it. In other words, whereas some assert that ethnic diversity as such is damaging for the redistributive welfare state, others blame the multicultural approach in particular – or, at the very least, that multiculturalism has only added to the problem (Banting, Johnston, Kymlicka, & Soroka, 2006, p. 49). Interestingly, as some have argued (Banting & Kymlicka, 2006; Banting et al., 2006; Crepaz, 2006), proponents of the progressive dilemma hypotheses made their claims at the time – that is, the early 2000s – on the basis of strikingly little empirical evidence. Nevertheless, studies on the topic have greatly increased in number since (Koning, 2013, p. 2-3), which warrants a re-evaluation of the available research.

A REVIEW OF THE EVIDENCE

Despite recent developments in empirical research on welfare chauvinism (Koning, 2013), such studies are still modest in number, and the literature overall is rather fragmented. This is partly because, as Banting & Kymlicka (2006, p. 3-4) stated, scholars have essentially been tasked with testing two distinct hypotheses: firstly, it must be determined whether ethnic heterogeneity as such reduces support for income redistribution; secondly, there is the question of the influence of multiculturalism, either direct or indirect, on public welfare state support. Additionally, recent studies employ different units of analysis and vary greatly in their chosen geographical context. While advocates of the progressive dilemma have mostly pointed at developments on the national scale (e.g. Goodhart, 2004, 2013; Klausen, 2000), the literature on diversity and the welfare state covers several levels of analysis; studies focusing on individual nations (Eger, 2010; Entzinger, 2006; Kraus & Schönwälder, 2006) and their neighborhoods or communities (Putnam, 2007; Tolsma, Van der Meer, & Gesthuizen, 2009) have increasingly been complemented by cross-country analyses (Burgoon, Koster, & Van Egmond, 2012; Mau & Burkhardt, 2009; Wright & Reeskens, 2013). This section will first discuss some of the inconsistencies in the research into welfare chauvinism, and will then review the cross-national comparisons and case studies which have been conducted on the subject.

ADDRESSING THE INCONSISTENCIES

Before the empirical evidence for either sub-hypothesis of the progressive dilemma can be examined, the ambiguity of the concepts to which they refer must be addressed. Firstly, if advocates of the hypotheses argue that 'diversity' or the 'multicultural approach' to such heterogeneity negatively affect the sustainability of 'the welfare state', they do so using definitions of these terms often left implicit. For instance, even though Goodhart (2013, p. 267-273) uses the term (income) 'redistribution' and the much more general 'welfare state' fairly interchangeably, it is actually only the former with which his theory is concerned. After all, those welfare provisions which require contribution as a condition for access, and which are grouped under the term social insurance (Sainsbury, 2006), are much less susceptible to the 'free rider' issue – where people are able to exploit the system without contributing to it (Goodhart, 2013, p. 269). Such problems may only arise for non-contributory programs funded by general taxes – also referred to as social assistance (Sainsbury, 2006) – and so it is here, on the redistributive side of the welfare state, where questions of fair use and entitlement emerge (Goodhart, 2013, p. 269-270). The social assistance system, Goodhart (2013, p. 270) claims, relies on contributors' trust that beneficiaries will use their support responsibly and will not simply take advantage; as a result of ethnic diversity, he continues, exactly this type of trust is eroded as "[too] many welfare recipients have become hard for ordinary citizens to identify with."

Scholars are similarly inconsistent in their use of the term 'multiculturalism' (Banting & Kymlicka, 2006, p. 28-9), and as Banting et al. (2006, p. 51) observe, many leave the concept undefined. In its most general form, multiculturalism refers to the recognition of minorities, yet Banting et al. (ibid.) draw attention to the fact that in the case of the progressive dilemma, only the policies implemented to promote such recognition are taken into account. One should therefore make a clear distinction between multiculturalism on the policy level, and multiculturalism as a discourse, demographic concept, or personal attitude, all of which have undergone very different developments in the Dutch case (Van de Vijver, Breugelmans, & Schalk-Soekar, 2008). Another point of imprecision is the selection of minorities targeted by multiculturalism; these can be categorized on the basis of an ethnic or sexual identity, or even characteristics not directly related to identity (e.g. disabilities) (Banting et al., 2006). It is only the first form of diversity, ethnic heterogeneity, which is considered in the model of the progressive dilemma. Still, countries differ greatly in their implementation of the multicultural ideal – the accommodation of (ethnic) minorities – and even countries with no comprehensive multicultural agenda can at times pursue typically multicultural strategies (ibid., p. 75; Kraus & Schönwälder, 2006, p. 210). As very few cross-country studies systematically define multiculturalism, this could certainly be a weakness in the literature on the relation between multicultural policies and income redistribution.

Furthermore, Banting et al. (2006, p. 52) note that criticism of multiculturalism is centered mostly on "policies that go beyond the protection of traditional individual rights of citizenship to provide some additional form of public recognition or support or accommodation of ethnic groups, identities, and practices." Therefore, in comparing multicultural policies and redistribution in Western states, they limit their definition of multiculturalism to include only the more extreme policies of support for ethnic minority groups, and thereby exclude more moderate measures such as anti-discrimination policies specifically targeting certain minorities (ibid., p. 51). One could argue that this is in fact too narrow a definition, since such anti-discrimination policies entail the same reaffirmation of ethnic boundaries lamented by critics such as Goodhart (2004, p. 36) or Barry (2001, p. 88) for prioritizing the recognition of minority groups over the enforcement of a unified national identity.

This leaves the question of how to define the 'ethnic diversity' which supposedly undermines support for the welfare state. In his study of post-war immigration to Great Britain, Goodhart (2013, p. xiv) focuses mostly on immigrant groups of non-Western origin, largely ignoring non-'visible' minorities (such as those from other European countries). Perhaps this approach might make sense considering Goodhart's (2004, p. 33) emphasis on the alleged lack of identification between welfare contributor and beneficiary, since visible minorities may indeed be perceived as 'more different' by the dominant group than those minorities not differentiated from the majority by physical markers. Nonetheless, studies of

2) The same phenomenon has also been referred to as the "progressive's dilemma" (Freeman, 2009) or the "New Liberal Dilemma" (Reeskens & Van Oorschot, 2012).

the relation between ethnic heterogeneity and social spending of states have used the concept of diversity to variously include recent immigrant groups, largely assimilated groups still identifying with a certain cultural heritage (e.g. Irish Americans or English Canadians), indigenous peoples and national minorities (Banting & Kymlicka, 2006, p. 29-30). The lack of a single particular definition of ethnic diversity in national-level datasets has been a sizeable obstacle for cross-country comparison; perhaps as a result, such studies have emerged only fairly recently.

ASSESSING THE EVIDENCE

Larger-scale, cross-national comparisons have mostly concentrated on the first of the two hypotheses, examining whether ethnic heterogeneity is inherently damaging to income redistribution. Very few have taken into account the potential influence of multicultural policies implemented in many Western states (Banting & Kymlicka, 2006). Moreover, the research in this tradition, concerned with Western welfare states specifically, has produced rather conflicting results, although cross-country studies generally support the existence of a negative correlation between ethnic heterogeneity in Western societies and support for the welfare state (Mau & Burkhardt, 2009; Senik, Stichnoth, & Van der Straeten, 2009; Wright & Reeskens, 2013).

As one of the first cross-national studies, Alesina and Glaeser (2004, p. 133-182) compare the US and virtually all member states of the EU on the ethnic heterogeneity of their populations and redistributive spending, asserting that the large gap between the 'residual' American approach to welfare and the more generous European welfare states can be attributed to their differences in institutional arrangements and centralization on the one hand, and demographic compositions on the other. Particularly racial fragmentation, which is most severe in the Netherlands compared to many other European nations (ibid., p. 139-40), is found to have a substantial negative effect on social spending, and is cited as the reason why the US lack a 'European-style' welfare state (ibid., p. 182). Unfortunately, their results are derived from the fractionalization index (p. 137) developed by Alesina et al. (2003), which has been heavily criticized for its reliance on inconsistent measures of diversity across countries.

Mau and Burkhardt (2009) have attempted to avoid this pitfall by examining not the potential effect of actual rates of foreign-born or minority populations, but public opinion surrounding the eligibility and deservingness of immigrants for social support. This way, the effect of perceived levels of immigration on support for more or less exclusionary welfare policies can also be tested more directly. For instance, using data collected in the European Social Survey 2002/2003, Mau and Burkhardt (2009) conclude that not all minority populations have an impact on public support for redistribution; only the non-Western, foreign-born groups have a negative effect on such support.

Another study based on the same survey data conducted by Senik et al. (2009) finds a similarly weak correlation between ethnic diversity and support for the welfare state when comparing these aspects in 22 European countries. While on the whole, the correlation between the ethnic fragmentation of a particular country and public support for redistribution is not very significant, they assert that there are considerable differences between countries (ibid., p. 359). Moreover, contrary to Mau and Burkhardt's (2009) study, Senik et al. (2009) have more attention for the influence of individual perceptions; for example, they use the share of immigrants in the overall population as perceived by participants rather than the actual portion of foreign-born people. Their attention for the level of the individual allows them to pinpoint another important finding: for people who dislike immigrants in general or worry about their economic impact, the perceived presence of immigrants does have a much more significant impact on their support for welfare provisions (ibid., p. 357).

However, as these 2002/2003 European Social Survey results date from before the emergence of the progressive dilemma in the political discourse of many European countries, one might argue they are rather outdated. Since the recent prevalence of this discourse may well be expected to reflect a change in, or have a significant impact on public opinion (Banting & Kymlicka, 2006, p. 44), more recent data must be considered to take these developments into account. An example is the research conducted by Wright and Reeskens (2013) using data from the 2008 European Values Study in 29 EU and OECD countries, which found that people who thought of their national community as bounded by cultural or ethnic characteristics were more liable to attitudes of welfare chauvinism. That is, those who define their national identity in terms of a particular ethnic or cultural identity appear to be especially prone to excluding others from national welfare provisions (ibid., p. 1457).

Nevertheless, as Burgoon (2011, p. 3-4) indicates, the general patterns identified in this small number of cross-country empirical studies could be weak because these studies have only considered one side of the story. Burgoon's (ibid., p. 1) analysis of the 2002 and 2008 waves of the European Social Survey demonstrates that the progressive dilemma may not apply to all segments of society in the same way, and that for some people increasing diversity heightens support for redistribution. Also drawing into question the idea that immigration could merely have a negative impact on support for the welfare state, Burgoon et al. (2012) support this alternative effect. They argue that personal interests may well weigh out the influence of national-level immigration by pointing out that those employed in sectors with relatively high levels of foreign-born workers are actually more likely to support redistributive policies due to the increased economic insecurity they believe themselves to be in (ibid., p. 292). Furthermore, Burgoon (2011) suggests that the progressive dilemma may not affect all immigrant classes equally: those who are socially and economically well integrated are less likely to become the target of welfare chauvinism.

If the socio-economic integration of immigrants plays such an important role in preventing a potential progressive dilemma, the question then becomes what the effect of integration policies has been in preventing or contributing to a progressive dilemma. Inspired by the writings of Goodhart (2004, 2013) and Barry (2001), Koopmans (2010, p. 2) has proposed that the reason for immigrants' disadvantaged labor market position and overreliance on welfare provisions in some states should be sought in those countries' multicultural approaches to diversity. In brief, Koopmans argues multicultural policies fail to sufficiently enforce cultural integration of immigrants and intergroup interaction, leading to the segregation and marginalization of immigrant groups in Western societies (ibid., p. 3). Drawing from a large range of quantitative data, he compares the welfare policies of eight Western European countries as well as the socio-economic integration of their ethnic minorities, and concludes that the combination of a generous welfare state and multicultural policies is related to more disadvantaged and welfare-dependent minorities (ibid., p. 20-21). As a consequence, Koopmans implies, these traditionally strong welfare states may choose to limit both the generosity of their provisions as well as the groups that have access to them (ibid., p. 22). These findings appear to support the second hypothesis of the progressive dilemma, which states that multicultural policies directly or indirectly undermine the welfare state.

Nevertheless, few studies have been conducted on this relationship, and empirical evidence for this claim has remained rather weak (Banting et al., 2006; Banting & Kymlicka, 2006, p. 5). Although Koopmans (2010, p. 11) offers a lot of seemingly robust quantitative evidence, he fails to systematically take into account alternative, potentially confounding variables (e.g. the level of human capital with which specific groups enter a host country, or macroeconomic developments which influence their ability to find suitable employment) that might cause the socio-economic marginalization of immigrants due to a lack of such specific data. Moreover, his study does not test the perceived correlation across more than a single point in time, which means one can only speculate about a causal relationship between multicultural policies and migratory pressures on the welfare state. This shortcoming is not unique to Koopmans' (2010) research, but extends to much of the cross-national research on ethnic diversity and the welfare state (e.g. Alesina et al., 2003; Burgoon et al., 2012; Wright & Reeskens, 2013), which includes virtually no longitudinal studies of the matter.[3]

These are, however, not the only flaws one can identify in Koopmans' (2010) work. Besides failing to consider alternative factors and to do so longitudinally, it appears Koopmans (2010, p. 3) may have also selected his cases rather conveniently: he leaves Canada, a long-time champion of multiculturalism (Winter, 2011, p. 195), out of his comparison on the ground that the Canadian experience with immigration started much earlier and was approached with more selective immigration policies. Yet if the particular immigration history and policy of a country is really crucial in allowing or preventing a progressive dilemma, the Canadian case serves as a clear counter-example to the anti-multicultural hypothesis

3) One of few exceptions to this is the research conducted by Banting et al. (2006), which found that those Western countries where the immigrant population was growing most rapidly also experienced the smallest growth rates in social spending. However, their study found no significant impact of multicultural policies on the erosion of Western welfare states (ibid., p. 83).

– which, after all, claims that multiculturalism systematically, regardless of a specific historical and geographical context, engenders pressure on the welfare state (Banting & Kymlicka, 2006, p. 31). Notably, the progressive dilemma discourse has not had a significant impact in Canada (Banting, 2010), indicating that it is not multiculturalism as such which undermines the foundations of the welfare state. Rather, a country's particular history of immigration and ethnic diversity, as well as its policies in other domains must be taken into account. This brings to the fore the limitations of the cross-national, comparative research discussed above: while it is an appropriate way of identifying the inherent effect of diversity on Western welfare states (or lack thereof), it insufficiently considers the influence of the specific context of individual countries (Banting & Kymlicka, 2006, p. 33) – to account for the latter, in-depth case studies are needed.

WELFARE CHAUVINISM THE U.S. AND GREAT BRITAIN

In the literature on the effect of diversity on the welfare state, the cases of the U.S. and Great Britain have received particular attention. To start with the first, the residual, 'bare-bones' character of the American welfare state has repeatedly been linked to its ethnic fragmentation (Alesina & Glaeser, 2004; Putnam, 2007). One of the more significant publications in this tradition, Putnam's (2007) influential article indicated that people living in ethnically diverse neighborhoods in the U.S. were less trusting of their neighbors and less supportive of redistributive policies, implying that diversity erodes the solidarity required for the welfare state. Similarly, the aforementioned study by Alesina and Glaeser (2004), which compared the U.S. welfare state to those of European countries, concluded that the greater ethnic diversity of the former explained a large part of its divergence from the latter. They argue that American anti-welfare demagogues have been able to successfully exploit ethnic divisions to depict the disadvantaged (particularly African-Americans) as lazy and undeserving of support, whereas their European counterparts, faced with relative ethnic homogeneity, had no such divisions to resort to (ibid., p. 180-1). Nonetheless, they predict that with increasing ethnic heterogeneity in many European countries, the same tensions that prevented the growth of the welfare state in the U.S. will emerge in Europe too, as reflected in the recent rise of the radical right wing (ibid., p. 182). Examining this prediction, Larsen (2011, p. 350-1) observes that the same in-group—out-group dynamics have already permeated European perceptions of non-Western immigrants, who are consequently seen as less deserving of welfare provisions.

It is perhaps unsurprising, then, that a similar literature has developed on the British experience – the second case to receive considerable scholarly scrutiny. Here, the 'American example' raised questions about the sustainability of the welfare state in the face of growing ethnic heterogeneity, and the efficacy of past, multicultural policy approaches (Goodhart, 2004; Klausen, 2000). Whereas British multiculturalists such as Parekh (2001, p. 695-6), in his famous Parekh Report, had insisted on the rights of minorities to maintain their own culture so as to not be treated as inferior to the dominant British culture, this position became fiercely criticized by conservatives and liberals alike for being hostile towards the British nation, national identity and institutions (ibid., p. 698-9; see also Barry, 2001, p. 89). For instance, Klausen (2000) condemns Parekh's disavowal of a singular, more or less enforced British identity for conflicting with the foundations of the welfare state. She claims that "what Marshall and Beveridge[4] saw as a precondition for a properly functioning welfare state," presumably shared values and norms, "Parekh views as domination," (ibid., para. 12). Similarly, Goodhart (2013, p. 296) faults British multiculturalism for having insufficiently inculcated minority groups with a unified national identity, which he argues has translated to a society composed of segregated groups that lack the mutual trust necessary to maintain a modern welfare state. Overall, the multicultural approach of the past advocated by Parekh (2001) was widely denounced as a mistake with potentially far-reaching consequences for British society and its institutions (McLaughlin and Neal, 2004).

Nonetheless, the question remains whether one can criticize multiculturalism as such for having an inherent propensity for the erosion of social cohesion, or if this criticism is concerned with a particular interpretation or historical application of multiculturalism. From Goodhart's (2013, p. 178-98) analysis of the development of British multicultural policies in the last few decades, it appears the latter might be the case. His argument echoes that of Barry (2001, p. 88), who asserts that the multicultural society where "groups live in parallel universes is not one well calculated to advance mutual understanding or encourage the cultivation of [...] sentiments of trust." It is this alleged segregationist nature of multicultural policies, their supposed separatist tendencies rather than their anti-discriminatory aspects, which Goodhart (2013, p. 178) finds fault with. According to Goodhart, British multiculturalists in the 1980s increasingly moved from encouraging anti-discrimination regulations to the creation of separate ethnic organizations, associations and education, heralding a new era of separatist multiculturalism which lasted well into the 1990s (ibid., p. 183-4). As a result, he argues, immigrants may have entered Great Britain expecting to have to accommodate to their new surroundings, but were subsequently discouraged from doing so by policies which emphasized their ethnic identity over a national one (ibid., p. 190). The multicultural focus on maintaining ethnic identity thus led to a segregated British society where each ethnic group is only concerned with its own welfare, at the cost of solidarity with others (ibid., p. 272). In Barry's (2001, p. 325) words, multicultural policies may have helped particular ethnic groups to establish rights shielding them from oppression, yet "this kind of particularistic focus will [...] make cultural minorities weak partners in endeavours to redistribute income from rich to poor across the board." Although these segregationist mechanisms may well emerge in other national settings, critics such as Goodhart (2004), Klausen (2000) and Barry (2001) clearly target the particular interpretation of multiculturalism that developed in Great Britain in the 1980s and 1990s.

APPLYING THE U.S. AND U.K. MODELS

Nonetheless, since this approach to multicultural integration policies was hardly unique to Great Britain (Koopmans, 2010, p. 6), one might wonder to what extent the same criticism could be applicable to other countries. As noted in the review of cross-national research, there is little evidence for a general tendency within multiculturalism to undermine the welfare state, yet certain parallels have been identified between the British and American cases and other European countries (e.g. Eger, 2010; Van der Waal, Achterberg, Houtman, De Koster, & Manevska, 2010; Van Oorschot, 2006). While these case studies have covered a wide range of countries, the scope of this project allows only for a general overview of their findings.

On a more general level, two important conclusions can be derived from the numerous national-level studies on the progressive dilemma. Firstly, it has been questioned whether a country's welfare regime (Sainsbury, 2006, p. 230-1) impacts the possibility for a progressive dilemma (Larsen, 2011). With regard to this issue, the case studies show that even though the more generous welfare regimes of Scandinavian countries witness more support for redistribution among their populations (ibid., 351), the progressive dilemma cuts across the boundaries of welfare regimes (Banting, 2010, p. 812). That is, ethnic diversity affects support for marginal welfare states such as the U.S. (Putnam, 2007) as well as their more generous counterparts – e.g. the Netherlands (Van der Waal et al., 2010) or Sweden (Eger, 2010). Even so, universal benefits (i.e. those provisions to which every citizen has access) are less susceptible to questions of who is truly deserving of support than those benefits targeting particular groups (Larsen, 2008). Nevertheless, the level and type of welfare provisions a state offers are not crucial in preventing or causing a progressive dilemma (Banting, 2010, p. 802).

Secondly, and contrary to what the first conclusion might suggest, the progressive dilemma is not wholly inevitable. Although increasing ethnic diversity seems to have affected, or is at least perceived as a problem for virtually all Western welfare states (Freeman, 2009), Canada has been a notable exception to this trend (Banting, 2010). This has been partly attributed to its emphasis on contributory welfare provisions, to which potential recipients have to contribute before they can receive the benefit; since the recipients of such a provision have helped finance it themselves, they are unlikely to be regarded as undeserving users (ibid., p. 813). Furthermore, Canadian immigration policies have traditionally given priority to those immigrants most likely to be able to sustain themselves economically (ibid., p. 806). Interestingly, Banting argues that Canadian multiculturalism has been an integrative force, instilling in the ethnically diverse Canadian population an inclusive national identity which reinforced social cohesion (ibid., p. 809-10).

4) Klausen (2000) refers to T. H. Marshall (1950/2010), who initiated the tradition of 'social rights', and William Beveridge (1942), author of the highly influential Beveridge report widely seen as the start of British welfare policies, here posited as the 'founding fathers' of the British welfare state.

Even though the Canadian application of multicultural ideals has been markedly different from that in many European countries, the Canadian case does suggest that a progressive dilemma can be avoided through the implementation of particular policies (as will be briefly discussed in the final section), and that multiculturalism need not be in conflict with the welfare state. However, a theoretical understanding of welfare chauvinism is necessary to explain why the issue has emerged in most welfare states, but not in all.

EXPLAINING WELFARE CHAUVINISM

While studies on the progressive dilemma generally support the link between ethnic heterogeneity and welfare chauvinism (Burgoon et al., 2012; Mau & Burkhardt, 2009; Nannestad, 2007; Senik et al., 2009), a plethora of explanations have been put forward as to why this might be the case. Studies have pointed to numerous diversity-related factors of lesser and greater impact on the erosion of support for the welfare state (e.g. Tolsma et al., 2009; Van der Waal et al., 2010), but nonetheless vary greatly in their overarching theoretical explanations. Freeman (2009) and Banting (2010) both outline three potential approaches to understanding welfare chauvinism: firstly, a continuation of Allport's (1954/1979) work on prejudice, focusing on out-group discrimination; secondly, the theory of reciprocal altruism, which emphasizes rational choice; and finally, a neo-Darwinian model which emerged from evolutionary biology. This section will propose Allport's theory on prejudice and intergroup contact as the most fruitful model for understanding the dynamics of interethnic antagonism and its expression in welfare chauvinism as well as potential solutions. To this end, Allport's theory will be critically reviewed and subsequently applied to the case of welfare chauvinism. Concluding this section, the merits of using this framework will then be discussed in relation to the two alternative paradigms.

INTERGROUP CONTACT THEORY

The oldest of the three models, Allport's (1954/1979) intergroup contact hypothesis has been subject to extensive empirical testing ever since its conception. In fact, Allport's contentions on the effects of interactions between different (ethnic) groups were themselves rooted in empirical studies of such contact. While Allport is generally credited with the articulation of the famous hypothesis, the first inklings of this theory can already be identified in earlier psychological studies (Dovidio, Gaertner, & Kawakami, 2003, p. 6). Under the dominance of Social Darwinism, which lasted well into the 1930s, scholars had almost exclusively framed contact between different groups in a context of intergroup competition and conflict (Pettigrew & Tropp, 2005, p. 262). Questioning the inevitability of intergroup conflict, studies in the 1930s started noting the positive potential of contact between different groups; nevertheless, scholars were unable to explain the contrasting effects of intergroup contact, conflict-inducing or prejudice-reducing, let alone predict which results one could expect in a given situation (Dovidio et al., 2003, p. 6). More systematic analyses of interethnic interactions in the 1940s presented the first considerations of situational aspects in determining the outcome of intergroup contact. For instance, positing prejudice as the primary source of interethnic conflict, Taba and Wilson (1946) stress the authoritative role of schools in de-emphasizing intergroup difference and farming interaction as a strategy towards mutual understanding. In the same volume, Brameld and Fish (1946) note the adverse effects of policies of segregation as a means of diminishing negative intergroup attitudes and conflict in schools, and underline the importance of the equal representation of ethnic groups in all strata of the organization of these schools.

As Katz (1991) suggests, it is important to take into account the academic climate in which Allport formulated his thesis about intergroup contact, as it provides highly significant insights into the development of the hypothesis as well as the role Allport assigns to prejudice. Contrary to the Social Darwinist focus on insurmountable biological difference between groups as a basis for violent competition, Allport (1954/1979, p. xv) attributes the "endless antagonism" between ethnic groups not to "a realistic conflict of interests" but instead to "fears of the imagination." Importantly, this prejudice emerges along the lines of a distinction between one's in-group and out-group (Brown & Zagefka, 2005). According to Allport (1954/1979, p. 20-3), people depend on generalized categories for judging everyday situations. From this manner of thinking, categorizing a diversity of elements into more or less clean-cut groups, stems the ascription of certain group identities to individuals who are thereby assigned to the in-group or out-group (Brown & Zagefka, 2005).

Although the exact dynamics of this process are beyond the scope of this brief overview, it is important to note that the distinction between the in-group and out-group relies on subjective perceptions of intergroup difference more than actual objective difference (ibid., p. 55).[5] In this regard, group membership should be viewed as situational rather than fixed, and while groups may define themselves hostilely in opposition to an 'enemy' out-group, this is not the inevitable outcome of real, irreconcilable difference. Essentially, the in-group serves to provide its members with security (Allport, 1954/1979, p. 42); hostility towards the out-group may be one strategy to reinforce group boundaries, but it is most certainly not the only way of attaining this goal. As Brewer (1999, p. 430) also contends, harboring positive feelings toward the in-group need not imply hostile attitudes toward the out-group.

REDUCING PREJUDICE

Prejudice against the out-group, although originating in intergroup distinctions, is therefore far from irresolvable (Gaertner & Dovidio, 2005, p. 72). Continuing the trend towards a systematic consideration of the context in which intergroup contact occurs in order to explain its contrasting effects, Allport (1954/1979) identified a number of conditions which, if all met, would constitute the necessary environment for interethnic interaction to have a positive effect. Concluding his chapter on the effect of intergroup contact, Allport states:

> *Prejudice [...] may be reduced by equal status contact between majority and minority groups in the pursuit of common goals. The effect is greatly enhanced if this contact is sanctioned by institutional support (i.e. by law, custom or local atmosphere), and provided it is of a sort that leads to the perception of common interests and common humanity between members of the two groups. (ibid., p. 281)*

In order to reduce prejudice, the environment where the interaction takes place must thus meet the following requirements: firstly, the groups interacting must be more or less equal in status in the particular situation; secondly, the interaction must be directed towards common objectives; thirdly, the contact must be cooperative rather than competitive; fourthly, the interaction must enjoy authoritative support (Pettigrew, 1998).

Despite the fact that, as Pettigrew and Tropp (2005) observe, subsequent studies have generally confirmed the significance of these conditions[6], thorough empirical testing of the intergroup contact hypothesis has led many to propose considerable refinement of the thesis. Koschate and Van Dick (2011, p. 771-2), for instance, note that simply having a common goal is not sufficient for successful interaction between groups, and that such goals must be "positively interdependent"; that is, the attainment of the goal must be dependent on the success of each group involved, so that the interaction is necessarily cooperative in nature instead of competitive. More ambitiously perhaps, studies such as those conducted by Cameron and Rutland (2006), and Herek and Capitanio (1996) extend the intergroup contact theory to the fields of disability and sexuality, respectively. In doing so, they demonstrate that the model is not strictly limited to groups differing only in terms of race and ethnicity (Pettigrew & Tropp, 2006).

Others have highlighted the character of intergroup contact as essential to its outcome. Pettigrew (1998) asserts that high-quality cross-group contact in the form of friendships generally guarantees the long-term, cooperative pursuit of common goals and as such is more significant than the initial acquaintance with outsiders. Moreover, a quantitative analysis of seven surveys, including one conducted in the Netherlands, shows a strong positive correlation between intergroup friendships and support for pro-outsider policies (Pettigrew, 1997, p. 179). The central role Pettigrew (1998, p. 76) assigns to the potential development of cross-group friendships is perhaps best illustrated by the fact that he elevates this factor to a fifth condition to be added to Allport's (1954/1979) original four – a revision also advocated by Dovidio, Gaertner and Kawakami (2003, p. 8). In this way, Pettigrew (1997, 1998) and Dovidio et al.

5) See also Barth's (1969) emphasis on ethnic identity as a process of boundary maintenance, whereby some cultural differences are accentuated and others suppressed depending on the context, rather than as bearer of a stable common culture.

6) See Pettigrew and Tropp (2005) and Kenworthy, Turner, Hewstone, and Voci (2005) for an elaborate overview of the academic reception of the intergroup contact hypothesis.

(2003) underscore the effect of the affective aspect of intergroup contact in addition to its merely cognitive implications (i.e. increased knowledge of the other group as a result of the interaction).

The Cameron and Rutland (2006, p. 470) study, which evaluated the effect of reading stories about disability on prejudice among children, also exemplifies a call for broadening the intergroup contact model, as it rejects the traditional primacy of direct interaction and emphasizes the effect of vicarious, indirect contact. The study represents a wider call for intergroup contact theory to abandon its singular focus on interpersonal contact and to take into account the effects of indirect interaction; in fact, direct intergroup contact can result in anxiety (Stephan & Stephan, 1985), which in turn may prevent the reduction of, or even affirm existing prejudice (Wright, Aron, McLaughlin-Volpe & Ropp, 1997, p. 74). Intergroup friendships as described by Pettigrew (1997, 1998) remain important, yet awareness of the existence of such relationships among members of either group can already diminish prejudice while avoiding potential anxiety (Wright et al., 1997). In other words, it is not necessary for all group members to engage in cross-group relationships in order to reduce intergroup antagonism, but it is vital that such contact does occur. For this reason, Dovidio et al. (2003, p. 8) add a sixth condition, "the opportunity for personal acquaintance between the members" – that is, at least some group members must be able to engage in cross-group interaction.

UNDERSTANDING WELFARE CHAUVINISM AND DESERVINGNESS

Seen as Allport's theory has often been applied to antagonistic intergroup relations in multi-racial, multi-ethnic societies (Pettigrew & Tropp, 2006), it has clear potential for contributing to an understanding of the processes underlying welfare chauvinism. While research in ethnically heterogeneous Western states has repeatedly observed a negative correlation between ethnic diversity and social spending or support for redistributive measures (Burgoon et al., 2012; Mau & Burkhardt, 2009; Senik et al., 2009), the intergroup contact theory allows for an analysis of the mechanisms by which the societal divisions underlying these issues are constituted in terms of in-group—out-group distinctions. After all, the immigrants perceived to constitute a threat to the 'native' welfare state tend to be of non-Western origin (Mau & Burkhardt, 2009) and differ therefore from the dominant group in several significant ways – ethnically, economically, socially; as Knigge (1998, p. 258) aptly states, immigrants "constitute an out-group par excellence [emphasis in original]: they are weak, vulnerable and powerless." While Allport (1954/1979) emphasized that intergroup conflict is not a necessary outcome of identification with a particular group, intergroup contact theory also provides a comprehensive framework for analyzing how such conflict may arise.

In the context of welfare chauvinism, it seems in-group favoritism and hostility towards the out-group are strongly intertwined. As Brewer (1999, p. 430) notes, it is often assumed that one necessarily implies the other, yet she asserts that they should be considered separately. While people are more likely to feel sympathetic vis-à-vis members of their in-group (those with whom they identify), they may well be mildly sympathetic or indifferent towards out-groups unless these constitute some sort of threat to the welfare of the in-group (ibid., 2001, p. 27). It is under such conditions, where the presence of an out-group is perceived as conflicting with the needs and maintenance of the in-group, that the out-group may evoke stronger feelings of resentment and even hatred. Especially when two groups compete for the same, limited resources, this presents an incendiary situation for intergroup conflict (ibid., p. 28). The heightened potential for intergroup antagonism in competitive situations might also explain why the progressive dilemma has emerged in so many Western welfare states, seeing as different ethnic groups can be said to compete here for the redistribution of a limited pool of wealth.

To some extent, such situations seem inevitable. In modern societies, Brewer (1999) points out, people always have to rely at least partly on others for their subsistence, by sharing resources, information or other types of aid. A person's survival thus becomes dependent on the willingness of others to reciprocate whatever they choose to share, making interpersonal trust an essential aspect of modern life (ibid., p. 433). One can however not be trusting of just anyone, as this would expose him or her to easy exploitation, and it is therefore safer to share resources only with those one recognizes as members of the same in-group (ibid., 2001, p. 29). In order to do so, groups must be clearly delineated and members must be marked so as to signal their belonging; processes of "assimilation within and differentiation between groups" serve to prevent "ingroup benefits [from being] inadvertently extended to outgroup members, and to ensure that ingroup members will recognize one's own entitlement to receive benefits," (ibid., 1999, p. 433-4). The degree to which one person identifies with another thus becomes an important determinant for their willingness to support each other.

The importance of such identification is translated into the way members of the out-group come to be seen as undeserving of support, as Van Oorschot's (2006) five criteria for deservingness illustrate. Drawing from a range of studies on perceptions of welfare recipients, Van Oorschot outlines five conditions which influence the perception of a welfare beneficiary as deserving or undeserving: "control over neediness", "level of need", "reciprocity", "identity", and "attitude" (ibid., p. 26). While the first two conditions, the degree to which a person can be held accountable for their need and the extent of their need[7], may apply to members of the in-group and out-group alike, the other three criteria underscore Brewer's (1999, 2001) account of group differentiation. The third criterion is that of reciprocity (Van Oorschot, 2006, p. 26); as already discussed above, this is an important ground on which to exclude members of the out-group from in-group benefits, and has incited the construction of clear group boundaries to avoid non-mutual support as much as possible (Brewer, 2001, p. 29). The fourth and fifth factors for determining deservingness, identity and attitude, refer to ways to signal such boundaries and identify members of the out-group so as to exclude them from in-group welfare. Echoing Brewer's argument (ibid., p. 28-9), Van Oorschot (2006, p. 26) asserts that people consider those who appear similar to themselves, and who act according to their own standards of behavior, as more deserving of support. In other words, people determine the deservingness of a welfare recipient by their membership in the in-group or out-group, as signaled by 'identity', meaning physical markers of difference, and 'attitude', referring to behavioral markers. This goes a long way towards explaining the possibility for welfare chauvinism in ethnically diverse societies, as it suggests that particularly non-Western immigrants, who can be expected to differ most from the dominant group in terms of their 'identity' and 'attitude', tend to be categorized as members of the out-group and therefore undeserving of 'native' (i.e. in-group) welfare provisions.

That is not to say, however, that the presence of ethnic 'others' in society will inevitably lead to an exclusionary situation dominated by ethnic in-group loyalties. Allport's (1954/1979, p. 43) view is that hypothetically, the in-group can be expanded indefinitely to include and extend loyalty and trust to members of former out-groups – yet he acknowledges that there are practical limits to this process. Although Allport (ibid., p. 43-6) provides few indications as to when and why the in-group ceases to expand, Brewer's (1991) concept of optimal distinctiveness helps pinpoint where the limits of inclusiveness lie. According to Brewer, people have the two contrasting needs to belong to a group on the one hand, and to be distinct on the other; whenever a group identity becomes too inclusive, it no longer satisfies one's need to differentiate oneself from others (ibid., p. 478). In an overly inclusive in-group, members are also less willing to protect the welfare of the in-group and will more likely act on personal interest (ibid., p. 479). This implies that smaller groups are more effective in ensuring trust and loyalty among their members (ibid., 1999, p. 434). Nevertheless, such groups can be subsumed under larger superordinate identities, a phenomenon Allport (1954/1979, p. 44) calls "concentric loyalties": people can 'belong' to a smaller sub-group, such as an ethnic group, within an overarching in-group, such as a national identity. In order for this to work and foster positive relations between sub-groups, the situation must abide by the same conditions for positive intergroup contact identified above. Particularly, the respective interests of the sub-groups must be seen as mutual rather than conflicting, and their interaction has to be structured under a common objective which must be attained cooperatively rather than competitively (Brewer, 1999, p. 434). In practice, however, implementing the notion of concentric loyalties in Western welfare states may prove an arduous task. As Allport (1954/1979, p. 44-5) also notes, the construction of non-antagonistic intergroup relations under overarching identities may be a definite possibility, but is still remains a long and complicated process.

7) According to Van Oorschot (2006, p. 26), those who are regarded as not personally accountable for their situation of need, and who also have the greatest need, tend to be seen as most deserving of support.

ALTERNATIVE APPROACHES

Before the intergroup contact approach can be applied to a concrete case of welfare chauvinism, the two alternative theoretical approaches and their shortcomings relative to intergroup contact theory must be addressed. Firstly, Alesina, Glaeser and Sacerdote (2001) have proposed the Trivers' (1971) theory of reciprocal altruism as a way to analyze welfare chauvinism by focusing on the idea that people are unwilling to support those who are unlikely to return this favor. This means that perceptions of beneficiaries as 'welfare leeches' who are exploiting the system will lead people to become unsupportive of the welfare state (Alesina et al., 2001, 237). As Van Oorschot (2006) indicated, reciprocity is indeed an important factor in determining whether a potential recipient is deserving of support – but it is not the only condition. Alesina et al.'s (2001, p. 227) own example of African Americans demonstrates that attitudes of welfare chauvinism are not so much based on a rational consideration of the risk that given support may not be reciprocated, but rather on the stereotypical, often racialized perceptions of welfare recipients as lazy. In other words, welfare chauvinism is informed by prejudice; what is crucial is one's perception of the probability of reciprocity rather than its actual likelihood. Unlike the intergroup contact framework, the model of reciprocal altruism fails to explain the emergence of such prejudicial views (ibid., p. 246).

Secondly, Salter (2002) has claimed that the explanation for welfare chauvinism should be sought in genetic interests. In Salter's (ibid., p. 111) view, people may attempt to protect their own gene pool by engaging in altruistic behavior towards their own ethnic group when faced with large inflows of migrants who form a rapidly growing population. In the context of multi-ethnic welfare states, the altruism of the ethnic majority – in the shape of income transfers – may also benefit ethnic minority groups, at least to the extent that those are eligible for the same welfare provisions. Since this welfare system is therefore not entirely in the 'genetic interest' of the dominant group, this may lead some to engage in ethnic nepotism, supporting only those of their own genetic background, and attempting to exclude others (ibid., p. 132). Again, this theory is in agreement with Van Oorschot's (2006) concept of deservingness, which takes into consideration physical characteristics ('identity') as a marker of difference and therefore un-deservingness. Nevertheless, the same criticism made of reciprocal altruism can be applied here: ethnic nepotism theory concentrates on only one aspect of the wider phenomenon of welfare chauvinism. As shown by Sniderman, Peri, De Figuereido, and Piazza (2000, p. 53) a larger genetic difference does not always translate to a greater perceived threat: for ethnic Italians, the arrival of Eastern European and African immigrants appeared equally threatening. Again, the vital aspect seems to be one's perception of a threat, which is influenced by a combination of factors (Van Oorschot, 2006, p. 26) instead of an actual genetic threat immigrants may constitute. Finally, by explaining the exclusion of minorities on the basis of genetic factors, Salter's (2002) theory risks representing intergroup antagonism as inevitable and 'natural'. Intergroup contact theory avoids this pitfall by emphasizing group identities as based on subjectively perceived rather than objective difference (Brown and Zagefka, 2006, p. 55).

Clearly, the theories of reciprocal altruism and ethnic nepotism touch on important factors in welfare chauvinism. Yet even though both theories make valuable contributions to the understanding of welfare chauvinism, both center overly on particular aspects of this issue, and fail to include the wider range of factors influencing support for the welfare state and perceptions of welfare beneficiaries. In this regard, intergroup contact theory is a more flexible framework to analyze the emergence of a progressive dilemma. Rather than portraying intergroup conflict as unavoidable, Allport's (1954/1979) theory focuses on intergroup relations as constructed and contingent on a wide array of factors, and may therefore also be better equipped to draw attention to possible solutions for the issue of welfare chauvinism.

WELFARE CHAUVINISM IN THE NETHERLANDS

While the intergroup contact theory has repeatedly been suggested as a potential way of understanding tensions between different ethnic groups in the welfare state (Alesina et al., 2001; Banting, 2010; Freeman, 2009), surprisingly few have systematically applied this theory to the issue of welfare chauvinism. As the Netherlands has recently experienced a radical shift towards anti-immigrant discourse, this appears to be an interesting case for further study. Studies suggest the Netherlands has been no exception to the general rise of the progressive dilemma in Western welfare states (Banting, 2010; Koning, 2013; Van der Waal et al., 2010); nonetheless, very few have studied the Dutch case specifically. In this section, the Dutch case will be examined from the intergroup contact perspective to explain emerging sentiments of welfare chauvinism in this context. Following an overview of the development of a progressive dilemma in the Netherlands, the recent rise of this issue will be analyzed using the tools of intergroup contact theory. Finally, potential solutions to the issue of welfare chauvinism will be considered.

As has often been noted, Dutch immigration discourse has undergone a rather extreme transition over the last few decades from a fairly lenient multiculturalism[8] to a sterner assimilationist discourse (Entzinger, 2006; Vasta, 2007). Even though other European countries have witnessed a similar harshening of the immigration discourse, the Netherlands has been noted for its particularly reproachful tone towards immigrants, who are often accused of exploiting the welfare system (Koning, 2013, p. ii). This reflects strong views among the Dutch public that (non-Western[9]) immigrants contribute little to the economy and should first and foremost adapt to Dutch society (Banting, 2010, p. 804). Considering the tremendous harshness which has recently characterized Dutch immigration discourse, it is perhaps striking how long it has taken for this anti-immigrant discourse to emerge. While unemployment among immigrants in the Netherlands had been a persistent problem throughout the 1980s and 1990s, and despite growing concern about the use of welfare provisions by immigrants, for a long time it remained politically incorrect to speak about these issues openly (Entzinger, 2006, p. 182). One of the first to discuss this problem openly was Paul Scheffer, a well-known member of the Dutch Labour Party, who criticized Dutch multicultural policies for having been to lenient in requiring immigrants to integrate (ibid., p. 184). Besides contributing to the poor socio-economic position of immigrants, their lack cultural integration, Scheffer (2000) argued, would endanger social cohesion in Dutch society; to combat this threat, the Dutch would have to form and enforce a stronger sense of national identity. Similar to Goodhart's (2004) criticism of British multiculturalism, Scheffer (2000) thus denounced Dutch immigration policies for having engendered a profound societal segregation, and advocated more unifying integration policies.

Although it was not until the early 2000s that this debate came into fruition in the Netherlands, the roots for the segregation of immigrant groups, so argue critics, had been laid down long before (Entzinger, 2006). As mentioned before, Dutch policies in the 1980s were directed towards the maintenance of minority cultures, paid little attention to cultural integration or Dutch language acquisition, and refrained from enforcing a unified Dutch national identity (Vasta, 2007, p. 717). Similar to 'separatist multiculturalism' in Great Britain (Goodhart, 2013, p. 190) immigrants in the Netherlands were encouraged to organize themselves within their own ethnic groups rather than in Dutch society as a whole (Duyvendak & Scholten, 2011, p. 337). This approach was a product of the Dutch history of pillarization, during which members of the four largest societal groups lived largely separate lives, each within their own 'pillar' (Ghorashi, 2010, p. 77). Although these pillars had already started to crumble in the 1960s and 1970s, this tradition still greatly influenced the way the Dutch government thought it should accommodate the newly arrived minority groups, by promoting among immigrants the maintenance of their culture and establishment of their own organizations (Duyvendak & Scholten, 2011, p. 339). Moreover, the experience of pillarization, which was characterized by strongly delineated cultural boundaries, created a 'pillarization habitus'; that is, it led the Dutch to essentialize cultural difference and to perceive immigrants solely in terms of their ethnic identities (Ghorashi, 2010, p. 78).

The Dutch government changed its position with regard to immigrant groups when the economic

8) It should be pointed out that, despite the general labeling of Dutch integration policies as multicultural, this has not gone undisputed. For instance, Duyvendak and Scholten (2011) argue that Dutch immigration policies have been characterized by multiple different discourses

9) Much of the literature used in this section addresses more recent immigrant groups, particularly the Turkish and Moroccan guest workers. Since this case study focuses on developments from the 1980s onwards, one can assume that the literature is largely based on the most prominent immigrant groups of that time (i.e. the guest workers and their families) – yet many studies leave this implicit.

restructuring of the 1980s saw the sectors where many guest workers had been working disappear, which left large numbers of immigrants unemployed (Snel & Scholten, 2005). These developments led to an overwhelming political consensus that multicultural policies had failed, causing a lack of integration of immigrants into Dutch society (Duyvendak & Scholten, p. 339). In response, integration policies were implemented in the 1990s to ensure that immigrants would be included in 'mainstream' institutions rather than remain in their ethnic minority equivalents (Vasta, 2007, p. 717). Meanwhile, as the societal marginalization of immigrants persisted, their ethnic minority identities became inextricably linked to a low socio-economic status, reinforcing perceptions of immigrants as fundamentally different (Eijberts, 2013, p. 124). The poor integration of especially Muslim immigrants in Dutch society was also increasingly blamed on the minority culture they had previously been encouraged to retain (Bevelander & Groeneveld, 2012); it was therefore concluded that integration policies should take a more assimilationist course by insisting on the acquisition of Dutch language and culture (Vasta, 2007, p. 718-9).

The retreat of multiculturalism in the 1990s coincided with a considerable retrenchment of the welfare state in the Netherlands, which might lead one to suspect that the combination of ethnic diversity, multicultural policies and a strong welfare state simply proved unsustainable (Entzinger, 2006). Although cuts were indeed made in welfare provisions which sustained a lot of immigrants, it should be pointed out that this was done as part of of a broader restructuring of the Dutch welfare state to encourage labor market participation (ibid., p. 198-200). Rather than attempting to curtail the use of welfare provisions by immigrants in particular, these welfare state reforms targeted all societal groups (Yerkes & Van der Veen, 2011). In fact, the retrenchment project mainly targeted the lenient Dutch disability benefits, which had for a long time been exploited by employers as a means to lay off employees without having to abide by strict regulations (ibid., p. 433).

However, more recent attempts to cut back on welfare provisions have targeted immigrants specifically, indicating that welfare chauvinism is starting to have an impact on Dutch welfare policies (Koning, 2013, p. 94-104). One example is the 2012 reform of eligibility requirements, whereby the initial phase during which new immigrants are ineligible for benefits was prolonged by two years to a total of five years; this means immigrants have no claim to welfare support for this period unless they can prove they have a 'durable link' to the country (ibid., p. 96). While it is left to the discretion of the judicial branch to determine what constitutes such a 'durable link', it is clear that the term refers to some demonstrable loyalty to the Netherlands (ibid., p. 97). In addition to delaying immigrants' access to welfare provisions, plans have also been proposed to make their eligibility for benefits contingent on their successful integration (ibid., p. 102). A clear example of this is the coalition agreement of the Rutte-Verhagen cabinet (Government of the Netherlands, 2010, p. 26), which outlines a proposal to curtail the access to unemployment benefits for those immigrants who "through behavior or dress effectively limit their chances on the labor market." Meanwhile, immigrants are also held gradually more accountable for their own integration; as mentioned before, a failure to integrate tends to be blamed on the immigrants themselves or their culture (Eijberts, 2013, p. 124). In recent years, integration policies have also increasingly stressed the immigrant's financial responsibility for the cost of civics and language courses (ibid., p. 133-4; Government of the Netherlands, 2010, p. 26). Overall, this brief overview – though far from exhaustive[10] – seem to reflect a view of immigrants as undeserving of welfare support, a growing prominence of integration as a condition for welfare eligibility, and an increasing emphasis on the responsibility of immigrants for their own integration into Dutch society.

AN INTERGROUP CONTACT ANALYSIS

When examined from the perspective of intergroup contact theory, it is perhaps unsurprising that the Dutch approach to immigration and ethnic diversity has resulted in intergroup antagonism. First of all, the separatist multiculturalism of the 1980s, informed by a history of pillarization, resulted in policies which encouraged the grouping of newly arrived ethnic minorities into a separate 'pillar', directly engendering social, physical, and cultural segregation from the rest of Dutch society (Koopmans, 2003). By reducing the likelihood of encounters between members of different ethnic groups, this pillarization-inspired model of a multicultural society can be said to have reduced the potential for positive intergroup contact. To examine this further, one must look at the conditions for prejudice reduction listed by Allport (1954/1979) and Pettigrew (1998): equal status, common objectives, cooperation, support from authorities, and cross-group friendship.

When considering the position of immigrant groups in the context of Dutch society, equal status in the interaction with the ethnic majority is unlikely; not only do these groups tend to occupy a lower socio-economic position bound to a negatively perceived ethnic identity (Eijberts, 2013, p. 124), they also have a disproportionate obligation to adapt to the majority group – as witnessed in policies which consider integration the (moral and financial) responsibility of the immigrant alone (ibid., p. 133-4). As for common objectives and cooperation, one could argue that these were prevented by separatist policies which left the immigrant groups to their own devices. By encouraging the establishment of 'ethno-specific organizations' (Vasta, 2007, p. 717), ethnic minority groups were incited to set their own objectives and obtain them within their own 'pillar', rather than in collaboration with other groups. With regard to institutional support for intergroup contact, it is clear that the governmental policies of the 1980s at least indirectly discouraged such interaction in the first place, as they instead encouraged a segregated social life in an ethnic group's own organizations (Koopmans, 2003). Finally, then, the likelihood for the development of intergroup friendships was slim, as social life was to take place within the immigrants' own pillar. While these conditions may in practice have applied differently to different immigrant groups – for instance, some may have had a less stigmatized socio-economic and ethnic identity than others – this analysis demonstrates that Dutch multicultural policies did little to promote positive intergroup contact. Although it cannot be said that they necessarily caused intergroup antagonism, it seems these policies made positive interaction more difficult; as the multicultural policies promoted the separation of immigrant groups from the ethnic majority rather than inter-ethnic interaction, they cut off both the cognitive (i.e., knowledge-increasing) as well as the affective (i.e., friendship-producing) pathway to prejudice reduction. Seen from this perspective, it seems 1980s Dutch multiculturalism made intergroup prejudice and antagonism more likely, by discouraging or complicating intensive and positive intergroup contact.

In the context of welfare chauvinism, one might expect the absence of intergroup contact to lead to very low levels of cross-group identification. This situation may have been exacerbated by pressures on Dutch identity from the part of globalization and growing immigrant populations (Eijberts, 2013, p. 129). In other words, the 'native' Dutch may have felt hostility towards the immigrant groups from the first instant, as they were perceived to represent a threat to their in-group identity. From the perspective of optimal distinctiveness (Brewer, 1991), the Dutch may have been unlikely to approach the newly arrived immigrant groups in a very inclusive manner, as an over-inclusive Dutch identity would have no longer served the need for differentiation. In addition, the 'pillarization habitus' (Ghorashi, 2010) may have led the native Dutch to regard immigrants in an essentialist way, as fundamentally different and therefore unfit for inclusion in Dutch national identity. At the same time, as Savelkoul, Scheepers, Tolsma, and Hagendoorn (2011, p. 744) suggest, the expanding immigrant out-group may have been perceived as an increasingly powerful interest group and thus a threat in the societal competition for scarce resources. Each of these pressures – the drive for distinctiveness, a pillarization habitus, and increasing competition – contributes to an explanation of the way the Dutch majority received immigrant groups; each of these accounts provide complementary factors which indicate that the Dutch may have been likely to exclude and feel hostile towards the immigrant groups from the outset (Savelkoul et al., 2011, p. 752).

Taken together, the unlikelihood of positive intergroup contact and the likelihood of the hostility towards ethnic minorities in the Dutch context made for an environment where welfare chauvinism could easily arise. Since immigrants were encouraged to live segregated from the Dutch majority group, and because the Dutch regarded the ethnically different immigrants in an essentialized way (Ghorashi, 2010, p. 77), the competition for resources in Dutch society (i.e. the redistribution of income) was likely to be drawn along ethnic lines. In this context, one would expect particularly non-Western, 'Muslim' immigrants, who

10) See Koning (2013, p. 94-104) for a more comprehensive review of Dutch welfare state reforms targeting immigrants specifically.

constituted a fundamental 'Other' (Eijberts, 2013, p. 124), to attract attention as a potential threat to the welfare of the Dutch in-group; being the ultimate out-group, these immigrants are an easy target for perceptions of being exploitative and undeserving of Dutch welfare, as Van Oorschot's (2006) conditions for deservingness would also indicate. While immigrants in general are already deemed least deserving of welfare provisions (Careja & Emmenegger, 2013, p. 147), Muslim immigrants may find themselves at the extreme end of the scale, as they tend to be clearly marked as members of the out-group by both physical indicators – signaling, in Van Oorschot's (2006) terms, an undesirable 'identity' – and behavioral indicators – a deviant 'attitude'. The aforementioned plans to exclude immigrants from welfare provisions based on divergent "behavior or dress" (Government of the Netherlands, 2010, p. 26) is a clear example of welfare chauvinism on the basis of these two criteria. Moreover, as mentioned before, the long-term socio-economic marginalization of Muslim immigrants has caused their ethnic identity to become linked to a poor economic status as well as perceptions of passivity or laziness (Eijberts, 2013, p. 124), which may have led the native Dutch to believe that this immigrant group would be unable to reciprocate any given support. In other words, in the context of Dutch separatist multiculturalism and an allegedly threatened Dutch identity, one could conclude that Muslim immigrants were highly likely to evoke hostility and perceptions of un-deservingness, as they tend to be evaluated more negatively by the native Dutch on the criteria of identity, attitude and reciprocity.

SOLVING INTERGROUP ANTAGONISM

If, as this reconstruction has demonstrated, separatist multicultural policies can be said to have stimulated ethnic fragmentation, intergroup antagonism, and the exclusion of (especially non-Western) minorities in the Netherlands, then the question inevitably emerges what the Dutch government could have done differently. In response to the marginalization and poor societal integration of immigrants, Dutch integration policies have recently taken a more assimilationist turn, obliging immigrants to follow courses for the acquisition of the Dutch language and culture (Vasta, 2007, p. 734). Goodhart (2004, 2013) advocates a similar policy: to restore societal unity, he argues, one must establish a singular, unified national identity. Yet there are reasons to be skeptical about the preferability of such assimilationist policies.

Firstly, it can be questioned whether, as critics of multiculturalism (Goodhart, 2013; Klausen, 2000; Scheffer, 2000) would have it, the segregation of minority groups can truly be blamed on multiculturalism as a whole, or only on a particular implementation of multiculturalism. In fact, contrary to such criticisms, there are indications that multiculturalism can be an integrative force (Banting, 2010; Winter, 2011). In Canada, the framework of multiculturalism has been used for the governmental support of multi-ethnic organizations (Banting, 2010). Already in the 1970s and 1980s, the Canadian government reformed its subsidies system for ethnic organizations, and started prioritizing the funding of multi-ethnic organizations over their mono-ethnic counterparts; in this way program meant for the support of minority cultures was used to establish an environment for positive intergroup contact (McAndrew, Helly, & Tessier, 2005). Rather than separating groups into different pillars and cutting off pathways to prejudice reduction, this policy has created spaces where ethnic groups must work together for the attainment of a common goal (ibid., 810). In this respect, Canadian multiculturalism may provide a model for constructing a pluralist national identity by bringing minority groups together under the banner of celebrating their cultural difference (Winter, 2011). However, this multicultural model has not been without its critics (Sharma, 2011). For example, Winter (ibid., p. 214) suggests that this model, instead of cherishing minority cultures, might erase group identities in the long term. This indicates that a systematic study of different interpretations of multiculturalism and their long-term effects is needed to determine the relative merits and faults of this 'integrative multiculturalism', and its potential for implementation elsewhere.

Contrastingly, Dutch integration policies seem to increasingly follow an assimilationist model, which attempts to suppress intergroup difference by enforcing a unified national identity and culture (Vasta, 2007). Proponents of this model point at the overrepresentation of immigrants in welfare provisions and crime, and argue that the only way to prevent this societal chaos from escalating is to restore cultural and moral unity (Schinkel & Van Houdt, 2010, p. 702-3). While to a certain degree, cultural knowledge and language acquisition may help immigrants participate in the host society (Vasta, 2007, p. 734), the way in which assimilationists tend to define national identity ethnically or culturally raises problems for a multi-ethnic welfare state. In the Dutch context, some have spoken of a 'culturalization of citizenship', where culture takes up a central role in the construction of citizenship and the integration process (Mepschen, Duyvendak & Tonkens, 2010, p. 964). As a result, cultural characteristics and attitudes become markers with which to distinguish 'real' citizens from 'inferior' ones (Schinkel, 2010). However, despite the fact that this might restore the unity and integrity of the in-group by supposedly reinstating the old group boundaries, there are nonetheless indications that a strong, culturalized national identity does not lead to a more harmonious society (Wright & Reeskens, 2013). To the contrary, a national identity defined in cultural terms may actually engender welfare chauvinism, as such an identity is by definition exclusive of immigrants (ibid., p. 1457-9).

This makes sense from an intergroup contact perspective (Brewer, 1999). The culturalization of citizenship establishes a single dimension of difference, one's cultural background, to distinguish 'real' members of the in-group from lesser ones and members of the out-group (ibid., p. 439). This is an incendiary situation rife with potential for intergroup conflict (Pelinka, 2007, p. 133), as the singular categorization invites "social comparison and perceptions of conflict of interest that give rise to negative attitudes toward outgroups," (Brewer, 1999, p. 439). Whenever groups are differentiated on a single characteristic alone, this creates a situation of 'parallel social cleavages' (Pelinka, 2007, p. 133), where the out-group is perceived as fundamentally different.

A more stable and harmonious society results from cross-cutting social cleavages, where multiple dimensions of difference cut across one another and people have multiple, potentially contrasting loyalties (Andeweg, 2000). An example is pillarization: while the four pillars in Dutch society where highly segregated on the basis of religious and political identities, this did not lead to fierce intergroup conflict, as the dimension of class cut across all four pillars (ibid., p. 509, 517). In the case of cross-cutting social cleavages, two people may belong to the same in-group along one cleavage, and to different groups across another; consequently, there is less distance between the members of different groups, and people are less likely to identify strongly enough with a single group for this to cause intergroup conflict (Brewer, 1999, p. 439). It thus seems that a more inclusive group identity which does not define itself in a particular ethnic or cultural way, but instead leaves people free to identify with multiple different groups, is necessary to prevent the kind of in-group—out-group dynamics which underlie welfare chauvinism. When societies are too cleanly separated into different groups, as in the case of an ethnicized or culturalized group identity, this increases the risk for intergroup hostility. What is needed instead is what Parekh (2001, p. 694) calls a "community of communities," not reliant on internal homogeneity, and where people are free to engage with any group and do not become "imprisoned within or defined by these communities." Allport (1954/1979) already hinted at this idea with his concept of 'concentric loyalties', where larger, superordinate identities can subsume smaller sub-groups. Yet as Brewer (1999, p. 439-40) rightly points out, the practical implementation of this ideal form of identification may prove problematic, as little is known about the dynamics of cross-cutting memberships. For instance, it is unclear how a membership in multiple social categories serves the simultaneous need to belong and to be different (ibid., p. 441). This is an important avenue for future research, since the current trend towards assimilation and the culturalization of citizenship will do little to solve the issue of intergroup antagonism and the exclusion of out-group minorities from in-group welfare.

CONCLUSION

The emergence of the progressive dilemma in the multi-ethnic Western welfare states had been a long time coming. Ever since the arrival of large groups of immigrants in countries such as Great Britain and the Netherlands, governments had been unable to prevent them from slowly sinking into marginalized positions. In both countries, attempts to let immigrant groups fend for themselves with little governmental interference turned out less beneficial than expected. High unemployment rates and welfare dependence among immigrants became a serious cause for concern; the contributors of the welfare state had to pick up the tab for new neighbors who had failed to introduce themselves. In a radical change of course, the Dutch multicultural tradition turned into one of the harshest assimilationist discourses in Europe; the same government which had encouraged immigrants to build up their own societal pillar only a few decades earlier, now attempted to dictate their manner of dress.

Where did it go wrong? The progressive dilemma offered two answers. First, ethnic diversity may have been an inherent weak spot for the welfare state all along. Implemented in highly homogenous societies, contributors could trust that whoever would receive their support would be in some way like them, facing a problem they might one day face themselves; now, they could no longer identify with their beneficiaries. The second answer put the blame on governments themselves: perhaps it had not been ethnic diversity per se, but the way it had been approached though overly lenient multicultural policies. Governments had been too relaxed about the cultural integration of immigrants, who as a result became segregated in their own communities.

While the literature on welfare chauvinism is recent and fairly limited, it shows a general support for the first hypothesis: there appears to be a weak negative correlation between ethnic heterogeneity and (support for) income redistribution. This relationship cuts across the boundaries of welfare regimes. There is however one notable exception: Canada, a long-time champion of multiculturalism has not seen a rise in welfare chauvinism. This case also serves as a counter-example for the second hypothesis: multiculturalism as such, it seems, is not damaging for the welfare state. There is however a possibility that a particular interpretation of multiculturalism does contribute to a progressive dilemma.

To examine this possibility, and arrive at a possible solution for the issue of welfare chauvinism, a theoretical understanding of the phenomenon is needed. This thesis proposed Allport's (1954/1979) intergroup contact theory, in combination with Van Oorschot's (2006) conditions for deservingness as a fruitful way of analyzing this problem.

By applying this theoretical framework to the Dutch case, it is found that the separatist multicultural policies of the 1980s contributed to an environment where positive, prejudice-reducing intergroup contact was unlikely. On the other hand, strains on Dutch national identity made the inclusion of immigrants into the native Dutch in-group improbable. Together, separatist multiculturalism and pressures on Dutch identity made for a situation where particularly non-Western or Muslim immigrants could easily be perceived as a threat to Dutch in-group welfare and undeserving of support. However, as the exclusion of immigrants from in-group welfare is based on perceived rather than objective, insurmountable difference, this means there may be ways of reducing or solving the issue of welfare chauvinism.

Interestingly, both pathways towards solutions identified in this thesis come from the side of multiculturalism. Firstly, it appears Canada has implemented a more integrative multicultural policy, which creates spaces for positive intergroup contact. By reducing intergroup antagonism this way, Canada has attempted to construct a pluralist national identity. Further research must systematically compare multicultural policies and their effects, to determine the merits of a more integrative multiculturalism and its potential for implementation elsewhere. Secondly, and slightly more radically, it seems a solution can be found in a more loosely defined national identity. In this regard, Parekh (2001, p. 694) has spoken of a "community of communities," a superordinate identity within which one can identify with multiple sub-groups. However, there is a lot that is still unknown about the implementation of such superordinate identities and the membership in multiple social categories, a gap which warrants future research.

This is a particularly pressing avenue for further study, as the current trends in Dutch integration policies toward a more assimilationist and culturalized approach clearly offer no solution to the problem of welfare chauvinism. The culturalization of citizenship is merely likely to create a situation of parallel social cleavages, in which intergroup conflict is a distinct possibility. Assimilation only serves to rigidify existing group boundaries, reinforcing the same exclusionary in-group—out-group dynamics. In other words, unless something is changed fundamentally about these dynamics, through positive intergroup contact or more inclusive identities, the current situation is unlikely to change. After all, the progressive dilemma might have taken several decades to emerge, but it may well be here to stay.

REFERENCES

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2), 155-194.

Alesina, A., & Glaeser, E. (2004). F*ighting Poverty in the US and Europe: A World of Difference.* Oxford: Oxford University Press.

Alesina, A., Glaeser, E., & Sacerdote, B. (2001). Why Doesn't the United States Have a European-Style Welfare State?. *Brookings Papers on Economic Activity, 2001(2),* 187-254.

Allport, G. W. (1979). *On the Nature of Prejudice* (Rev. ed.). Reading, MA: Addison-Wesley Publishing (Original work published 1954).

Alvarez, R. M., & Butterfield, T. L. (2000). The Resurgence of Nativism in California? The Case of Proposition 187 and Illegal Immigration. *Social Science Quarterly*, 81(1), 167-179.

Andeweg, R. B. (2000). Consociational Democracy. *Annual Review of Political Science*, 3(1), 509-536.

Banting, K. (2010). Is There a Progressive's Dilemma in Canada? Immigration, Multiculturalism and the Welfare State. *Canadian Journal of Political Science*, 43(4), 797-820.

Banting, K., Johnston, R., Kymlicka, W., & Soroka, S. (2006). Do Multiculturalism Policies Erode the Welfare State? An Empirical Analysis. In K. Banting & W. Kymlicka (Eds.), *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies* (pp. 49-91). New York, NY: Oxford University Press.

Banting, K., & Kymlicka, W. (2006). Introduction. In K. Banting & W. Kymlicka (Eds.), Multiculturalism and the Welfare State: *Recognition and Redistribution in Contemporary Democracies* (pp. 1-45). New York, NY: Oxford University Press.

Barry, B. (2001). *Culture and Equality: An Egalitarian Critique of Multiculturalism*. Cambridge: Polity Press.

Barth, F. (1969). Introduction. In Ethnic Groups and Boundaries: T*he Social Organization of Culture Difference* (pp. 9-37). Boston: Little, Brown and Company.

Bevelander, P., & Groeneveld, S. (2012). How Many Hours Do You Have to Work to Be Integrated? FullÐ Time and PartÐTime Employment of Native and Ethnic Minority Women in the Netherlands. *International Migration*, 50(S1), e117-e131.

Beveridge, W. (1942). *Report of the Inter-Departmental Committee on Social Insurance and Allied Services*. London: HMSO

Brameld, T., & Fish, E. (1946). School Administration and Intercultural Relations. *The Annals of the American Academy of Political and Social Science*, 244, 26-33.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, 17(5), 475-482.

Brewer, M. B. (1999). The Psychology of Prejudice: Ingroup Love and Outgroup Hate?. *Journal of Social Issues*, 55(3), 429-444.

Brewer, M. B. (2001). Ingroup Identification and Intergroup Conflict: When Does Ingroup Love Become Outgroup Hate?. In R.D. Ashmore, L. J. Jussim, & D. Wilder (Eds.), *Social Identity, Intergroup Conflict, and Conflict Reduction*, (pp. 17-41). Oxford: Oxford University Press.

Brown, R., & Zagefka, H. (2005). Ingroup Affiliations and Prejudice. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the Nature of Prejudice: Fifty Years after Allport* (pp. 54-70). Oxford: Blackwell.

Burgoon, B., Koster, F., & Van Egmond, M. (2012). Support for Redistribution and the Paradox of Immigration. *Journal of European Social Policy*, 22(3), 288-304.

Burgoon, B. (2011, June). *Immigration, Integration and Support for Redistribution in Europe* (IC3JM working paper 2011/260). Madrid: The Carlos III-Juan March Institute. Retrieved January 24, 2014, from http://www.march.es/ceacs/Publicaciones /working/archivos/WP-260-Burgoon.pdf

Cameron, L., & Rutland, A. (2006). Extended Contact through Story Reading in School: Reducing Children's Prejudice Toward the Disabled. *Journal of Social Issues*, 62(3), 469-488.

Careja, R., & Emmenegger, P. (2013) Keeping Them Out! Migration and Social Policies in the 'Reluctant Countries of Immigration'. In A. M. Suszycki & I. P. Karolewski (Eds.), *Citizenship and Identity in the Welfare State* (pp. 139-164). Baden-Baden: Nomos.

Carnoy, M., & Castells, M. (2001). Globalization, the Knowledge Society, and the Network State: Poulantzas at the Millennium. *Global Networks*, 1(1), 1-18.

Castells, M. (2000). Toward a Sociology of the Network Society. *Contemporary Sociology*, 29(5), 693-699.

Castles, S. (1997). Multicultural Citizenship: A Response to the Dilemma of Globalisation and National Identity? *Journal of Intercultural Studies*, 18(1), 5-22.

Choudhry, S. (2000). Distribution vs. Recognition: The Case of Anti-Discrimination Laws. *George Mason Law Review*, 9(1), 145-178.

Crepaz, M. (2006). 'If You Are My Brother, I May Give You a Dime!' Public Opinion on Multiculturalism, Trust, and the Welfare State. In K. Banting & W. Kymlicka (Eds.), *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies* (pp. 92-120). New York, NY: Oxford University Press.

Dovidio, J. F., Gaertner, S. L., & Kawakami, K. (2003). Intergroup Contact: The Past, Present, and the Future. *Group Processes & Intergroup Relations*, 6(1), 5-21.

Duyvendak, J. W., & Scholten, P. W. (2011). Beyond the Dutch "Multicultural Model". *Journal of International Migration and Integration*, 12(3), 331-348.

Easterly, W., & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics*, 112(4), 1203-1250.

Eger, M. A. (2010). Even in Sweden: The Effect of Immigration on Support for Welfare State Spending. *European Sociological Review*, 26(2), 203-217.

Eijberts, M. (2013). *Migrant Women Shout it Out Loud: The Integration/Participation Strategies and Sense of Home of First- and Second-Generation Women of Moroccan and Turkish Descent.* Amsterdam: VU University Press.

Entzinger, H. (2006). The Parallel Decline of Multiculturalism and the Welfare State in the Netherlands. In K. Banting & W. Kymlicka (Eds.), *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies* (pp. 177-201). New York, NY: Oxford University Press.

Freeman, G. P. (2009). Immigration, Diversity, and Welfare Chauvinism. *The Forum*, 7(3), article 7.

Gaertner, S. L., & Dovidio, J. F. (2005). Categorization, Recategorization and Intergroup Bias. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the Nature of Prejudice: Fifty Years after Allport* (pp. 71-88). Oxford: Blackwell.

Ghorashi, H. (2010). From Absolute Invisibility to Extreme Visibility: Emancipation Trajectory of Migrant Women in the Netherlands. *Feminist Review*, 94(1), 75-92.

Goodhart, D. (2004). Too Diverse?. *Prospect Magazine*, 95, 30-37.

Goodhart, D. (2013). *The British Dream: Successes and Failures of Post-War Immigration*. London: Atlantic Books.

Government of the Netherlands (2010). *Vrijheid en Verantwoordelijkheid*: Regeerakkoord VVD-CDA [Liberty and Responsibility: Coalition Agreement VVD-CDA]. Retrieved from http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/rapporten /2010/09/30/regeerakkoord-vvd-cda/regeerakkoord-vvd-cda.pdf

Herek, G. M., & Capitanio, J. P. (1996). "Some of My Best Friends": Intergroup Contact, Concealable Stigma, and Heterosexuals' Attitudes Toward Gay Men and Lesbians. *Personality and Social Psychology Bulletin*, 22(4), 412-424.

Katz, I. (1991) Gordon Allport's T*he Nature of Prejudice. Political Psychology*, 12(1), 125-157.

Kenworthy, J. B., Turner, R. N., Hewstone, M., & Voci, A. (2005). Intergroup Contact: When Does it Work, and Why? In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), On the Nature of Prejudice: *Fifty Years after Allport* (pp. 278-292). Oxford: Blackwell.

Knigge, P. (1998). The Ecological Correlates of Right-Wing Extremism in Western Europe. *European Journal of Political Research*, 34(2), 249-279.

Klausen, J. (2000, December 20). *Other People*. Prospect (Dec.), 28–33. Retrieved from https://www.prospectmagazine.co.uk/magazine/otherpeople

Koning, E. A. (2013). *Selective Solidarity: The Politics of Immigrants' Social Rights in Western Welfare States* (Doctoral Thesis, Queen's University, Kingston, Canada). Retrieved from http://qspace.library.queensu.ca/handle/1974/7922

Koopmans, R. (2003). Good Intentions Sometimes Make Bad Policy: A Comparison of Dutch and German Integration Policies. In R. Cupeurs, K. A. Duffek & Kandel J. (Eds.), *The Challenge of Diversity: European Social Democracy Facing Migration, Integration, and Multi- Culturalism*. Innsbruck: Studien Verlag.

Koopmans, R. (2010). Trade-offs Between Equality and Difference: Immigrant Integration, Multiculturalism and the Welfare State in Cross-National Perspective. *Journal of Ethnic and Migration Studies*, 36(1), 1-26.

Koschate, M., & Van Dick, R. (2011). A Multilevel Test of Allport's Contact Conditions. *Group Processes & Intergroup Relations*, 14(6), 769-787.

Kraus, P. A., & Schönwälder, K. (2006). Multiculturalism in Germany: Rhetoric, Scattered Experiments and Future Chances. In K. Banting & W. Kymlicka (Eds.), *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies* (pp. 202-221). New York, NY: Oxford University Press.

Larsen, C. A. (2008). The Institutional Logic of Welfare Attitudes: How Welfare Regimes Influence Public Support. *Comparative Political Studies*, 41(2), 145-168.

Larsen, C. A. (2011). Ethnic Heterogeneity and Public Support for Welfare: Is the American Experience Replicated in Britain, Sweden and Denmark?. *Scandinavian Political Studies*, 34(4), 332-353.

Lechner, F. J., & Boli, J. (2010). World Culture: Origins and Consequences. In G. Ritzer & Z. Atalay (Eds.), Readings in Globalization: *Key Concepts and Major Debates* (pp. 410-421). Chichester: Wiley-Blackwell. (Reprinted from World Culture: Origins and Consequences, by F. J. Lechner and J. Boli, 2005, Maiden, MA: Blackwell)

Mann, M. (1997). Has Globalization Ended the Rise and Rise of the Nation-State?. *Review of International Political Economy*, 4(3), 472-496.

Marshall, T. H. (2010). Citizenship and Social Class. In C. Pierson & F. G. Castles (Eds.), The Welfare State Reader (pp. 32-41). Cambridge: Polity Press. (Reprinted from *Citizenship and Social Class and Other Essays*, by T. H. Marshall, Ed., 1950, Cambridge: Cambridge University Press).

Mau, S., & Burkhardt, C. (2009). Migration and Welfare State Solidarity in Western Europe. *Journal of European Social Policy*, 19(3), 213-229.

McAndrew, M., Helly, D., & Tessier, C. (2005). Pour un débat éclairé sur la politique Canadienne du multiculturalisme: une analyse de la nature des organismes et des projets subventionnés (1983-2002). *Politique et Sociétés*, 24(1), 49-71.

McLaughlin, E., & Neal, S. (2004). Misrepresenting the Multicultural Nation: The Policy-Making Process, News Media Management and the Parekh Report. *Policy Studies*, 25(3), 155-174.

Mepschen, P., Duyvendak, J. W., & Tonkens, E. H. (2010). Sexual Politics, Orientalism and Multicultural Citizenship in the Netherlands. *Sociology*, 44(5), 962-979.

Mewes, J., & Mau, S. (2013). Globalization, Socio-Economic Status and Welfare Chauvinism: European Perspectives on Attitudes toward the Exclusion of Immigrants. *International Journal of Comparative Sociology,* 54(3), 228-245.

Miller, D. (2006). Multiculturalism and the Welfare State: Critical Reflections. In K. Banting and W. Kymlicka (Eds.), *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies* (pp. 323-338). New York, NY: Oxford University Press.

Nannestad, P. (2007). Immigration and Welfare States: A Survey of 15 Years of Research. *European Journal of Political Economy*, 23(2), 512-532.

Pelinka, A. (2007). Language as a Political Category: The Viewpoint of Political Science. *Journal of Language & Politics,* 6(1), 129-143.

Pettigrew, T. F. (1997). Generalized Intergroup Contact Effects on Prejudice. *Personality and Social Psychology Bulletin*, 23(2), 173-185.

Pettigrew, T. F. (1998). Intergroup Contact Theory. *Annual Review of Psychology*, 49(1), 65-85.

Pettigrew, T. F., & Tropp, L. R. (2005). Allport's Intergroup Contact Hypothesis: Its History and Influence. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the Nature of Prejudice: Fifty Years after Allport* (pp. 262-277). Oxford: Blackwell.

Pettigrew, T. F., & Tropp, L. R. (2006). A Meta-Analytic Test of Intergroup Contact Theory. *Journal of Personality and Social Psychology,* 90(5), 751-783.

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science*, 48(4), 849-863.

Putnam, R. D. (2007). E Pluribus Unum: Diversity and Community in the TwentyÐFirst Century. The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, 30(2), 137-174.

Reeskens, T., & Van Oorschot, W. (2012). Disentangling the 'New Liberal Dilemma': On the Relation between General Welfare Redistribution Preferences and Welfare Chauvinism. *International Journal of Comparative Sociology*, 53(2), 120-139.

Sainsbury, D. (2006). Immigrants' Social Rights in Comparative Perspective: Welfare Regimes, Forms in Immigration and Immigration Policy Regimes. *Journal of European Social Policy*, 16(3), 229-244.

Salter, F. (2002). Estimating Ethnic Genetic Interests: Is It Adaptive to Resist Replacement Migration?. *Population and Environment*, 24(2), 111-140.

Sassen, S. (1991). *The Global City: New York, London, Tokyo*. Princeton, NJ: Princeton University Press.

Savelkoul, M., Scheepers, P., Tolsma, J., & Hagendoorn, L. (2011). Anti-Muslim Attitudes in the Netherlands: Tests of Contradictory Hypotheses Derived from Ethnic Competition Theory and Intergroup Contact Theory. *European Sociological Review*, 27(6), 741-758.

Scheffer, P. (2000, January 29). Het Multiculturele Drama. *NRC Handelsblad*. Retrieved from http://retro.nrc.nl/W2/Lab/Multicultureel/scheffer.html

Schinkel, W. (2010). The Virtualization of Citizenship. *Critical Sociology*, 36(2), 265-283.

Schinkel, W., & Van Houdt, F. (2010). The Double Helix of Cultural Assimilationism and NeoÐLiberalism: Citizenship in Contemporary Governmentality. The *British Journal of Sociology*, 61(4), 696-715.

Senik, C., Stichnoth, H., & Van der Straeten, K. (2009). Immigration and Natives' Attitudes towards the Welfare State: Evidence from the European Social Survey. *Social indicators research*, 91(3), 345-370.

Sharma, N. (2011). Canadian Multiculturalism and its Nationalisms. In M. Chazan, L. Helps, A. Stanley & S. Thakkar (Eds.), *Home and Native Land: Unsettling Multiculturalism in Canada* (pp. 85-101). Toronto: Between the Lines.

Snel, E., & Scholten, P. W. A. (2005). Van Gastarbeider tot het Multiculturele Drama: Integratie als Hardnekkig Beleidsprobleem. In M. Arentsen & W. Trommel (Eds.), *Moderniteit en Overheidsbeleid; Hardnekkige Beleidsproblemen en hun Oorzaken* (pp. 165-189). Bussem: Countinho.

Sniderman, P. M., Peri, P., de Figuereido, R. J. P., & Piazza, T. (2000). *The Outsider.* Princeton, NJ: Princeton University Press.

Soysal, Y. N. (1994). *Limits of Citizenship: Migrants and Postnational Membership in Europe*. Chicago, IL: University of Chicago Press.

Soysal, Y. N. (2000). Citizenship and Identity: Living in Diasporas in Post-War Europe?. *Ethnic and Racial Studies*, 23(1), 1-15.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup Anxiety. *Journal of Social Issues*, 41(3), 157-175.

Taba, H., & Wilson, H. E. (1946). Intergroup Education Through the School Curriculum. *The Annals of the American Academy of Political and Social Science* 244, 19-25.

Tolsma, J., Van der Meer, T., & Gesthuizen, M. (2009). The Impact of Neighbourhood and Municipality Characteristics on Social Cohesion in the Netherlands. *Acta Politica*, 44(3), 286-313.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, 46(1), 35-57.

Van de Vijver, F. J., Breugelmans, S. M., & Schalk-Soekar, S. R. (2008). Multiculturalism: Construct Validity and Stability. *International Journal of Intercultural Relations*, 32(2), 93-104.

Van der Waal, J., Achterberg, P., Houtman, D., De Koster, W., & Manevska, K. (2010). 'Some Are More Equal Than Others': Economic Egalitarianism and Welfare Chauvinism in the Netherlands. *Journal of European Social Policy*, 20(4), 350-363.

Van Oorschot, W. (2006). Making the Difference in Social Europe: Deservingness Perceptions among Citizens of European Welfare States. *Journal of European Social Policy*, 16(1), 23-42.

Van Oorschot, W. (2008). Solidarity Towards Immigrants in European Welfare States. *International Journal of Social Welfare*, 17(1), 3-14.

Vasta, E. (2007). From Ethnic Minorities to Ethnic Majority Policy: Multiculturalism and the Shift to Assimilationism in the Netherlands. *Ethnic and Racial Studies*, 30(5), 713-740.

Winter, E. (2011). *Us, Them, and Others: Pluralism and National Identities in Diverse Societies*. Toronto: University of Toronto Press.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The Extended Contact Effect: Knowledge of Cross-Group Friendships and Prejudice. *Journal of Personality and Social Psychology*, 73(1), 73.

Wright, M., & Reeskens, T. (2013). Of What Cloth Are the Ties That Bind? National Identity and Support for the Welfare State across 29 European Countries. *Journal of European Public Policy*, 20(10), 1443-1463.

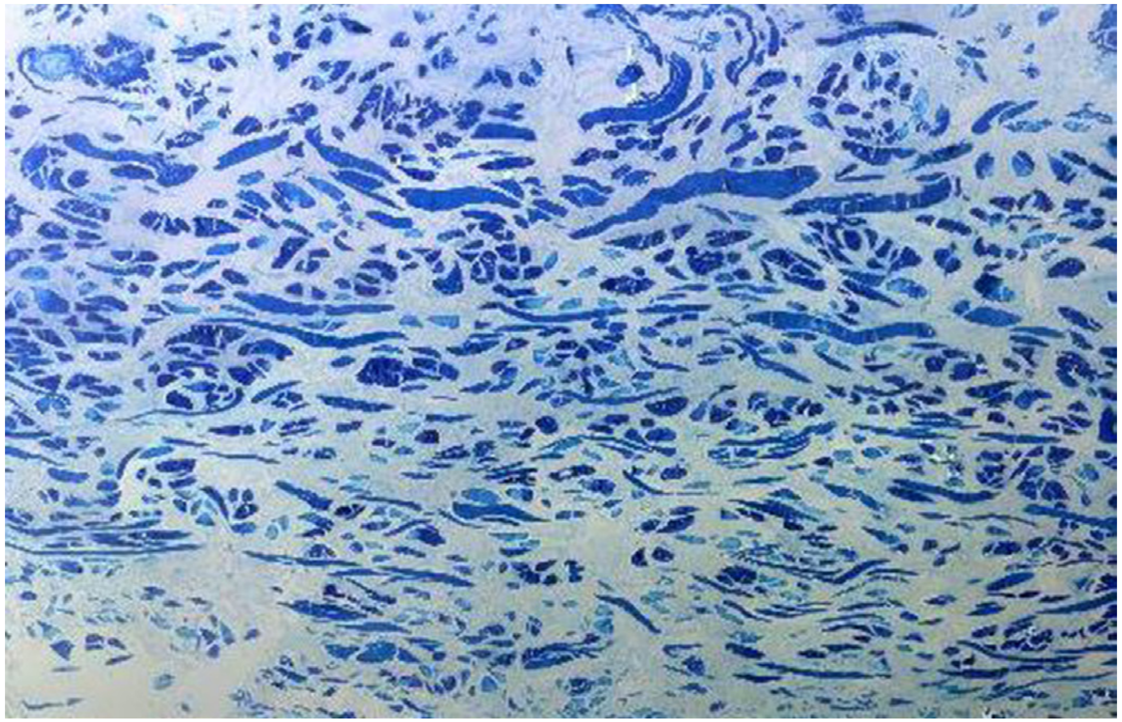
WRR. (2001). *Nederland als Immigratiesamenleving* [The Netherlands as a Country of Immigration]. Retrieved from http://www.wrr.nl/fileadmin/nl/publicaties/PDF-Rapporten/Nederland_als_immigratiesamenleving.pdf

Yerkes, M., & Van der Veen, R. (2011). Crisis and Welfare State Change in The Netherlands. *Social Policy & Administration*, 45(4), 430-444.

The Effect of Collagen Cross-Linking on Foreign Body Giant Cell Formation

Jelle Koopsen

Supervisors: Prof. V. Everts & Mr. B. ten Harkel (Academisch Centrum Tandheelkunde Amsterdam)



ABSTRACT

BACKGROUND: Collagen-based biomaterials are regularly used in regenerative medicine, but little is known about the immune response, known as the *foreign body response*, they elicit in vivo. Foreign body giant cells (FBGCs) are usually formed in the presence of these biomaterials and can affect biocompatibility. The aim of this study was to determine the effect of cross-linking these collagen based biomaterials on FBGC formation, and to describe notable morphological features of the FBGCs.
METHODS: Non-cross-linked dermal sheep collagen (NDSC) (n=6) and hexamethylene diisocyanate cross-linked dermal sheep collagen (HDSC) disks (n=6) were subcutaneously implanted in mice. The number of FBGCs formed at day 21 was assessed using light microscopy. Morphological features were analyzed with transmission electron microscopy (TEM).
RESULTS: HDSC elicited the formation of more FBGCs than NDSC. We hypothesize that the difference between HDSC and NDSC is largely due to the fact that HDSC is more difficult to degrade. This can lead to *frustrated phagocytosis*, which is hypothesized to be a driving force of cell-cell fusion, and will therefore lead to the formation of more giant cells. FBGCs were capable of completely engulfing collagen and extracellular degradation of both HDSC and NDSC was observed. Internalization of collagen particles was absent in both conditions. FBGCs displayed distinct finger-like structures on the cell-borders. The function of these structures is yet to be determined.
CONCLUSIONS: We conclude that cross-linking a collagen based biomaterial can aggravate the immune response against the biomaterial. Furthermore, by clarifying the behavior and morphology of FBGCs in the presence of differently cross-linked biomaterials, we increase our understanding of the foreign body response and its implication on biocompatibility of biomedical devices.

INTRODUCTION

Collagen-based biomaterials are commonly used in regenerative medicine and reconstructive surgical procedures. They are often used to augment tissue repair and reinforce soft tissue. These biomaterials are often composed of mammalian extracellular matrix (ECM), and are taken from several species of mammals (e.g. equine, bovine, porcine, or human) depending on the purpose of the tissue. Clearly, these materials are allogeneic or xenogeneic and can elicit an immune response that plays a role in biocompatibility. Even though these biomaterials are extensively used, this immune response is not yet fully understood[1].

The immune response is initiated when the biomaterials come in contact with the host tissue. Instantly, a layer of host proteins is formed on the material. The type of proteins depend on the type of biomaterial. These proteins modulate the inflammatory response by interacting with the inflammatory cells. The interaction leads to the formation of a thrombus. This thrombus releases chemo attractants, such as transforming growth factor (TGF-Ð) and CXCL4, which attract macrophages to the implementation site[2]. Mast cells found in the thrombus have also been shown to play a role in the recruitment of macrophages[3]. The presence of macrophages at the implementation site also attracts additional macrophages by producing, among others, tumor necrosis factor (TNF-α), interleukin-6 (IL-6) and granulocyte macrophage colony stimulating factor (GM-CSF), creating a positive feedback loop. Via integrin signaling, these macrophages interact with the layer of host proteins on the biomaterial. During this interaction, macrophages fuse under the influence of interleukins. Both IL-4 and IL-13 have been described to initiate macrophage fusion[4,5]. IL-4 and IL-13 are released by mast cells[2].

Macrophage fusion is a highly regulated process which occurs rarely in mammals[6]. It can lead to either osteoclast formation, when activated by the cytokines M-CSF and RANKL, or giant cell formation, when activated by the previously mentioned IL-4 or IL-13. When this fusion is the result of a foreign material implementation, the fused macrophages are referred to as *foreign body giant cells* (FBGCs). This process is referred to as the *foreign body response* (FBR).

Many questions about the role of FBGCs in the immune response remain. Although many studies have investigated the mediators involved in FBGC formation, little is known about the mechanisms by which these giant cells function. It has been demonstrated that FBGCs have the potential to respond to cellular signalling[7] and are capable of cytokine production[8]. It was demonstrated that FBGCs are actively involved in the inflammatory response by secreting IL-1α and TNF-α[8]. The interaction between the collagen-based biomaterials and FBGCs, however, has not been extensively studied.

The interaction between FBGCs and collagen is therefore an interesting field of research. Previous studies have shown that FBGCs are capable of degrading collagen through a process called *frustrated phagocytosis*[9,10]. This is a process that occurs when the biomaterial is too large to be fully engulfed by the immune cells. During frustrated phagocytosis degradative enzymes, reactive oxygen species, or other products are secreted in an attempt to degrade the biomaterial[11]. This process has also been hypothesized to be a driving force of the multi-nucleation of FBGCs[12]. It has been demonstrated that FBGCs are also capable of internal degradation of collagen in vivo[13]. The extent and speed of degradation is dependent on whether or not the collagen is chemically cross-linked. Crosslinking of several types of collagen strengthens the ECM scaffold and inhibits its degradation significantly[14,15]. Both cross-linked and non-cross-linked collagen are used in commercially available devices[1]. The effect of the increased scaffold persistence on the FBR is not yet fully determined. This study will investigate the effect of collagen-based biomaterials on foreign body giant cell formation. We will use dermal sheep collagen as a model for a collagen-based biomaterial. Cross-linked collagen will be compared to non-cross-linked collagen.

2. LITERATURE REVIEW

2.1. COLLAGEN CROSS-LINKING

Collagen cross-linking is used to stabilize collagen-based biomaterials in order to decrease enzymatic degradation. Collagenases can degrade these biomaterials by attacking the triple helical molecule at the N-terminal and C-terminal fragments of collagen. These fragments then become transformed to gelatin and are cleaved to oligopeptides by other naturally occurring enzymes[16]. The physical properties of these biomaterials are greatly determined by the method of cross-linking. Cross-linking can be done either physically or chemically. Physical cross-links can be obtained by, for instance, ultraviolet irradiation or dehydrothermal treatment[17]. Chemical cross-linking is used most often nowadays. Glutaraldehyde (GA) is the chemical reagent most generally used[18]. However, van Luyn et al. demonstrated detrimental calcification of GA treated collagen subcutaneously[19]. Furthermore, GA is known to elicit toxic effects[20]. In contrast, collagen cross-linked with hexamethylene diisocyanate (HMDIC) showed similar physical properties to GA, without the toxic effects[20]. HMDIC interacts with the N-terminals of the polypeptide chains of the collagen, thereby bridging them and creating cross-links within the collagen fibers. This protects the collagen against the collagenase attack on the N-terminals. Hence cross-linking collagen decreases degradation rate of collagen-based biomaterials.

The increased stability of the cross-linked collagen has an effect on the physiological tissue repair process that follows biomaterial implantation. In addition, the method of cross-linking also effects the repair process. Lastly, these effects are altered in different implant locations[21]. Although the effect on the FBR is not completely clear, some progress has been made in this area of research. Ye et al. showed that GA cross-linked dermal sheep collagen disks (GDSC) showed a different response than HMDIC cross-linked dermal sheep collagen (HDSC) when implanted subcutaneously in rats[22]. The FBR in the two settings differed in several aspects. Most importantly, they showed that GDSC was almost completely degraded after 28 days of implementation, while HDSC showed almost no degradation. Furthermore, little to no FBGCs were formed in GDSC samples while a vast amount of giant cells were formed on the HDSC samples[22]. This led to the hypothesis that the inability to degrade collagen disk leads to the formation of FBGCs. This is concordant with the hypothesis that frustrated phagocytosis due to the inability to degrade the biomaterial is a driving force behind the multi-nucleation of giant cells.

It remains unclear whether HMDIC cross-linking of collagen is preferable to GA cross-linking: while it does not elicit a toxic response, it does induce the formation of FBGCs. Giant cell formation is an integral part of the FBR, which is associated with implant failure[23].

2.2. MACROPHAGE FUSION

One of the key questions in FBR research is why macrophages fuse when they encounter a non-phagocytozable material. As previously mentioned, one hypothesis is that the "frustrated" macrophages fuse because they are unable to phagocytize the material. In this case, fusing would be a mechanism that increased the phagocytic capabilities of the cell. Another suggestion is that the macrophages fuse to increase their surface area, similar to osteoclasts. FBGCs are known to cover up to 25% of the implant's surface area[24]. This might be in order to protect the body from the chronic inflammation near the biomaterial. Neither of these hypotheses have been sufficiently studied.

Many aspects of macrophage fusion are not fully understood: signaling pathways, molecular mechanisms of fusion, and significance of multi-nucleation remain to be somewhat of an enigma[12]. In the last two decades, however, some progress has been made on elucidating the signaling pathway leading to FBGC formation. Besides the previously mentioned IL-4 and IL-13, Ð-tocopherol (a vitamin) was also found to be a potent inducer of FBGC formation24. Their specific mechanisms of action are yet to be discovered.

Some conditions on a molecular level have been found to be necessary for fusion, but the driving force is not yet known. McNally and Anderson found that FBGC formation and macrophage fusion require macrophage adhesion to the adhesion proteins on the surface of the biomaterial[12]. Whether or not the macrophages adhere to the biomaterial is dependent on the chemistry of the surface and the adsorbed blood proteins[25]. Cytokine expression of the adherent macrophages is also dependent on the surface chemistry of the biomaterial2. Furthermore, Yagi et al. determined that dendritic cell–specific transmembrane protein (DC-STAMP), was necessary for macrophage fusion in both osteoclasts and FBGCs[26].

During macrophage fusion, many cytoskeletal changes occur necessary for fusion. Especially the formation of microfilaments, consisting of F-actin, is essential: macrophage fusion was inhibited when microfilaments were disrupted[27]. McNally and Anderson described the fusion process as a 'phagocytic mechanism', because of the similar dependence on F-actin and the fact that many phagocytosis markers were found during macrophage fusion[12].

It is clear that macrophage fusion is a highly regulated process, of which some mechanisms and key processes have been discovered, but that many questions currently remain unanswered.

2.3. THE ROLE OF MMPS

Macrophage fusion is also modulated by matrix matelloproteases (MMPs). MMPs are endopeptidases capable of degrading extra-cellular matrix components by cleaving peptide bonds. Several enzymes of the MMP family are collagenases, enzymes that can degrade collagen. Because ECM remodeling and degradation occurs in the presence of FBGCs, the role of MMPs in the foreign body reaction has been investigated. The outcomes of these studies, however, are not conclusive. Jones et al. showed high numbers of MMP-9, a gelatinase and collagenase, in FBGC cultures after 7 days, compared to other MMPs[28]. These FBGCs were cultured *in vitro*, on surfaces coated with different biomaterials. Because MMP-9 was so prevalent, they tested the effect of MMP-9 inhibition on fusion and adhesion. They found that pharmacological inhibition of MMP-9 did not affect adhesion of macrophage fusion[28].

In contrast to the findings by Jones et al., MacLauchlan et al. did find a role of MMP-9 in macrophage fusion29. They showed that FBGC formation was compromised in MMP-9 deficient mice. In the MMP-9 null mice, the macrophages did not undergo the changes in cell-shape necessary for cell fusion. Furthermore, when MMP-9 antibodies were introduces in an *in vitro system*, significantly smaller FBGCs were formed, again suggesting a role for MMP-9 in fusion[29]. Although controversy remains about the role of MMP-9 in macrophage fusion and FBGC formation, the results of the knock-out study seem more convincing than the study by Jones et al.

While Jones et al. did not find MMP-9 involvement in macrophage fusion, their findings on different MMPs did suggest a role for several other MMPs in macrophage fusion. Furthermore, they found differences in MMP expression between the several biomaterials, suggesting a role for the surface chemistry in macrophage related MMP secretion[28]. Lastly, Luttikhuizen et al. demonstrated the *in vivo* presence of collagenases MMP-8 and MMP-13, and gelatinases MMP-2 and MMP-9 in macrophages adhering to biomaterials[21]. It has been proposed that the combination of these collagenases and gelatinases is partly responsible for the ECM remodeling and collagen degradation. It is interesting to note that when IL-4, the FBGC inducing cytokine, is introduced in a monocyte/macrophage cell-culture, MMP-9 expression drops greatly *in vitro*[28,30]. This suggests that the induction of FBGC formation might lower de degradative capabilities of the monocytes/macrophage. More research is necessary to determine the exact role of MMPs in the FBR.

2.4. COLLAGEN DEGRADATION

The MMPs are probably released in the process referred to as "frustrated phagocytosis". In addition to these degradative enzymes, reactive oxygen species (ROS) and acid have also been proposed to contribute to ECM remodeling[2].

Besides the MMP action within the microenvironment between the FBGC and the biomaterials, it has also been suggested that FBGCs can internalize collagen and degrade it there. In the previously mentioned study by van Wachem et al. internalization of collagen was observed[13]. They also showed engulfment of collagen by the FBGC without internalization. They found that both intracellular and surrounded collagen was degraded by detachment of the collagen fibrils. They referred to the giant cells as either fibroblast-like foreign body multinucleated giant cells or macrophage-like foreign body multinucleated giant cells. Interestingly, Ye et al. showed no degradation of HDSC disks when implanted subcutaneously in rats[22], while van Wachem et al. did show HDSC disk degradation under the same conditions[13]. Because the setting of both experiments was almost identical, these results were unexpected, and a conclusive explanation is yet to be found. So even under similar circumstances, it seems that the FBR can differ, and even the degradative capabilities of FBGCs could be affected.

The previously mentioned study by Ye et al. demonstrated no degradation of HDSC even though the number of FBGCs was the highest in the HDSC samples[22]. Similar results were observed by the same research group in NDSC samples: while FBGCs were abundant in the NDSC samples, little degradation was observed. It has been established that FBGCs are capable of collagen degradation through the aforementioned frustrated phagocytosis, which makes these findings rather unexpected.

Our study will therefore determine whether degradation of HDSC and NDSC by FBGCs is visible though electron microscopy. Furthermore, to this date, the study by van Wachem et al. is the only study reporting HDSC internalization by FBGCs. Therefore, our study will also investigate whether collagen internalization is visible on an electron microscopical level.

3. MATERIALS & METHODS

3.1. IN VIVO FBGC FORMATION

We obtained all samples from colleagues at the Universitair Medisch Centrum in Groningen, The Netherlands. They implanted both non-cross-linked dermal sheep collagen (NDSC) discs and hexamethylene diisocyanate cross-linked dermal sheep collagen (HDSC) discs subcutaneously in 6 mice (3 male, 3 female). After 21 days, all discs were removed. The preparation of NDSC and the cross-linking methods used to obtain HDSC disks are described by Olde Damink et al[20]. There were 12 samples in total: 1 HDSC disc and 1 NDSC disc per mouse.

3.2. SAMPLE PREPARATION

All samples were prefixed in 1% glutaraldehyde plus 4% paraformaldehyde in 0.1 M sodium cacodylate buffer and post fixed with a solution of 1% OsO4 in cacodylate buffer. Subsequently, the specimens were dehydrated in an alcohol series and embedded into Epon.

3.3. CELL STAINING

Light microscopy: 1 micrometer sections were collected on object glasses and stained with Richardson staining solution (self-made, Mixture of "azuur", methylene blue and borax) and mounted with Epon resin.
TEM: Ultrathin sections were collected on formvar-coated grids and counterstained with uranil acetate and lead citrate. Images were acquired using a Transmission electron microscope Philips Cm-10 at several magnifications.

3.4. LIGHT MICROSCOPY ANALYSIS

We made 2 stained sections per sample of 1 micrometer thickness, amounting to 24 samples: 6 male cross-linked, 6 male non-cross-linked, 6 female cross-linked, and 6 female non-cross-linked. Per sample, the clearest section was chosen for analysis, which led to a total of 12 samples. All multinucleated cells were counted on each of the 12 samples. The mono-nucleated cells which were adjacent to the collagen were also counted as FBGC, for it is possible that one or more nuclei are not visible due to the cutting of the sample. Cell surface area and collagen density were analyzed using ImageJ software.

3.5. TEM ANALYSIS

Several ultrathin sections per sample were analyzed using transmission electron microscopy. Images were acquired using iTEM (Olympus) and were assessed on several locations per sample. All images were assessed on whether collagen degradation is visible, on cell interactions, and other distinct morphological features.

3.6. STATISTICAL ANALYSIS

Statistical analysis was performed in order to determine differences in male versus female mice and cross-linked versus non-cross-linked collagen. The data was analyzed with statistical software (GraphPad Prism, version 5.00, GraphPad Software Inc.). Significance of differences was evaluated by paired t tests and unpaired t tests with Welch's correction. Statistical difference was defined as having a p-value < 0.05.

4. RESULTS

4.1. FBGC FORMATION NEAR CROSS-LINKED AND NON-CROSS-LINKED COLLAGEN.

Formation of FBGCs was observed in all samples. The formation of FBGCs was assessed microscopically by counting all FBGCs formed on each sample and measuring the collagen area. Values represent the number of FBGCs per 100% collagen (i.e. the number of FBGCs divided by the collagen density of the sample). More FBGCs were formed on the samples of cross-linked collagen, compared to non-cross-linked collagen (Fig. 1). Per 100% collagen, an average of 346.4 (± 286.7) FBGCs were found on cross-linked-collagen, compared to 125.0 (± 113.5) FBGCs on non-cross-linked collagen. Cross-linked samples derived from a specific mouse, all showed more FBGCs than the non-cross-linked sample in the same mouse (Fig. 2).
One out of the six mice was excluded from the analysis, because we were unable to count the cells on one of the samples derived from this mouse. All multinucleated cells were included in the FBGC count. Furthermore, all mononuclear cells adjacent to collagen were also counted as FBGCs in all samples, for it is probable that only one nucleus is visible, while in reality more nuclei are present.

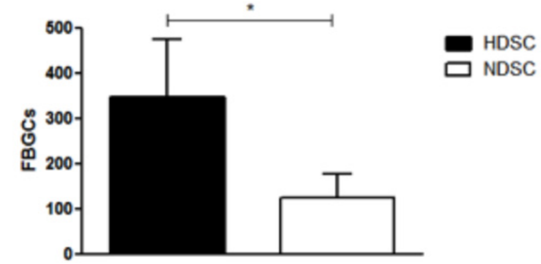


Fig. 1. *FBGC formation near HDSC and NDSC. More FBGCs were formed in the presence of HDSC than NDSC. * p<0.05*

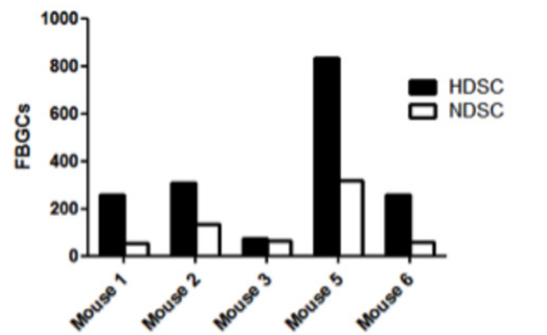


Fig. 2. *FBGC formation near HDSC and NDSC per mouse. Within every sample, more FBGCs were formed near HDSC than NDSC. A high variance between samples was observed.*

4.2. FBGC FORMATION IN MALE AND FEMALE MICE

No differences in number of FBGCs were observed between male and female mice. Although the mean in male mice was higher (307.6 ± 316.1) than in female mice (147.5 ± 108.9) (Fig. 3.), this difference is not significant due to the high variance in the male sample (Fig. 2). We excluded one sample from the assessment, because we were not able to count the cells on this sample.

4.3. COLLAGEN DENSITY

Collagen density was calculated as percentage collagen of the whole surface area of the sample. The collagen density of the cross-linked samples (34.6% ± 7.7%) showed no difference with the density of the non-cross-linked samples (36.1% ± 7.5%) (Fig. 4). No differences were found between collagen densities of male samples (34.2% ± 9.5%) and female samples (35.5% ± 5.1%) (Fig. 5). Ten out of twelve samples showed a lower collagen density at the border of the sample than in the center (Fig. 6). We did not statistically test for this difference.

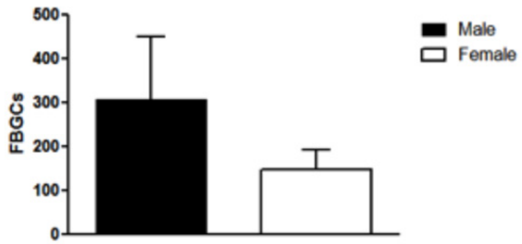


Fig.3 *FBGC formation in male and female mice. No differences were found between male and female mice regarding FBGC formation.*

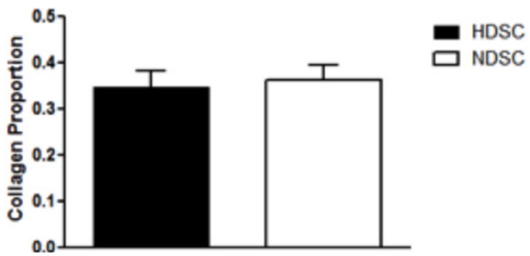


Fig.4 *Collagen density of HDSC and NDSC samples. No differences were found in collagen density between HDSC and NDSC samples.*

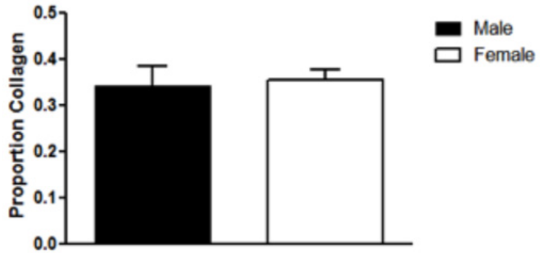
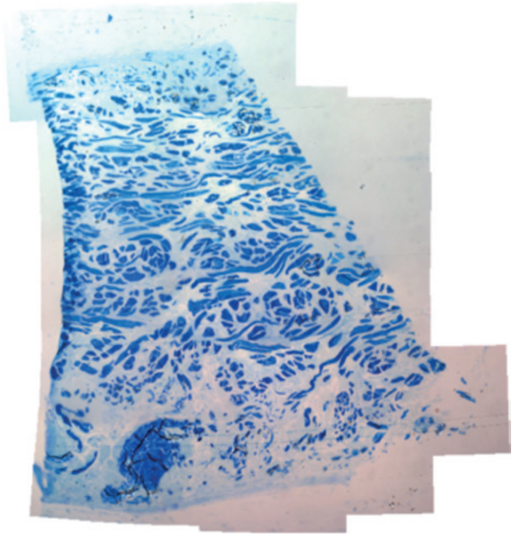
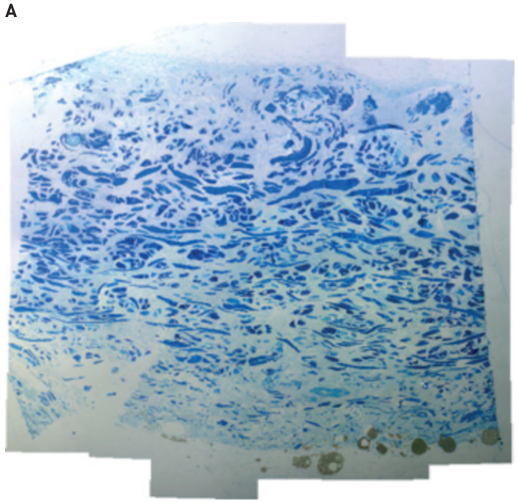


Fig.5 *Collagen density of male and female samples. No differences were found in collagen density between male and female samples.*



A



B

Fig.6 *Collagen density at the border of the sample is lower than in the center. (A) HDSC implanted in a female mouse. (B) HDSC implanted in a male mouse. Magnification=40×.*

4.4. FBGC MORPHOLOGY NEAR CROSS-LINKED AND NON-CROSS-LINKED COLLAGEN

FBGCs were observed in eleven out of twelve samples. The sample that not included FBGCs was excluded from analysis. In all cross-linked samples, and three out of five non-cross-linked samples finger-like structures on the FBGCs were present (Fig.7). These structures were rarely observed at the site of the neighboring collagen, but often at the opposite side of the FBGC (Fig. 8). We observed intertwining fingers of different FBGCs in all cross-linked samples, but in only one of five non-cross-linked samples. All cross-linked samples showed loosened collagen particles near the FBGC (Fig. 9), which is an indication for collagen degradation. Two of the five non-cross-linked samples showed this loosened collagen. In all samples the FBGCs partly or fully engulfed the collagen (Fig. 8). Four samples (all cross-linked) showed complete encapsulation of collagen parts (Fig. 10). The collagen fibrils appear loosened within the encapsulation, demonstrating degradation (Fig 11).

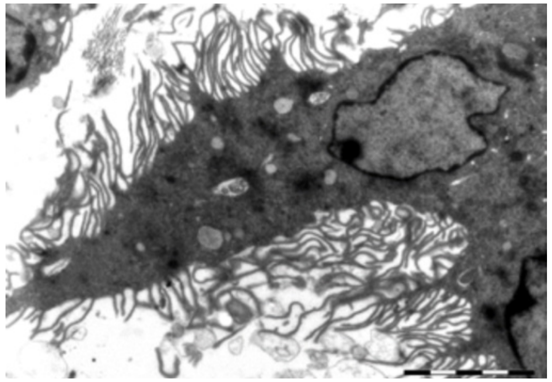


Fig. 7. *Finger-like structures are present on the border of the FBGCs. These structures were formed in the presence of NDSC in a male mouse.*

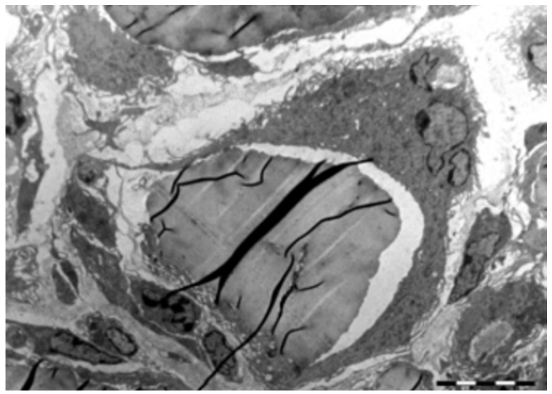


Fig. 8. *Partial engulfment of HDSC by a FBGC in a male mouse. Finger-like structures are not present at the collagen-FBGC interface, but on the opposite side of the cell.*

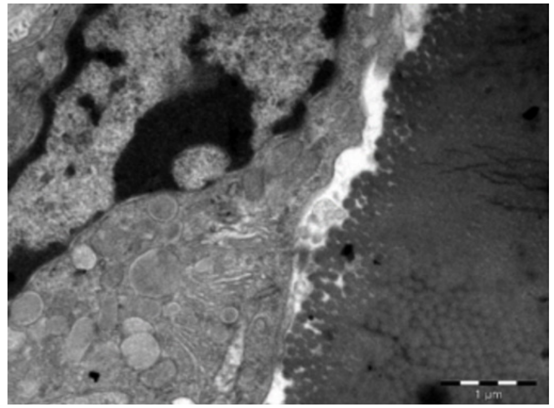


Fig. 9. *FBGC interaction with NDSC in a female mouse. FBGC are capable of loosening the collagen fibrils are the cell-collagen interface. This loosening is characteristic for collagen degradation.*

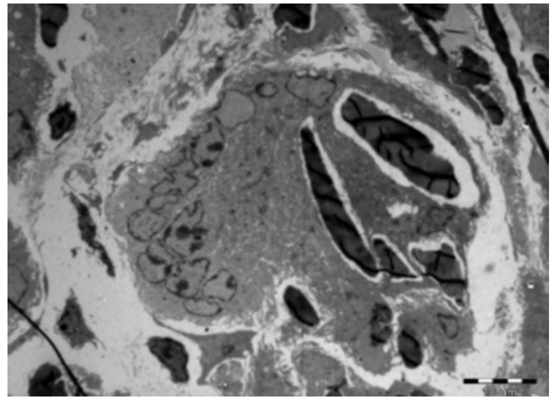


Fig. 10. *Complete engulfment of HDSC by FBGC in a male mouse. FBGCs are capable of completely engulfing collagen.*

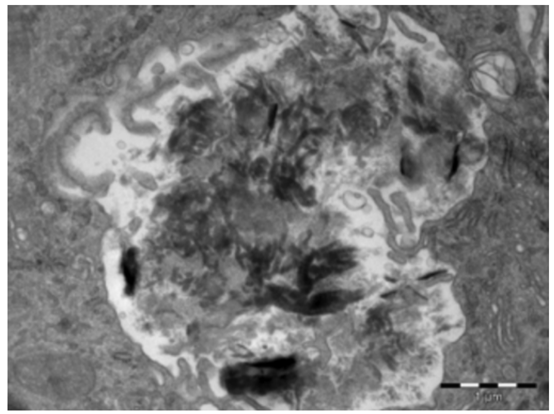


Fig. 11. *NDSC degradation after complete engulfment. After surrounding the collagen fibrils, FBGCs are capable of extracellular degradation of these fibrils.*

5. DISCUSSION

Due to the controversies and uncertainties about FBGC formation in the presence of differently cross-linked collagen, the aim of our study was to investigate the role of collagen cross-linking on FBGC formation and morphology in vivo.

5.1. HDSC ELICITS THE FORMATION OF MORE FBGCS THAN NDSC

A possible explanation of the finding that HDSC leads to the formation of more FBGCs than NDSC is that HDSC disks are more difficult to degrade due to the cross-linking. This is in line with the current hypothesis that materials which are difficult to degrade elicit a more severe FBR. One explanation for this phenomenon could be that frustrated phagocytosis occurs more in these cross-linked materials, which was hypothesized to be a driving force of multi-nucleation[12]. Because multiple studies have shown that macrophage fusion, and hence FBGC formation, is dependent of the surface proteins of the biomaterials[6,29,31], it is possible that the surface proteins of the two different disks are unalike. More research is necessary to determine these possible differences.

Secondly, it is interesting to note that large differences in FBGC formation were observed between similar samples. For instance, 88 FBGCs were observed on a cross-linked sample of mouse 2, while only 25 FBGCs were observed on a cross-linked sample of mouse 3. Corrected for collagen density and cell area, the difference is even fourfold: 308 in mouse 2 versus 75 in mouse 3. Both mouse 2 and 3 were female mice. This shows that huge differences exist between mice, when all other conditions are the same. Because FBGC formation (and macrophage fusion) is regulated by so many factors, these differences were inevitable. Again, heterogeneity in biomaterial surface proteins probably plays a significant role in the differences between the two mice.

Lastly, our findings contradict studies previously conducted. In 2010 and 2011, the research group led by Ruud Bank published 2 studies; the first comparing FBGC formation near HDSC and GDSC[22] and the second study comparing FBGC formation near NDSC and gelatin disks[32]. The first study showed giant cell formation of just under 400 FBGCs per mm^2 near HDSC at day 21[22]. The second study demonstrated that a similar number of FBGCs (just shy of 400) were formed in the presence of NDSC at day 21[32]. Both studies implanted the disks subcutaneously in mice. The disks had equal size and were obtained from the same location. Operation procedures and histological techniques were equal is both studies. Although we obtained our samples from the same batch of collagen disks, our study did find a significant difference between HDSC and NDSC at day 21.

5.2. FBGC FORMATION IS NOT SEX DEPENDENT

The inflammatory response after implantation of a biomaterial can be modulated by hormones. Lescure et al. demonstrated that implants containing progesterone, a sex hormone, delayed the healing process[33]. Foreign body induced sarcomas also differed between male and female mice[34]. This suggests that sex differences might exist in the FBR. The fact that our study showed no differences between male and female mice in either FBGC formation or collagen density is concordant with the findings of Dalu et al. who investigated sex differences more thoroughly: he showed that the infiltration of neutrophils and macrophages did not differ significantly in male and female mice[35]. So although sex differences are probably present in other inflammatory processes, it seems that they do not influence FBGC formation after collagen implantation in mice.

5.3. COLLAGEN DENSITY DOES NOT DIFFER IN HDSC AND NDSC DISKS

Although one could imagine that non-cross-linked collagen is more easily degraded and non-cross-linked samples would show lower collagen density, we did not find this difference. Because some controversy exists on whether FBGCs are in essence capable of extensive collagen degradation, it is possible that only minimal amounts of collagen were degraded. We did not study the extent of collagen degradation, only whether it was microscopically visible, so it is possible that the degradation only occurred in tiny amounts on the collagen surface. It is possible that no significant part of the collagen was degraded in either HDSC or NDSC setting. This could explain the similarities in collagen densities. Furthermore, the two studies conducted by Ye et al. also found similar degradation of HDSC and NDSC: little to none[22,36]

The fact that the collagen density at the border of the samples is lower than in the center could be an indication of more collagen breakdown at the border than in the center. This could be explained by the fact

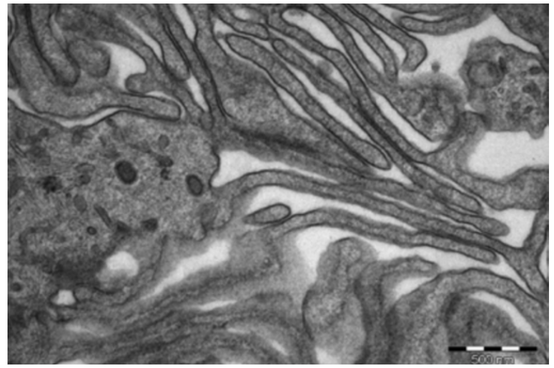


Fig. 13. *Intertwining of finger-like structures near HDSC in a female mouse. Although the fingers can intertwine, no characteristics of fusion seem present.*

that the collagen degrading cells infiltrate the tissue from the outside towards the inside of the sample. Because we showed on an electron microscopical level that FBGCs were capable of collagen degradation, it is possible that they are responsible for de difference between the sample border and center. Interestingly, this phenomenon was visible in both cross-linked and non-cross-linked samples, even though we showed that more FBGCs are present in cross-linked samples. We did not quantify the differences in the number of FBGCs at the border and in the center, so we cannot make any conclusive statements on the role of the FBGCs in this difference. Another possibility of the difference between border and center of the sample might be a different alignment of the collagen fibrils.

5.4. FBGCS HAVE DISTINCT FINGER-LIKE STRUCTURES ON THE CELL BORDER

To our knowledge, the finger-like structures we found on many FBGCs have not been described in the literature before. The role of these finger-like structures can only be hypothesized. Initially we believed that the structures were involved in cell-cell fusion, because the fingers of different cells were often intertwined (Fig. 12). Images on a higher magnification, however, showed that the fingers never fuse, but that they are in contact (Fig. 13). It could be the case, however, that the intertwining of the fingers is a method used by the FBGCs to cling to each other and in that way facilitate cell fusion. Almost without exception, FBGCs that adhere to collagen show the finger-like structures. Studies have shown that adhesion to the biomaterial and interaction with the surface proteins is necessary for cell-cell fusion[26,37]. Hence the fact that finger-like structures are present on adherent FBGCs might suggest that finger formation is a part of the fusion process, regulated by surface proteins. In order to prove this, these cells should be compared to FBGCs not adherent to collagen. However, these cells could not be microscopically assessed on the presence of finger-like structures, because they were free floating in the medium which was discarded. The fact that some cell not adhering to collagen also show finger-like structures (Fig. 14) could be because they have just 'released' the collagen, or adhere to collagen in the vertical plane, while we only see the horizontal plane. More research is required to determine a possible relationship between finger formation and fusion.
We discarded the idea that the fingers were involved in the degradation process, because the FBGCs did not show fingers near the collagen, but on the opposite side of the cell (Fig. 8). Currently, we are working on creating 3D tomography images of the finger-like structures in order to get a better understanding of their morphology. Preliminary results show a plate-like

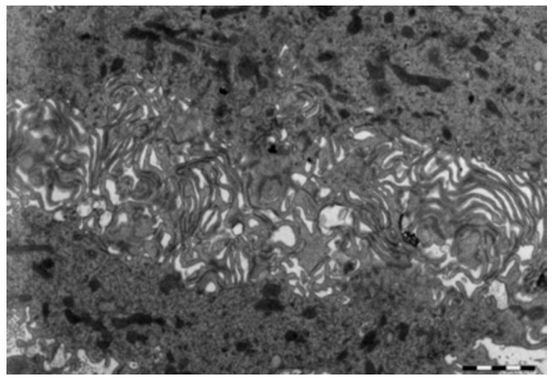


Fig. 12. *The finger-like structures of different FBGCs can intertwine.*

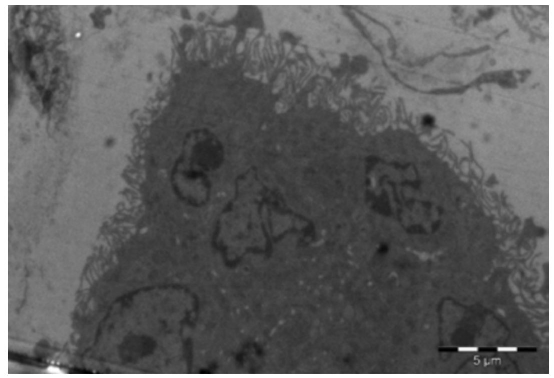


Fig. 14. *The formation of finger-like structures in the absence of collagen in a female mouse.*

shape of the fingers rather than

a cylindrical shape, as we previously assumed. More imaging studies are required to get a better grasps of the functionality of the finger-like structures.

5.5. FBGCS DISPLAY ENGULFMENT OF COLLAGEN, BUT NO INTERNALIZATION.

The collagen engulfment capabilities of FBGCs were previously demonstrated by TEM studies conducted by van Wachem et al.[13] They showed that FBGCs are capable of both intra- and extracellular degradation of collagen. While our study demonstrated the extracellular degradation by FBGCs, indications of intracellular degradation were not observed. This is possibly due to the absence of Endo180 and membrane bound metalloproteinase MT1-MMP in NDSC and HDSC. It has been proposed that collagen internalization receptor Endo180 plays an important role in the FBR, and mediates collagen internalization and degradation[32,38]. Ye et al. hypothesized that co-expression of Endo180 and MT1-MMP is necessary for collagen internalization. They demonstrated that collagen phagocytosis was not present in HDSC and attributed that to the fact that Endo180 expression in HDSC was absent[38]. In a different study, Ye et al. showed that after NDSC implantation, usually no co-expression of MT1-MMP and Endo180 is observed[32]. Furthermore, they showed that interferon-gamma (IFN-Ð) induced Endo180 and MT1-MMP expression[38]. Hence they concluded that collagen phagocytosis and subsequent internalized degradation was INF-γ dependent, and low levels of INF-γ were responsible for absence of collagen internalization in HDSC.

Because our study also failed to show collagen internalization in HDSC samples, we could hypothesize that low IFN-γ is responsible for this observation. However, Khouw et al. showed that monoclonal inhibition of IFN-γ delayed the formation and function of FBGC in the presence of dermal sheep collagen[39]. This could suggest high levels of IFN-γ in our HDSC samples, for they showed many FBGCs. However, IFN-γ is only one among many factors that influence FBGC formation and therefore low interferon-gamma levels could still be an explanation for the absence of internal collagen degradation.

Hence, even though we observed that FBGCs can completely engulf collagen (Fig. 10 & Fig. 11), we believe this is also a component of extracellular degradation, rather than phagocytosis, due to the absence of Endo180 and/or MT1-MMP. This is supported by the fact that collagen fibrils are already loosening while engulfed, demonstrating degradation takes place extracellularly (Fig. 11).

6. CONCLUSIONS

Foreign body giant cell (FBGC) formation in the presence biomaterials is a well-documented occurrence. Here we report that the FBGC formation near HMDIC cross-linked dermal sheep collagen disks (HDSC) and non-cross-linked dermal sheep collagen disks (NDSC) showed significant differences. At day 21 after subcutaneous implantation in mice, HDSC elicited the formation of more FBGCs than NDSC. No differences were observed in male and female mice. The driving force behind macrophage fusion into FBGCs is hypothesized to be the inability to break down the biomaterial, but this topic needs further research. We demonstrated that FBGCs are capable of extracellular degradation of both HDSC and NDSC on a transmission electron microscopical level. Internalization of collagen was not observed. To our knowledge, this is the first study to describe the finger-like structures of which the function is yet to be discovered. We conclude that cross-linking of a biomaterial elicits a more severe foreign body response that manifests itself by the formation of more FBGCs. By elucidating the FBGC behavior and morphology in the presence of different biomaterials we increase our understanding of the foreign body response and its implications on biocompatibility of biomedical devices.

ACKNOWLEDGEMENTS

This study was facilitated by the Academic Centre for Dentistry Amsterdam (ACTA) and the Academic Medical Centre, Amsterdam (AMC). The author wishes to thank Mr. Bas ten Harkel and Prof. Vincent Everts for their excellent support and guidance during this study. Both were an inspiration and a constant source of knowledge. Furthermore, the author would like to thank Mrs. Daisy Picavet and Mr. Henk van Veen for their outstanding technical support at the Academic Medical Centre. Many thanks to Tjardo Postma, who read my revisions and was a constant sounding board throughout the writing process. Lastly, the author thanks the colleagues at the ACTA, in particular Mr. Ton Schoenmaker and Dr. Teun de Vries, who were always willing to help out where possible.

REFERENCES

1. Badylak SF, Gilbert TW. Immune response to biologic scaffold materials. Seminars in immunology. 2008 April;20(2):109–16.

2. Anderson JM, Rodriguez A, Chang DT. Foreign body reaction to biomaterials. Seminars in immunology. 2008 April;20(2):86–100.

3. Tang L, Jennings T a, Eaton JW. Mast cells mediate acute inflammatory responses to implanted biomaterials. Proceedings of the National Academy of Sciences of the United States of America. 1998 July 21;95(15):8841–6.

4. McNally AK, Anderson JM. Interleukin-4 induces foreign body giant cells from human monocytes/ macrophages. Differential lymphokine regulation of macrophage fusion leads to morphological variants of multinucleated giant cells. The American journal of pathology. 1995 November;147(5):1487–99.

5. DeFife K, Jenney C. Interleukin-13 induces human monocyte/macrophage fusion and macrophage mannose receptor expression. The Journal of Immunology. 1997;158(7):3385–3390.

6. Chen EH, Grote E, Mohler W, Vignery A. Cell-cell fusion. FEBS letters. 2007 May 22;581(11):2181–93.

7. Neale SD, Athanasou NA. Cytokine receptor profile of arthroplasty macrophages, foreign body giant cells and mature osteoclasts. Acta orthopaedica Scandinavica. 1999 October;70(5):452–8.

8. HernandezÐPando R. Inflammatory cytokine production by immunological and foreign body multinucleated giant cells. Immunology. 2000;100(3):352–359.

9. Henson P. The immunologic release of constituents from neutrophil leukocytes I. The role of antibody and complement on nonphagocytosable surfaces or phagocytosable particles. The Journal of Immunology. 1971;107(6):1535–1546.

10. Henson P. The immunologic release of constituents from neutrophil leukocytes II. Mechanisms of release during phagocytosis, and adherence to nonphagocytosable surfaces. The Journal of Immunology. 1971;107(6):1547–1557.

11. Anderson J. Biological responses to materials. Annual Review of Materials Research. 2001;(31):81–110.

12. McNally AK, Anderson JM. Macrophage fusion and multinucleated giant cells of inflammation. Advances in experimental medicine and biology. 2011 January;713:97–111.

13. Van Wachem PB, van Luyn MJ, Nieuwenhuis P, Koerten HK, Olde Damink L, Ten Hoopen H, Feijen J. In vivo degradation of processed dermal sheep collagen evaluated with transmission electron microscopy. Biomaterials. 1991 March;12(2):215–23.

14. Van der Laan JS, Lopez GP, van Wachem PB, Nieuwenhuis P, Ratner BD, Bleichrodt RP, Schakenraad JM. TFE-plasma polymerized dermal sheep collagen for the repair of abdominal wall defects. The International journal of artificial organs. 1991 October;14(10):661–6.

15. Jarman-Smith ML, Bodamyali T, Stevens C, Howell JA, Horrocks M, Chaudhuri JB. Porcine collagen crosslinking, degradation and its capability for fibroblast adhesion and proliferation. Journal of materials science. Materials in medicine. 2004 August;15(8):925–32.

16. Ratner BD. Biomaterials Science: An Introduction to Materials in Medicine. Academic Press; 2004 pp. 1–851.

17. Weadock KS, Miller EJ, Bellincampi LD, Zawadsky JP, Dunn MG. Physical crosslinking of collagen fibers: comparison of ultraviolet irradiation and dehydrothermal treatment. Journal of biomedical materials research. 1995 November;29(11):1373–9.

18. Charulatha V, Rajaram A. Influence of different crosslinking treatments on the physical properties of collagen membranes. Biomaterials. 2003 February;24(5):759–67.

19. Luyn MJA, Wachem PB, Dijkstra PJ, Olde Damink LHH, Feijen J. Calcification of subcutaneously implanted collagens in relation to cytotoxicity, cellular interactions and crosslinking. Journal of Materials Science:

Materials in Medicine. 1995;6(5):288–296.
20. Olde Damink L, Dijkstra P, Luyn M Van, Wachem PB, Nieuwenhuis P, Feijen J. Crosslinking of dermal sheep collagen using hexamethylene diisocyanate. Journal of Materials Science: Materials in Medicine. 1995;6:429–434.

21. Luttikhuizen DT, van Amerongen MJ, de Feijter PC, Petersen AH, Harmsen MC, van Luyn MJA. The correlation between difference in foreign body reaction between implant locations and cytokine and MMP expression. Biomaterials. 2006 December;27(34):5763–70.

22. Ye Q, Harmsen MC, van Luyn MJ a, Bank RA. The relationship between collagen scaffold cross-linking agents and neutrophils in the foreign body reaction. Biomaterials. 2010 December;31(35):9192–201.

23. Tang L, Eaton J. Natural responses to unnatural materials: A molecular mechanism for foreign body reactions. Molecular medicine. 1999;5:351–358.

24. McNally AK, Anderson JM. Foreign body-type multinucleated giant cell formation is potently induced by alpha-tocopherol and prevented by the diacylglycerol kinase inhibitor R59022. The American journal of pathology. 2003 September;163(3):1147–56.

25. Dadsetan M, Jones JA, Hiltner A, Anderson JM. Surface chemistry mediates adhesive structure, cytoskeletal organization, and fusion of macrophages. Journal of biomedical materials research. Part A. 2004 December 1;71(3):439–48.

26. Yagi M, Miyamoto T, Sawatani Y, Iwamoto K, Hosogane N, Fujita N, Morita K, Ninomiya K, Suzuki T, Miyamoto K, et al. DC-STAMP is essential for cell-cell fusion in osteoclasts and foreign body giant cells. The Journal of experimental medicine. 2005 August 1;202(3):345–51.

27. DeFife KM, Jenney CR, Colton E, Anderson JM. Disruption of filamentous actin inhibits human macrophage fusion. The FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 1999;13:823–832.

28. Jones J, McNally A. Matrix metalloproteinases and their inhibitors in the foreign body reaction on biomaterials. Journal of biomedical materials research Part A. 2008;84(1):158–166.

29. MacLauchlan S, Skokos EA, Meznarich N, Zhu DH, Raoof S, Shipley JM, Senior RM, Bornstein P, Kyriakides TR. Macrophage fusion, giant cell formation, and the foreign body response require matrix metalloproteinase 9. Journal of leukocyte biology. 2009 April;85(4):617–26.

30. Chizzolini C, Rezzonico R. IL-4 enhance matrix metalloproteinase-1 (MMP-1) while decreasing MMP-9 production by granulocyte-macrophage colony-stimulating factor-differentiated human. The Journal of Immunology. 2000;164(11):5952–5960.

31. Vignery A. Osteoclasts and giant cells: macrophage-macrophage fusion mechanism. International journal of experimental pathology. 2000 October;81(5):291–304.

32. Ye Q, Harmsen MC, Ren Y, Bank RA. The role of collagen receptors Endo180 and DDR-2 in the foreign body reaction against non-crosslinked collagen and gelatin. Biomaterials. 2011;32:1339–1350.

33. Lescure F, Gurny R, Doelker E, Anderson JM. Long-term histopathological study of new polypeptidic biomaterials. Biomaterials. 1992 January;13(14):1009–1011.

34. Michelich VJ, Brand KG. Effects of Gonadectomy on Foreign-Body Tumorigenesis in CBA/H Mice. J Natl Cancer Inst. 1980 April 1;64(4):807–808.

35. Dalu A, Blaydes BS, Lomax LG, Delclos KB. A comparison of the inflammatory response to a polydimethylsiloxane implant in male and female Balb/c mice. Biomaterials. 2000 October;21(19):1947–1957.

36. Ye Q, Harmsen MC, Ren Y, Bank RA. The role of collagen receptors Endo180 and DDR-2 in the foreign body reaction against non-crosslinked collagen and gelatin. Biomaterials. 2011;32:1339–1350.

37. Dittmar T, Zänker K. Cell Fusion in health and disease.; 2011.

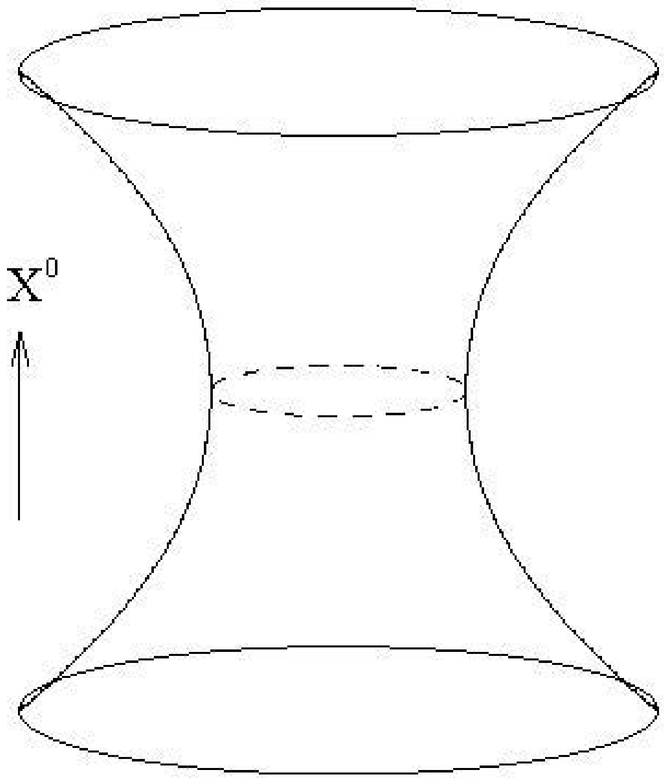
38. Ye Q, Xing Q, Ren Y, Harmsen MC, Bank RA. Endo180 and MT1-MMP are involved in the phagocytosis of collagen scaffolds by macrophages and is regulated by interferon-gamma. European cells & materials. 2010 January;20:197–209.

39. Khouw IMSL, Wachem PB Van, Leij LFMH De, Luyn MJA Van. Inhibition of the tissue reaction to a biodegradable biomaterial by monoclonal antibodies to IFN-gamma. Journal of biomedical materials research. 1997;41(2):202–210.

The Infrared Stability of de Sitter Space in the Presence of a Scalar Field

Danu Thung

Supervisors: Prof. dr. J. de Boer (UvA)



ABSTRACT

In this thesis, we study a recent proposal to solve the infamous cosmological constant problem. It has been claimed by Polyakov that a massive, scalar field in de Sitter space will give rise to an explosive production of infrared particles, inducing an instability of spacetime itself. The contribution to the energy-momentum tensor from these particles could counteract the effect of a high cosmological constant, screening it down to a low effective value. This mechanism could provide an explanation as to why we observe a cosmological constant which is many, many orders of magnitude smaller than a naive theoretical estimate would suggest.

Our aim is to understand the content of Polyakov's claim in a quantitative fashion. Although we assume previous knowledge of the fundamentals of quantum field theory and general relativity, much of this work is devoted to introducing the required tools and concepts to understand and perform calculations within the framework of quantum eld theory on a de Sitter space background. This includes a review of the particle interpretation in curved spacetimes, a discussion of the solutions to the equation of motion and the two point functions as well as an introduction to the Keldysh-Schwinger technique for non-equilibrium quantum field theoretical calculations.

After covering the necessary preparatory material, we reproduce the main results that have been obtained thus far by those who support Polyakov's position on the (in)stability of de Sitter space. In our considerations, we restrict ourselves to the expanding Poincare patch of de Sitter space. This choice is motivated by the observation that, although our universe is modeled well by the expanding section of de Sitter space, there seems to be no conclusive evidence that we live in a 'globally' de Sitter universe. We find that there is substantial evidence that interesting infrared effects take place in the expanding Poincare patch. However, it is not entirely clear how to correctly interpret the results. We conclude that it seems too early to claim with reasonable certainty that de Sitter space carries the seeds of its own destruction.

FOREWORD

The topic of this thesis falls within the category of quantum field theory (QFT) in curved spacetime or semiclassical gravity. Therefore, familiarity with (scalar) QFT and general relativity (GR) will be assumed to avoid having to cover extremely large amounts of preparatory material. In the introduction to this thesis I will point out a number of textbooks that could be used to obtain the prerequisite knowledge. Furthermore, a basic understanding of complex analysis and differential equations is required. This text should be accessible to graduate students who have taken at least one course in both QFT and GR, and perhaps undergraduates in their senior year.

A few remarks on the notation and conventions used in this work: Since the topic of this thesis is part of relativistic quantum theory, we will work in reduced Planck units where $\hbar = c = \sqrt{8\pi G} = 1$ unless otherwise indicated. We will use metric signature – + + +...so that the Minkowski metric reads $\eta_{\mu\nu}$= diag(–1, 1, 1, 1, ...). Tensor notation and the Einstein summation convention will be employed throughout, with Greek indices (e.g. $\mu, \upsilon, \rho, \ldots$) running over all spacetime labels 0, 1, 2, 3,...N – 1 in an N-dimensional spacetime while Latin indices (e.g. $i, j, k, \ldots$) run over spatial labels 1, 2, 3, ...N – 1 only. Sometimes the arrow notation will be employed for spatial vectors for aesthetic reasons. Partial derivatives are denoted by ∂_μ and covariant derivatives by ∇_μ. We will occasionally use the common notational shorthand $\Xi\, g^{\mu\upsilon}\partial^\mu\partial^\upsilon$

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

During the twentieth century, spectacular progress was made in uncovering the essential physics at both the smallest and largest scales of our universe. Despite these successes, there are still many unanswered questions. Some of the most important theoretical questions lie in the domain where both quantum physics and gravitational effects are expected to play an important role. In the absence of a fully developed theory of quantum gravity, these problems can be attacked by performing quantum field theoretical calculations on a fixed, curved spacetime background or using semiclassical methods, which take the backreaction on the spacetime into account and can be seen as a first approximation to the full quantum theory.

In this context, the solution to Einstein's equations known as 'de Sitter spacetime' (after Dutch physicist Willem de Sitter) is probably the single most interesting and important spacetime. It possesses great simplicity, being one of very few maximally symmetric solutions to Einstein's equations. Moreover, cosmological data from supernovae shows that our universe is reasonably well approximated by de Sitter space [1]. Furthermore, inflationary theory, which has recently received experimental support from the BICEP2 experiment [2], asserts that the early universe went through a de Sitter-like epoch during which it underwent exponential expansion, driven by a large (quasi-)constant energy density. Thus, there is a strong cosmological motivation for investigating the behavior of quantum fields in a de Sitter universe.

1.2 MOTIVATION

In this thesis, we will study one particular aspect of QFT in de Sitter space, namely the 'infrared' (i.e. long-wavelength) behavior of a massive, scalar quantum field. The motivation for doing this is provided by the (in)famous cosmological constant problem. It is conventionally considered natural for the cosmological constant|regarded as the energy density of the vacuum|to be set by the 'ultraviolet' (i.e. short-wavelength) cuto. This cutoff is usually taken to be at the Planck scale, leading to an expected value for the cosmological constant on the order of $E_{\mathrm{Planck}} L^{-3}_{\mathrm{Planck}}$ (in SI units). The observed value of the cosmological constant, however, is lower by a factor of approximately 10^{122}.

Several solutions to this *vacuum catastrophe* have been proposed (e.g. Weinberg's `anthropic principle' [3]). Some have argued that the value of the cosmological constant should be set by infrared physics, rather than a UV cuto. In particular, infrared effects may provide a dynamical mechanism for screening a high cosmological constant down to a low effective value that we observe today. For a review that describes a number of distinct approaches to dynamical screening of the cosmological constant, the interested reader is referred to [4].

This thesis explores one incarnation of this idea, namely dynamical screening of the cosmological constant due to the explosive production of massive spin-zero particles. This proposal seems to have garnered a signicant amount of support since recent work on this topic was carried out by the prominent theoretician Dr. A. M. Polyakov from Princeton University, who claims to have shown that de Sitter space is fundamentally unstable [5{14]. On the other hand, there are a number of authors that claim that there is no such thing as an instability of de Sitter space (e.g. [15-18] and references therein). In this thesis we aim to form a clear picture of the content and implications of the claims put forward by Polyakov and supporters.

1.3 GENERAL OUTLINE

A large part of this thesis is dedicated to developing the necessary concepts and tools to deal with the problems at hand. The discussion offered in this work will largely follow Mukhanov and Winitzki's textbook [19] and a set of lecture notes by Akhmedov on scalar field theory in de Sitter space [20], although we have drawn from numerous other sources as well.

First, we outline some of the basic features of de Sitter space that we will need later on. Then, we will briefly review the particle interpretation of QFT in Minkowski space and outline the difficulties that the concept of a particle runs into when working in curved spacetime. Chapter 4 treats scalar quantum fields in a de Sitter background.

The equations of motion and their solutions in the Poincare patches and global de Sitter space are discussed. We also consider two point functions in both the Poincare patches and global de Sitter space. In chapter 5 a brief introduction to the Keldysh-Schwinger (also known as 'in-in') formalism for performing calculations within the context of non-equilibrium QFT is presented. In Appendix A, which the interested reader should read before chapter 6, we present an introduction to semiclassical field theory with a derivation of the semiclassical field equations of gravity. This completes our preparation for a number of important calculations that lie at the heart of the question of the stability of de Sitter space. These are shown in chapter 6. In the final section of that chapter, we summarize our findings and comment on the interpretation of the results.

1.4 RECOMMENDED READING MATERIAL

Finally, some guidance for preparatory reading material. For an introduction to the fundamentals of GR and QFT, both of which will be assumed as previous knowledge in this work, we refer the reader to a more or less random sample of the many available textbooks, which each cover all that is needed for our purposes and much more [21-27]. It is recommended to draw from a number of different sources instead of relying on a single textbook, especially when studying QFT. A thoroughly enjoyable general introductory text about QFT in curved spacetime is Mukhanov and Winitzki's 'Introduction to Quantum Effects in Gravity' [19], while Birrell and Davies' classic textbook [28] provides a more in-depth treatment for the interested reader. Knowledge of these works is not required to read and understand this thesis.

CHAPTER 2

DE SITTER GEOMETRY

N-dimensional de Sitter space can be constructed as the surface in $N+1$-dimensional Minkowski space satisfying

$$\eta_{\mu\nu} X^\mu X^\nu = -(X^0)^2 + (X^j)^2 = H^{-2} \tag{2.1}$$

where H is Hubble's constant. This is schematically shown in fig. 2.1. From its construction, it is clear that SO(N; 1), the Lorentz group of the ambient Minkowski space, is the isometry group of de Sitter space. Alternatively, one can view de Sitter space as simply the solution to Einstein's equations in the vacuum with a positive cosmological constant:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \tag{2.2}$$

Contracting Einstein's equations yields

$$\Lambda = \frac{N-2}{2N} R \tag{2.3}$$

It follows from the N-dimensional Friedmann equation [29] that we can express the cosmological constant in terms of H as

$$\frac{(N-1)(N-2)}{2} H^2 \tag{2.4}$$

so that we can express R, and H in terms of each other.

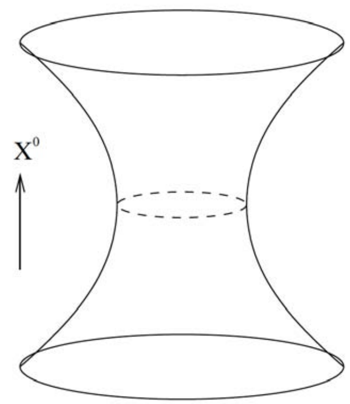


Figure 2.1: *A schematic graphical representation of the embedding of de Sitter space in a Minkowski space with one more dimension. All but two of the spatial dimensions are suppressed. Image taken from [30].*

2.1 GLOBAL COORDINATES

Using (2.1) we can induce a metric on de Sitter space. The equation clearly admits the following solution:

$$X^0 = H^{-1}\sinh(Ht), \qquad X^j = H^{-1}v^j\cosh(Ht) \tag{2.5}$$

with v^i an N-dimensional unit vector:

$$\begin{aligned}
v_1 &= \cos\theta_1 \\
v_2 &= \sin\theta_1\cos\theta_2 \\
v_3 &= \sin\theta_1\sin\theta_2\cos\theta_3 \\
&\vdots \\
v_{N-1} &= \sin\theta_1\sin\theta_2\ldots\sin\theta_{N-2}\cos\theta_{N-1} \\
v_N &= \sin\theta_1\sin\theta_2\ldots\sin\theta_{N-2}\sin\theta_{N-1}
\end{aligned}$$

All angles range from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$, except the last, which runs from $-\pi$ to π. The metric on the *N*-dimensional de Sitter space is obtained through what is known as the pullback of the metric on the ambient Minkowski space [21]. We find

$$ds^2 = -dt^2 + H^{-2}\cosh^2(Ht)d\Omega^2_{N-1} \tag{2.6}$$

This is known as the global de Sitter metric. From this construction, it is clear that the constant t slices of de Sitter are (*N* - 1)-dimensional spheres. We can find the Penrose diagram of global de Sitter space by performing the substitution

$$\cosh(Ht) \equiv \frac{1}{\cos\tau}$$

with $-\frac{\pi}{2} < \tau < \frac{\pi}{2}$, which transforms the metric to

$$ds^2 = \frac{1}{H^2\cos^2\tau}\left(-d\tau^2 + d\Omega^2_{N-1}\right) = \frac{1}{H^2\cos^2\tau}\left(-d\tau^2 + d\theta_1^2 + \sin^2\theta_1 d\Omega^2_{N-2}\right) \tag{2.7}$$

this is conformally equivalent to a spherically symmetric, static Universe. We can use the and 1 coordinates to draw the Penrose diagram (shown in figure 2.2) of de Sitter space, which contains all the information about its causal structure.

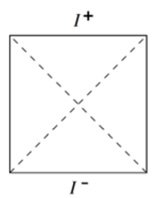


Figure 2.2: *In this Penrose diagram of global de Sitter space, each point represents a (N - 2)-dimensional sphere. τ increases vertically, while θ_1 increases left-to-right. The dashed lines represent the past and future horizons for an observer at $\theta_1 = \frac{\pi}{2}$: It can be seen that not the entire space is accessible. Image taken from [30].*

2.2 THE EXPANDING AND CONTRACTING POINCARE PATCHES

There are, of course, other possibilities when choosing coordinates on de Sitter space. One particularly useful solution is the following:

$$\begin{aligned}
HX^0 &= -\sinh(H\tau_+) + \frac{(Hx^i_+)^2}{2}e^{H\tau_+} \\
HX^i &= Hx^i_+ e^{H\tau_+} \qquad i = 1, 2, \ldots, N-1 \\
HX^D &= -\cosh(H\tau_+) + \frac{(Hx^i_+)^2}{2}e^{H\tau_+}
\end{aligned} \tag{2.8}$$

This results in the spatially flat metric [20]

$$ds^2_+ = -d\tau^2_+ + e^{2H\tau_+}d\vec{x}^2_+ \tag{2.9}$$

However, this only covers half of de Sitter space. In deriving this metric, we have silently imposed the condition $X^0 \geq X^N$ on the Minkowski coordinates, as can be easily checked. This part of de Sitter space is known as the expanding Poincare patch (EPP). The other half ($X^0 \leq X^N$), called the contracting Poincare patch (CPP), is covered by the coordinates

$$ds^2_- = -d\tau^2_- + e^{-2H\tau_-}d\vec{x}^2_- \tag{2.10}$$

In each patch we can introduce the conformal time coordinate $\eta_\pm \equiv H^{-1}e^{\mp H\tau_\pm}$ so that the metric is of the same form:

$$ds^2_\pm = \frac{1}{(H\eta_\pm)^2}\left(-d\eta^2_\pm + d\vec{x}^2_\pm\right) \tag{2.11}$$

Note that η_+ ranges from ∞ to 0 (!) while η_- ranges from 0 to ∞ as we let τ run from - ∞ to ∞. The EPP corresponds to the upper left half of the Penrose diagram of global de Sitter space (see fig. 2.2), while the CPP corresponds to the bottom right triangle. We can therefore obtain a Penrose diagram for both the Poincare patches from the global de Sitter diagram. This is done by 'unwrapping' the original diagram to double its width. The result is shown in fig. 2.3. For our purposes, the global and

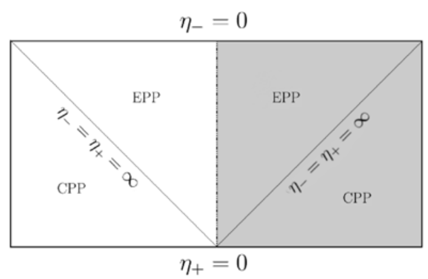


Figure 2.3: *A conformal diagram of the Poincare patches Note that η- = 0 and η+ = ∞ correspond to the innite past of the CPP and EPP respectively, while η- = 1 and η+ = 0 correspond to the innite future. The boundary between the Poincare patches therefore lies at η+ = η- = ∞ The shaded area is the original full de Sitter space diagram fig. 2.2. Image adapted from [20].*

Poincare patch metrics are all that is needed, although there are of course several other different coordinates that one can choose to study de Sitter space (see for instance [30] for a number of options).

2.3 PHYSICAL VERSUS COMOVING QUANTITIES

In the context of general relativity, there are two useful notions of spatial volume. The first is the comoving volume, which is simply the volume form $d^{N-1}V$ with respect to the spatial metric. This volume remains constant throughout time. On the other hand, the physical volume, which is defined as the comoving volume times the factors multiplying the spatial components of the metric, varies drastically with time. We can thus also distinguish between the physical and comoving density of particles in de Sitter space. As we will see in chapter 6, these two notions play an important role in the analysis of the (in)stability of de Sitter space. For other quantities, we can also distinguish between a comoving and a physical version. For instance, in the Poincare patches we can consider either the physical k or comoving momentum $k\eta$.

2.4 GEODESIC DISTANCE

Consider two points, x_1 and x_2, on an N-sphere of radius L, embedded in Euclidean space of dimension $N + 1$ (the embedding equation being $\delta_{ij}X^iX^j = L^2$). Then, the only $SO(N + 1)$ invariant quantity we can associate with this pair of points is the geodesic distance D, or equivalently the angle $\Theta = \frac{D}{L}$, between the points. We can then define z, which satisfies:

$$z(x_1, x_2) = \frac{\delta_{ij} x_1^i x_2^j}{|x_1||x_2|} = \cos\theta = \cos\frac{D}{L} \tag{2.12}$$

Although it is not as easily visualized, one can do the same thing and obtain an exactly analogous relation in N-dimensional de Sitter space embedded in $N + 1$-dimensional Minkowski space (intuitively, this makes sense when one views the situation as simply analytic continuation from the sphere in Euclidean space). We obtain

$$Z = Z(x_1, x_2) = \frac{\eta_{\mu\nu} x_1^\mu x_2^\nu}{|x_1||x_2|} = \cos(HD)$$

From now on, we will set $H = 1$ for convenience. Z will turn out to be a useful quantity in terms of which to express certain things. It can be explicitly worked out in any coordinate system. For instance, in global coordinates, we find

$$Z_{\text{global}}(x_1, x_2) = -\sinh(t_1)\sinh(t_2) + \cosh(t)\cosh(t_2)\vec{n}_1 \cdot \vec{n}_2$$

CHAPTER 3

THE PARTICLE INTERPRETATION IN CURVED SPACETIME

3.1 THE PARTICLE INTERPRETATION IN FLAT SPACE

The equation of motion for a real scalar field Φ in Minkowski space, i.e. the Klein-Gordon (KG) equation, reduces to the harmonic oscillator equation in Fourier space. Denoting the Fourier transform of Φ by Φk, we have

$$(\partial_t^2 + \omega_k^2)\phi_k = 0, \qquad \omega_k \equiv \sqrt{k^2 + m^2} \tag{3.1}$$

Working in the Heisenberg picture, the field Φ (x, t) can then be quantized by using the mode expansion:

$$\phi(x,t) = \int \frac{d^{N-1}k}{(2\pi)^{(N-1)/2}\sqrt{2\omega_k}} \left(a_k^- e^{ik_\mu x^\mu} + a_k^+ e^{-ik_\mu x^\mu}\right) \tag{3.2}$$

where $a^\pm$ are the creation and annihilation operators, which are multiplied by the plane wave solutions of the KG equation in Minkowski space. A crucial aspect of these plane wave 'mode functions' is that they unambiguously distinguish positive and negative frequencies. This decomposition is invariant under Lorentz transformations [21]. The equal-time canonical commutation relations between the field and the conjugate momentum

$$[\phi(x,t), \pi(y,t)] = i\delta^{N-1}(x-y)$$

together with (3.2) imply the commutation relation between the creation and annihilation operators:

$$[a_k^-, a_{k'}^+] = \delta^{N-1}(k - k')$$

The particle interpretation is based on these operators. The Hamiltonian is diagonal:

$$H = \int d^{N-1}k\ \omega_k a_k^+ a_k^- = \int d^{N-1}k\ \omega_k N_k \tag{3.3}$$

Here, $N_k \equiv a_k^+ a_k^-$ is the number operator for particles with momentum k. We postulate the existence of a vacuum state which satisfies

$$a_k^-|0\rangle = 0, \quad \forall k \tag{3.4}$$

The state which has n_i particles with momentum k_i can then be unambiguously defined as

$$|n_1, n_2, \ldots\rangle = \prod_i \frac{(a_{k_i}^+)^{n_i}}{\sqrt{n_i!}}|0\rangle \tag{3.5}$$

here, i runs over the total number of different k-values which have nonzero occupation numbers. The (infinite dimensional) Hilbert space is spanned by the totality of the possible vectors $|n_1, n_2, \ldots\rangle$, which together make up the so-called Fock basis [19].

3.2 MODE FUNCTIONS IN CURVED SPACETIME

In a general curved spacetime the KG equation, derived from the scalar field Lagrangian which we will introduce in the next chapter, is more complicated:

$$(\Box - m^2)\phi = \frac{1}{\sqrt{|g|}}\partial_\nu\left(\sqrt{|g|}g^{\mu\nu}\partial_\mu\phi\right) - m^2\phi = 0 \quad (3.6)$$

where $|g|$ is the absolute value of the determinant of the metric $g_{\mu\nu}$. We would still like to Fourier expand the field using its mode functions v_k (here, we assume that the mode functions are isotropic and hence only depend on $k = |k^j|$) as

$$\phi(x,t) = \int \frac{d^{N-1}k}{(2\pi)^{(N-1)/2}}\left(a_k^- v_k^* e^{ik_jx^j} + a_k^+ v_k e^{-ik_jx^j}\right) \quad (3.7)$$

In Minkowski space the mode functions are solutions to the harmonic oscillator equation, but for curved spacetimes the equation for Φ_k corresponding to (3.1) is not as simple. For the sake of concreteness, we will now specialize to an important class of solutions of Einstein's equations, characterized by a (spatially) flat Friedmann-Lemaître-Robertson-Walker (FLRW) metric. The metric is then of the form

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2 = a^2(\eta)(-d\eta^2 + d\vec{x}^2) \quad (3.8)$$

For the second equality we defined $\eta = \int \frac{dt}{a}$. The KG equation then takes on the form

$$\left(-\partial_\eta^2 + \Delta + (2-N)\frac{\partial_\eta a}{a}\partial_\eta - m^2a^2\right)\phi = 0 \quad (3.9)$$

Here, Δ is the (N – 1)-dimensional Laplacian. If we define $\varphi = a\Phi$ and transform to Fourier space we nd, denoting ∂_η by a prime,

$$-\varphi_k'' + (4-N)\frac{a'}{a}\varphi_k' - \left(m^2a^2 + k^2 - \left[\frac{a''}{a} + (4-N)\frac{a'^2}{a^2}\right]\right)\varphi_k = 0 \quad (3.10)$$

For concreteness and simplicity we will examine the four-dimensional case, which is of course usually the one of greatest interest. The equation reduces to a time dependent harmonic oscillator equation, with

$$\omega^2(\eta) \equiv m^2a^2(\eta) + k^2 - \frac{a''(\eta)}{a(\eta)} = k^2 + m_{\text{eff}}^2(\eta)$$

If we have two linearly independent real solutions $u_{1,k}(\eta)$ and $u_{2,k}(\eta)$ then the solution $v_k(\eta) \equiv u_{1,k}(\eta) + iu_{2,k}(\eta)$ and its complex conjugate $v_k^*(\eta)$ form a basis of the space of (complex) solutions. We can normalize v_k with the condition

$$2\text{Im}(v'v^*) = \frac{v'v^* - v'^*v}{i} = \frac{W(v,v^*)}{i} = -W(u_1,u_2) = 1 \quad (3.11)$$

The appearance of the Wronskian W shows that $2\text{Im}(v'v^*)$ is indeed a nonzero real constant that we can normalize to 1. With this normalization, v and v^* are called mode functions, because it is exactly the normalization one needs to make the standard commutation relations the eld and conjugate momentum compatible with those for the creation and annihilation operators that appear in the mode expansion of the field, as can be shown using equation (3.7).

3.3 BOGOLYUBOV TRANSFORMATIONS

It is clear from (3.7) that $a_k^\pm$ can be expressed in terms of Φ and v_k. Therefore, different choices for v_k will lead to different creation and annihilation operators. As the complex solution of a second order dierential equation with one normalization condition, the possible vk are determined only up to a complex parameter. Realizing that multiplication by a constant phase factor does not lead to physically distinguishable results tells us that we have a family of physically inequivalent v_k, distinguished by a single real parameter. Each of these gives rise to a different set of $a_k^\pm$, and thus to a different definition of the vacuum and corresponding particle states, which are defined by (3.4) and (3.5). We are thus led to the conclusion that there is no unique choice of the vacuum and the definition of a particle in a curved background. In Minkowski space, the existence of a global timelike Killing vector allows us to pick out a natural set of modes that unambiguously distinguish positive- and negative-frequency solutions and thus particles, but this does not carry over to general spacetimes [21].

Different mode functions can be related to each other using a so-called Bogolyubov transformation. Suppose we have two sets of mode functions: v_k and w_k. Since each of them (with their complex conjugates) form a complete basis of the space of solutions, we can express each one in terms of a linear combination of the other:

$$w_k^* = \alpha_k v_k^* + \beta_k v_k \quad (3.12)$$

It follows from the normalization condition (3.11) that the Bogolyubov coecients α_k and β_k satisfy the condition

$$|\alpha_k|^2 - |\beta_k|^2 = 1$$

The field can expressed in terms of either set of mode functions. Equating these expansions and using (3.12) allows one to relate the creation and annihilation operators $a_k^\pm$ and $b_k^\pm$, respectively) to each other. We find

$$b_k^- = \alpha_k a_k^- + \beta_k^* a_{-k}^+, \qquad b_k^+ = \alpha_k^* a_k^+ + \beta_k a_{-k}^- \quad (3.13)$$

Because each set of annihilation operators defines a different vacuum, we can expect b-particles to appear in the a-vacuum and vice versa. Indeed, a quick calculation shows that this is the case:

$$\begin{aligned}\left\langle 0^a \left| N_k^b \right| 0^a \right\rangle &= \left\langle 0^a \left| b_k^+ b_k^- \right| 0^a \right\rangle = \left\langle 0^a \left| \left(\alpha_k^* a_k^+ + \beta_k a_{-k}^-\right)\left(\alpha_k a_k^- + \beta_k^* a_{-k}^+\right) \right| 0^a \right\rangle \\ &= \left\langle 0^a \left| |\beta_k|^2 a_{-k}^- a_{-k}^+ \right| 0^a \right\rangle = |\beta_k|^2 \delta^3(0)\end{aligned}$$

The divergent delta function is associated with the infinite volume of our space. By dividing by it we can define a mean density of b-particles with momentum k:

$$n_k = |\beta_k|^2 \quad (3.14)$$

This striking result shows explicitly that what looks like the vacuum from one perspective is filled with particles from another.

3.4 THE NOTION OF A PARTICLE IN CURVED SPACETIMES

Although in some situations, it is possible to identify a natural choice of mode functions, there is no overall 'best' prescription for the mode functions in general spacetimes. We can intuitively understand why the particle interpretation breaks down in curved spacetimes by this heuristic argument: In QFT, the particle interpretation is based on decomposing field excitations into plane waves $e^{ik_\mu x^\mu}$. A particle with momentum p is really a wavepacket with a spread of momentum δp. For the particle to be well-defined, we require $\delta p \ll p$. Using the uncertainty principle, we can relate the spread in momentum to the spatial size of the particle: $\delta x \delta p \sim 1 \rightarrow \delta x \sim \frac{1}{\delta p} \gg \frac{1}{p}$.

It is clear that, if the curvature of the spacetime is such that the geometry changes signicantly over a distance δx, the plane wave solution is not a good approximation. The notion of a particle of (physical) momentum p therefore breaks down if the spacetime is not extremely flat on length and time scales of order p^{-1}. This is in agreement with the intuition that spacetime curvature should not affect the shortwavelength (UV) behavior of a theory, because space is almost exactly flat at short distances so that the Minkowski mode functions (and hence the particle interpretation) remain valid. Generally, we must be very careful when interpreting results from QFT in curved spacetime in terms of particles. In light of these considerations, it is often preferred in the literature to work purely in terms of correlation functions of the fields. For a more detailed discussion of the breakdown of the particle interpretation in de Sitter space, see the first section of [10].

CHAPTER 4

FREE SCALAR QFT IN DE SITTER SPACE

We consider a noninteracting, massive, real scalar field minimally coupled to gravity on a de Sitter background. The action is

$$S = -\frac{1}{2}\int d^N x\sqrt{|g|}\left(g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi + m^2\phi^2\right) \tag{4.1}$$

The curved spacetime version of the KG equation, derived from the above action, is

$$\frac{1}{\sqrt{|g|}}\partial_\nu\left(\sqrt{|g|}g^{\mu\nu}\partial_\mu\phi\right) - m^2\phi = 0$$

We will now study the KG equation and its solutions in both the Poincare patches and global de Sitter space.

4.1 THE FREE FIELD IN THE POINCARE PATCHES

In the EPP or CPP (we can treat both simultaneously here), the equation of motion (a special case of (3.9)) has the form

$$\left(-\eta^2\partial_\eta^2 + \eta^2\Delta + \eta(N-2)\partial_\eta - m^2\right)\phi = 0 \tag{4.2}$$

transforming to Fourier space, we find

$$\left(\eta^2\partial_\eta^2 + \eta(2-N)\partial_\eta + \eta^2 k^2 + m^2\right)\phi_k = 0 \tag{4.3}$$

Using the ansatz $\phi_k = \eta^{\frac{N-1}{2}} f(k\eta)$, this can be transformed into Bessel's equation of order $i\mu = \frac{1}{2}\sqrt{(N-1)^2 - 4m^2}$ for f.

We also define the quantities $h_\pm = \frac{1}{2}(N-1)$ i, which will be useful later. The general solution has the following asymptotic behavior [20]:

$$f(k\eta) = \begin{cases} \frac{1}{\sqrt{k\eta}}\left(Ae^{ik\eta} + Be^{-ik\eta}\right), & k\eta \rightarrow \infty \\ C(k\eta)^{i\mu} + D(k\eta)^{-i\mu}, & k\eta \rightarrow 0 \end{cases} \tag{4.4}$$

A, B, C, D are arbitrary complex constants. Recalling that $\eta = e^{\pm\tau}$ we see that we can interpret $k\eta^{\pm i\mu}$ as a plane wave in the future (past) innity of the EPP (CPP) if μ is real. For this to be the case the eld must be 'heavy': $m > \frac{N-1}{2}$. It is then said to belong to the 'principal series'. The mode functions then oscillate and decay to zero as $\eta^{h\pm}$ when $\eta \rightarrow 0$. For `light' fields ($m \leftarrow \frac{N-1}{2}$, which belong to the `complementary series', this decay to zero is homogeneous since μ is imaginary. When we expand the field in terms of the mode functions, using the normalization condition (3.11) to ensure the compatibility of the essential commutation relations, we have:

$$\phi(\vec{x},\eta) = \int \frac{d^{N-1}k}{(2\pi)^{(N-1)/2}}\eta^{\frac{N-1}{2}}\left(a_k^- f(k\eta)e^{ik_j x^j} + a_k^+ f^*(k\eta)e^{-ik_j x^j}\right) \tag{4.5}$$

Now, we have to choose a basis of the two-dimensional space of solutions to the equation of motion (4.3). For instance, one can use Bessel functions of the first or second kinds ($J_{i\mu}$ and $Y_{i\mu}$). This corresponds to choosing mode functions such that either C or D is zero, while A and B are both nonzero. Thus, these Bessel harmonics represent plane waves in the future (past) infinity in the EPP (CPP), and are called out- (in-)harmonics. Instead, one might write the solution in terms of the Hankel functions:

$$H^{(1)}_{i\mu}(x) = J_{i\mu}(x) + iY_{i\mu}(x), \qquad H^{(2)}_{i\mu}(x) = J_{i\mu}(x) - iY_{i\mu}$$

From them, we can construct the following mode functions:

$$f(k\eta) = \frac{\sqrt{\pi}}{2} e^{-\pi\mu} H^{(1)/(2)}_{i\mu}(k\eta) \tag{4.6}$$

These mode functions correspond to setting A or B to zero, with non-vanishing C and D. They are called the Bunch-Davies (BD) modes [31] and behave as plane waves in the past (future) infinity of the EPP (CPP), as ! 1. Remembering that the physical momentum is $k\eta$, we see that the BD modes are the right mode functions to reproduce the correct behavior in the UV limit, where spacetime curvature should not play a role. Alternatively, one can show that it is only possible to diagonalize the Hamiltonian in the limit $\eta \to \infty$ when using the BD modes. It is not possible to diagonalize the Hamiltonian at all as $\eta \to 0$, as could be expected from the fact that this corresponds to the infrared limit of the theory, where gravitational effects are important [20].

4.2 THE FREE FIELD IN GLOBAL DE SITTER SPACE

In the global de Sitter coordinates (3.6) becomes

$$\left(-\partial_t^2 + (N-1)\tanh(t)\partial_t + \frac{\Delta(\Omega)}{\cosh^2(t)} - m^2\right)\phi = 0 \tag{4.7}$$

here, $\Delta(\Omega)$ is the (N - 1)-dimensional spherical Laplacian. This equation can be solved by separation of variables [32]:

$$\phi = y_\ell(t) Y_{\ell\vec{m}}(\Omega)$$

where $Y_{\ell\vec{m}}(\Omega)$ are the spherical harmonics on S^{N-1}, with $m = (m_1, \ldots, m_{N-2})$ as a shorthand denoting their indices. These satisfy

$$\Delta(\Omega) Y_{\ell\vec{m}} = -\ell(\ell + N - 2) Y_{\ell\vec{m}} \tag{4.8}$$

Combining (4.7) and (4.8) yields

$$\left(\partial_t^2 + (1-N)\tanh(t)\partial_t + m^2 + \ell[\ell + N - 2]\right) y_\ell = 0 \tag{4.9}$$

which can be recast as a hypergeometric equation. This equation is solved using the hypergeometric function F($a; b; c; d$) [20, 33]. One possible set of solutions is

$$y^{\text{in}}_\ell(t) = \frac{2^{\ell+(N-2)/2}}{\sqrt{\mu}} \cosh^\ell t\; e^{(\ell+h_\pm)t} F\left(\ell + \frac{N-1}{2}, \ell + h_\mp; 1 \mp i\mu; -e^{2t}\right) \tag{4.10}$$

The superscript `in' added to these solutions is justied by noting their plane wave form as $t \to -\infty$, where $F \to 1$:

$$t \to -\infty \Rightarrow y^{\text{in}}_\ell \sim e^{h_\mp t}$$

We can also define a related solution $y^{\text{out}}_\ell(t) = (y^{\text{in}}_\ell)^*(-t)$, which behaves as a plane wave in the distant future.hm Therefore, we can define the in- (out-) harmonics as

$$\phi^{\substack{\text{in}\\ \text{out}}}_{\ell\vec{m}} = y^{\substack{\text{in}\\ \text{out}}}_\ell Y_{\ell\vec{m}}(\Omega) \tag{4.11}$$

which are the positive frequency modes with respect to t at the past (future) boundary, representing incoming (outgoing) particle states. Note that, $\eta_\pm = e^{\mp t}$ after the identification we see that the asymptotic behavior of these modes corresponds to the in- (out) harmonics in the CPP (EPP). This is perhaps not surprising, since the metric of global de Sitter space is well approximated by that of the Poincare patches in the distant past and future, respectively. As we will see, there is a more general correspondence between Poincare patch modes and global de Sitter modes.

The so-called Euclidean modes are another important set of solutions. These are obtained by requiring that the modes are regular under analytical continuation to the Euclidean sphere. Technically, one defines a Euclidean time coordinate $\vartheta = it + \frac{\pi}{2}$ and requires the mode functions not to 'blow up' at $t = -i\frac{\pi}{2}$, the Euclidean `south pole' [5]. The original motivation for such a procedure came from the Hartle-Hawking 'no-boundary' proposal [34]. The Euclidean modes are given by [20]:

$$y^E_\ell(t) = \frac{2^{\ell+N/2-1} i^{-\ell+\frac{N-1}{2}}}{\sqrt{\mu}} \cosh^\ell t\; e^{(\ell+h_\pm)t} F\left(\ell + \frac{N-1}{2}, \ell + h_\pm; 2\ell + N - 1; 1 + e^{2t}\right) \tag{4.12}$$

The vacuum state defined by the Euclidean mode functions is invariant under the isometry group of de Sitter space.

The free Hamiltonian in global de Sitter space cannot be diagonalized as $t \to \pm\infty$, as this corresponds to the IR limit: The physical momentum $k_{\text{phys}} = k / \cosh t$ goes to zero and gravitational effects cannot be neglected.

4.3 TWO POINT FUNCTIONS

Each choice of mode functions gives rise to a certain vacuum state $|0_\Omega\rangle$. We can study two point functions with respect to this vacuum. An important example is the Wightman function:

$$G^W_\Omega(x_1, x_2) = \langle 0_\Omega | \phi(x_1)\phi(x_2) | 0_\Omega \rangle \tag{4.13}$$

There are many other possible two point functions that one might study, but these can all be obtained from the Wightman function [30]. The Wightman function obeys the KG equation:

$$(\Box - m^2) G^W(x_1, x_2) = 0 \tag{4.14}$$

When dealing with a mode function that defines a de Sitter invariant vacuum state, the Wightman function can only depend on the geodesic distance Z between the two points. In terms of this quantity, the KG equation becomes [35]

$$\left((Z^2-1)\partial_Z^2 + NZ\partial_Z + m^2\right)G^W(Z) = 0 \tag{4.15}$$

Changing variables to $\tilde{Z} \equiv \frac{1+Z}{2}$ this takes on the standard form for the hypergeometric differential equation [36]:

$$\left[\tilde{Z}(\tilde{Z}-1)\partial_{\tilde{Z}}^2 + N\left(\tilde{Z}-\frac{1}{2}\right)\partial_{\tilde{Z}} + m^2\right]G^W(\tilde{Z}) = 0$$

With the general solution (again in terms of the hypergeometric function)

$$G^W(Z) = C_1 F\left(h_+, h_-; \frac{N}{2}; \frac{1+Z}{2}\right) + C_2 F\left(h_+, h_-; \frac{N}{2}; \frac{1-Z}{2}\right) \tag{4.16}$$

The first of the two linearly independent solutions is the standard one given the coefficients in the hypergeometric equation, while the second is obtained by observing that (4.15) is invariant under $Z \to -Z$. They each have two branch points: one where the fourth argument takes on the value 1 (i.e. $Z = \pm 1$) and one at infinity. We therefore make two branch cuts, from $Z = \pm 1$ to infinity. As $Z \to \infty$, the general solution behaves as [20]

$$G^W(Z) \sim K_1 Z^{-h_-} + K_2 Z^{-h_+} \tag{4.17}$$

where $K_{1,2}$ are complex constants that depend on $C_{1,2}$. After these general remarks on the general solution we should study the Wightman function in some more detail in both global de Sitter space and the Poincare patches.

4.3.1 THE WIGHTMAN FUNCTION IN GLOBAL DE SITTER SPACE

The Euclidean Green's function, which is de Sitter invariant, can be obtained by analytic continuation of the (unique) Green's function on the sphere in Euclidean space [30]. The in/out harmonics in global de Sitter coordinates can be related to the Euclidean ones by a so-called Mottola-Allen transformation [32, 37]:

$$\phi^{(\alpha)}_{\ell,\vec{m}}(x) = \frac{1}{\sqrt{1-e^{2\mathrm{Re}(\alpha)}}}\left[\phi^E_{\ell,\vec{m}}(x) + e^{\alpha}\phi^{E*}_{\ell,\vec{m}}(x)\right], \qquad \mathrm{Re}(\alpha) < 0 \tag{4.18}$$

they are part of the one-parameter (α) family of mode functions that give rise to the so-called α-vacua. These α-harmonics are the most interesting solutions to the equation of motion in global de Sitter space because the α-vacuum states are de Sitter invariant (at least in the free theory). This is shown explicitly by Bousso et al. [32].The idea is to use (4.18) to show that the two point function corresponding to an α-vacuum can be expressed entirely in terms of the Euclidean two point function G^W_E. The result is:

$$G^W_\alpha(x_1,x_2) = \frac{1}{1-e^{2\mathrm{Re}(\alpha)}}\left[G^W_E(x_1,x_2) + e^{2\mathrm{Re}(\alpha)}G^W_E(x_2,x_1) + e^{\alpha^*}G^W_E(x_1,x_{2,A}) + e^{\alpha}G^W_E(x_{1,A},x_2)\right] \tag{4.19}$$

Here, we have introduced the notation $x_{i,A}$ to denote the point antipodal to x_i. This shows that the G^W_α, and hence the corresponding vacuum states $|0_\alpha\rangle$, are all $SO(N,1)$ invariant.

In terms of the general solution (4.16), the Euclidean modes correspond to $C_2 = 0$.The other constant is then fixed by examining the behavior of the hypergeometric function near its singularity at $Z = 1$, which corresponds to the short distance limit (recall $Z = \cos D$). Here, one requires that the solution behaves as the propagator in flat space, which also fixes the ϵ-prescription to be the same as in Minkowski space [20, 30]:

$$G^W_E(Z) \longrightarrow G^W_E\left(Z - i\epsilon\,\mathrm{sgn}(\Delta t)\right) \qquad \Delta t \equiv t_2 - t_1 \tag{4.20}$$

The Wightman functions corresponding to any of the α-vacua can be found by applying (4.19). These all have another singularity at $Z = -1$ because $C_2 \neq 0$.

4.3.2 THE WIGHTMAN FUNCTION IN THE EPP/CPP

In the Poincare patches, the general form of the Wightman function is easily obtained from the mode expansion (4.5), taking into account that $a^-_{k,(\Omega)}|0_\Omega\rangle = \langle 0_\Omega|a^+_{k,(\Omega)} = 0$ and using the commutation relations for the creation and annihilation operators:

$$G^W_\Omega = \int \frac{d^{N-1}k}{(2\pi)^{N-1}}(\eta_1\eta_2)^{\frac{N-1}{2}} f^{(\Omega)}(k\eta_1) f^{(\Omega)*}(k\eta_2) e^{ik_j(x_1^j - x_2^j)} \tag{4.21}$$

The particular integral that one must evaluate for each choice of mode functions (e.g. in-, out-, or BD harmonics) can be looked up in a table of standard integrals such as [38]. Each different choice of $f^{(\Omega)}(k\eta)$ turns out to correspond to different values of C_1 and C_2 in (4.16), establishing a one-to-one correspondence with the global de Sitter α-harmonics.

In particular, the BD harmonics yield the solution where $C_2 = 0$ and therefore correspond to the Euclidean modes in global de Sitter space. Hence, they can be expected to produce the correct UV behavior near the singularity at $Z = 1$ with the ϵ-prescription

$$G^W_{\mathrm{BD}}(Z) \longrightarrow G^W_{\mathrm{BD}}\left(Z \mp i\epsilon\,\mathrm{sgn}(\Delta\eta_\pm)\right) \tag{4.22}$$

Note the appearance of the minus in the EPP, which is due to the reversed time flow. Moreover, the BD mode functions can be used in equations exactly analogous to (4.18) and (4.19) to obtain the Green's functions for other harmonics in the EPP/CPP [20]. Just like in global de Sitter coordinates, the other mode functions have a second singularity as well as a dierent residue at $Z = 1$.

CHAPTER 5

KELDYSH-SCHWINGER FORMALISM

The Hamiltonian in both the Poincare patches and global de Sitter coordinates depends on time. The 'standard' QFT technique of evaluating correlation functions is therefore no longer appropriate, as it assumes the vacuum at future infinity is the same as the initial vacuum state. Now, there is no reason to believe that the final state is the same as the initial state. We would therefore like to avoid any mention of the final state altogether. The Keldysh-Schwinger (also referred to as the in-in, or closed time path) formalism [39, 40] was developed in order to deal with this type of non-equilibrium problems. The formalism has a wide range of applications in condensed matter theory, as well as cosmology [41{43]. Here, we will present a short introduction to the formalism. For a more in-depth review, see e.g. [42–46].

5.1 CLOSING THE TIME CONTOUR

To begin, let us review a few more concepts from QFT in flat space. Consider the simple case of a real, scalar field. The generating functional $Z[J]$ is given by

$$Z[J] = \int \mathcal{D}\phi \exp\left[\frac{i}{2}\int d^N x\left(-g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2 + 2J\phi\right)\right] \tag{5.1}$$

$$= \int \mathcal{D}\phi \exp\left[i\int d^N x\left(\frac{1}{2}\phi\left[g^{\mu\nu}\partial_\mu\partial_\nu - m^2\right]\phi + J\phi\right)\right] \tag{5.2}$$

(5.3)

Using the Fourier transform of ϕ and J and performing some standard manipulations, we find

$$Z[J] = \mathscr{N}\exp\left[\frac{i}{2}\int d^N x\, d^N y\, \left(J(x)G_F(x-y)J(y)\right)\right] \tag{5.4}$$

where $\mathscr{N}$ is just a normalization constant, independent of J. We normalize by requiring $Z[0] = \langle 0j0i\rangle_{J=0} = 1$, so that $\mathscr{N} = 1$. $G_F(x - y)$ is the Feynman propagator, which satisfies

$$(\Box - m^2)G_F(x-y) = \delta^N(x-y)$$

This is just one example of the identity

$$Z[J] = \int \mathcal{D}\phi \exp\left[i\int d^N x\left(\frac{1}{2}\phi A\phi + J\phi\right)\right] = \mathscr{N}\exp\left[\frac{i}{2}\int d^N x\, d^N y\, \left(J(x)A^{-1}J(y)\right)\right] \tag{5.5}$$

which can be used to quickly extract the Feynman propagator. When there are interactions it receives corrections, which are usually calculated using Feynman diagrams and other standard techniques from perturbation theory.

The propagator is an important tool in evaluating n-point functions. Alternatively, these can be calculated using functional derivatives:

$$\langle 0\left|T\left\{\phi(x_1)\dots\phi(x_n)\right\}\right|0\rangle = \left(\frac{1}{i}\right)^n \left.\frac{\delta^n Z[J]}{\delta J(x_1)\dots\delta J(x_n)}\right|_{J=0} \tag{5.6}$$

From this formula one can quickly derive that the Feynman propagator is equal to the time-ordered 2-point function:

$$\frac{1}{i}G_F(x-y) = \langle 0\left|T\left\{\phi(x_1)\phi(x_2)\right\}\right|0\rangle \tag{5.7}$$

A general expectation value takes the form (easily generalizable to an arbitrary number of operators):

$$E(t_1,t_2) = \left\langle \Omega \left| T\left\{\bar{T}e^{i\int_{t_0}^{t} H dt'} O_2 O_1 T e^{-i\int_{t_0}^{t} H dt'}\right\}\right| \Omega\right\rangle \tag{5.8}$$

where t_0 is an initial time, Ω the state at that time and is a $(\bar{T})\ Te^{\int x(t)dt}$ |(anti-) timeordered exponential. In the interaction picture, we have

$$\begin{aligned}E(t_1,t_2) &= \langle 0\left|T\left\{S^+(t_0,-\infty)S^+(t_2,t_0)O_2(t_2)S(t_2,t_1)O_1(t_1)S(t_1,t_0)S(t_0,-\infty)\right\}\right|0\rangle\\ &= \langle 0\left|T\left\{S(-\infty,t_0)S(t_0,t_2)O_2(t_2)S(t_2,t_1)O_1(t_1)S(t_1,t_0)S(t_0,-\infty)\right\}\right|0\rangle\end{aligned}$$

We introduced $S(t,t_0) = S^+(t_0,t) = Te^{-i\int_{t_0}^{t} H^{\text{int}}dt'}$, which is responsible for adiabatically switching 'on' the interactions (H^{int} being the interaction Hamiltonian). Furthermore used that $|\,\Omega\,| = S(t_0, -\infty)|0\rangle$, with $|0\rangle$ the ground state of the free Hamiltonian H_0, which is the initial state (at $t = \infty$). In regular QFT, we now assume that the final state (evolved to $t = \infty$) is unique, independent of the procedure of switching 'on' and 'o' interactions and *equal to the initial state up to a phase factor*: $\langle 0|\ S(\infty, -\infty)0\rangle = e^{i\theta}$. This implies that $\langle 0|S(-\infty,\ t_0) = e^{-i\theta}\langle 0|\ S(\infty,\ -\infty)S(-\infty,\ t_0)$ [44]. Rearranging the time ordered terms, which is always allowed, we then arrive at

$$E(t_1,t_2) = \frac{\langle 0\left|T\left\{O_2(t_2)O_1(t_1)S(\infty,-\infty)\right\}\right|0\rangle}{\langle 0\left|S(\infty,-\infty)\right|0\rangle} \tag{5.9}$$

In the non-equilibrium case, however, when the interactions are switched 'on' and eventually 'off' again, the state has evolved unpredictably to some new, unknown state. Therefore, equation (5.9) does not hold; the need for a theory that does not refer to the final state arises. Schwinger had the crucial idea to evolve the system from $t = -\infty$ to $t = \infty$ *and back again*, so that one returns to the initial state no matter what happens at t = 1: the time contour is closed. The switching 'on' and 'off' both take place in the past (on the forward and backward branch, respectively). However, this construction forces one to consider fields living on each of the two branches of the closed time contour. We will label fields living on the forward time branch with a `plus' subscript, while those on the backward time branch receive a 'minus' subscript. We then have to work with the following generating functional

$$\begin{aligned}Z[J_+,J_-] &= \int \mathcal{D}\phi_+\mathcal{D}\phi_- \exp\left[i\int_{t_i}^{t_f} dt\int d^{N-1}x\sqrt{|g|}\left(\mathcal{L}[\phi_+] + J_+\phi_+\right)\right.\\ &\qquad\left. + i\int_{t_f}^{t_i} dt\int d^{N-1}x\sqrt{|g|}\left(\mathcal{L}[\phi_-] + J_-\phi_-\right)\right]\\ &= \int \mathcal{D}\phi_+\mathcal{D}\phi_- \exp\left[i\int_{t_i}^{t_f} dt\int d^{N-1}x\sqrt{|g|}\left(\mathcal{L}[\phi_+] - \mathcal{L}[\phi_-] + J_+\phi_+ - J_-\phi_-\right)\right]\end{aligned} \tag{5.10}$$

In our case of free field Lagrangians, we can write this as

$$Z[J_+, J_-] =$$
$$\int \mathcal{D}\phi_+ \mathcal{D}\phi_- \exp\left[i \int d^N x \sqrt{|g|} \left(\frac{1}{2} \begin{pmatrix} \phi_+ & \phi_- \end{pmatrix} \begin{pmatrix} (\Box - m^2) & 0 \\ 0 & -(\Box - m^2) \end{pmatrix} \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} + \begin{pmatrix} \phi_+ & \phi_- \end{pmatrix} \begin{pmatrix} J_+ \\ -J_- \end{pmatrix} \right)\right] \tag{5.11}$$

Introducing the density matrix ρ in order to generalize to mixed states, an expectation value of an operator (5.8) is expressed in terms of the generating functional as

$$\begin{aligned} E(t_1, t_2) &= \mathrm{Tr}(\rho O_1 O_2) \\ &= \int \mathcal{D}\varphi^+ \langle \varphi^+ | \rho O_1 O_2 | \varphi^+ \rangle \\ &= \int \mathcal{D}\varphi^+ \mathcal{D}\varphi^- \langle \varphi^+ | \rho(t_{\mathrm{in}}) | \varphi^- \rangle \langle \varphi^- | O_1 O_2 | \varphi^+ \rangle \\ &= \int \mathcal{D}\varphi^+ \mathcal{D}\varphi^- \langle \varphi^+ | \rho(t_{\mathrm{in}}) | \varphi^- \rangle \int_{\varphi^+}^{\varphi^-} \mathcal{D}\phi^+ \mathcal{D}\phi^- O_2(t_1) O_1(t_2)\; e^{i(S[\phi^+] - S[\phi^-])} \end{aligned}$$

where the generating functional is now

$$Z[J_-, J_+, \rho(t_{\mathrm{in}})] = \int \mathcal{D}\varphi^+ \mathcal{D}\varphi^- \langle \varphi^+ | \rho(t_{\mathrm{in}}) | \varphi^- \rangle \int_{\varphi^+}^{\varphi^-} \mathcal{D}\phi^+ \mathcal{D}\phi^- \; e^{i(S[\phi^+] - S[\phi^-])} \tag{5.12}$$

We can appeal to the obvious generalization of (5.5) to immediately find:

$$\begin{aligned} Z[J_+, J_-, \rho(t_{\mathrm{in}})] &= \mathcal{N} \int \mathcal{D}\varphi^+ \mathcal{D}\varphi^- \langle \varphi^+ | \rho(t_{\mathrm{in}}) | \varphi^- \rangle \\ &\times \exp\left[\frac{i}{2} \int d^N x \, d^N y \, \begin{pmatrix} J_+(x) & -J_-(x) \end{pmatrix} \begin{pmatrix} G^{++}(x,y) & G^{+-}(x,y) \\ G^{-+}(x,y) & G^{--}(x,y) \end{pmatrix} \begin{pmatrix} J_+(y) \\ -J_-(y) \end{pmatrix} \right] \end{aligned} \tag{5.13}$$

Here we can again require $Z[0, 0, \rho] = 1$ to fix $\mathcal{N}$. The matrix of propagators satisfies

$$\begin{pmatrix} (\Box - m^2) & 0 \\ 0 & -(\Box - m^2) \end{pmatrix} \begin{pmatrix} G^{++}(x,y) & G^{+-}(x,y) \\ G^{-+}(x,y) & G^{--}(x,y) \end{pmatrix} = \delta(x-y) \mathbb{1} \tag{5.14}$$

Through the generalization of (5.6),

$$\begin{aligned} &\langle 0 | \bar{T}\{\phi_-(x_1) \ldots \phi_-(x_m)\} T\{\phi_+(x_{m+1}) \ldots \phi_+(x_{m+n})\} | 0 \rangle \\ &\qquad = i^{m-n} \frac{\delta^{m+n} Z[J_+, J_-]}{\delta J_-(x_1) \ldots \delta J_-(x_m) \delta J_+(x_{m+1}) \ldots \delta J_+(x_{m+n})} \bigg|_{J_+ = J_- = 0} \end{aligned} \tag{5.15}$$

we can relate each propagator to a 2-point function:

$$\begin{aligned} \frac{1}{i} G^{++}(x,y) &= \langle 0 | T\{\phi_+(x)\phi_+(y)\} | 0 \rangle \\ &= \theta(t_x - t_y) \left\langle 0 \middle| \hat{\phi}(x)\hat{\phi}(y) \middle| 0 \right\rangle + \theta(t_y - t_x) \left\langle 0 \middle| \hat{\phi}(y)\hat{\phi}(x) \middle| 0 \right\rangle \\ \frac{1}{i} G^{+-}(x,y) &= \langle 0 | \phi_+(y)\phi_-(x) | 0 \rangle \\ &= \left\langle 0 \middle| \hat{\phi}(y)\hat{\phi}(x) \middle| 0 \right\rangle \\ \frac{1}{i} G^{-+}(x,y) &= \langle 0 | \phi_-(x)\phi_+(y) | 0 \rangle \\ &= \left\langle 0 \middle| \hat{\phi}(x)\hat{\phi}(y) \middle| 0 \right\rangle \\ \frac{1}{i} G^{--}(x,y) &= \langle 0 | \bar{T}\{\phi_-(x)\phi_-(y)\} | 0 \rangle \\ &= \theta(t_y - t_x) \left\langle 0 \middle| \hat{\phi}(x)\hat{\phi}(y) \middle| 0 \right\rangle + \theta(t_x - t_y) \left\langle 0 \middle| \hat{\phi}(y)\hat{\phi}(x) \middle| 0 \right\rangle \end{aligned} \tag{5.16}$$

Note that, when considering the Poincare patches, the argument of the step functions should be suitably modied: $t \to \eta$, while in the EPP the sign of the argument also needs to be changed due to the reversed flow of time. The propagators are not all independent. They satisfy the following identity:

$$G^{++}(x,y) + G^{--}(x,y) = G^{+-}(x,y) + G^{-+}(x,y)$$

which is clear from the equations (5.16) where the subscripts + and – have been dropped after the last equality signs. This identity can be exploited by changing to what is knowns as the Keldysh basis.

5.2 THE KELDYSH ROTATION

In order to minimize the inconvenience we experience due to the doubling of the degrees of freedom of our theory we can use a simple change of variables, known as the Keldysh rotation:

$$\begin{pmatrix} \phi_{\mathrm{cl}} \\ \phi_{\mathrm{q}} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\phi_+ + \phi_-) \\ \phi_+ - \phi_- \end{pmatrix} = M \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}, \qquad M \equiv \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & -1 \end{pmatrix} \tag{5.17}$$

The labels 'cl' and 'q', which stand for 'classical' and 'quantum', come from the condensed matter literature. The free field action can be rewritten as

$$S = -\int d^N x \sqrt{|g|} \left(\partial_\mu \phi_{\mathrm{cl}} \partial^\mu \phi_{\mathrm{q}} + m^2 \phi_{\mathrm{cl}} \phi_{\mathrm{q}} \right) \tag{5.18}$$

and the matrix of propagators becomes

$$M \begin{pmatrix} G^{++}(x,y) & G^{+-}(x,y) \\ G^{-+}(x,y) & G^{--}(x,y) \end{pmatrix} M^T = \begin{pmatrix} iG_K(x,y) & G_R(x,y) \\ G_A(x,y) & 0 \end{pmatrix} \tag{5.19}$$

The vanishing of the bottom right entry is a direct consequence of the relation between the four propagators, which the Keldysh rotation is designed to take advantage of.
The new propagators are related to the original ones by

$$G_R(x,y) = \left(G^{++}(x,y) - G^{+-}(x,y)\right) = i\theta(t_x - t_y)\left\langle 0 \left| \left[\hat{\phi}(x), \hat{\phi}(y)\right] \right| 0 \right\rangle$$
$$G_A(x,y) = \left(G^{++}(x,y) - G^{-+}(x,y)\right) = i\theta(t_y - t_x)\left\langle 0 \left| \left[\hat{\phi}(y), \hat{\phi}(x)\right] \right| 0 \right\rangle$$
$$G_K(x,y) = \frac{1}{2i}\left(G^{+-}(x,y) + G^{-+}(x,y)\right) = \frac{1}{2}\left\langle 0 \left| \left\{\hat{\phi}(x), \hat{\phi}(y)\right\} \right| 0 \right\rangle \tag{5.20}$$

where {·,·} and [·,·] are the (anti-)commutator. The three propagators are called the Retarded, Advanced and Keldysh propagators. Note that the Keldysh rotation is nothing more than a smart change of basis; calculations may just as well be performed with the original propagators and the choice is a matter of convenience.

5.3 PERTURBATION THEORY

We can now add interactions to the free theory and use the Keldysh-Schwinger formalism to study the effects these induce (e.g. corrections to the propagators) perturbatively. One can write down the corresponding Feynman rules and use a diagrammatic technique to perform calculations in a fashion that is analogous to the standard method from QFT in at space. In most papers on scalar QFT in de Sitter space cubic or quartic self-interactions are used to create 'toy model' field theories to maximize the clarity of exposition while capturing the essential physics. In chapter 6 we will perform a few one-loop calculations using perturbation theory.

CHAPTER 6

IS DE SITTER SPACE STABLE?

During the early 1950's, Schwinger investigated the vacuum polarization due to an external electric field [47]. Using semiclassical methods, he found that the presence of a constant background field leads to the spontaneous production of particles from the vacuum. A brief introduction to semiclassical field theory can be found in Appendix A. It was later realized by others that a time-dependent gravitational background should have the same effect [4]. This provides the basic motivation behind Polyakov's proposed solution to the cosmological constant problem [5{8].

Using the semiclassical field equations, it can be shown that Schwinger's spontaneous particle production due to an 'electric cosmological constant' will cause this external field to decay to zero. One can develop the treatment of this problem in analogy with the gravitational cosmological constant problem. This analogy has been worked out to great depth by a number of authors [7, 9, 10], who interpret the final result as indirect evidence that a similar decay must occur in the gravitational case. As we saw earlier, it is not easy to consistently define the notion of a particle in de Sitter space. Nevertheless, Polyakov and others (see [11] and references therein) claim to have shown that an explosive production of massive particles takes place in de Sitter space. In this chapter, we will describe the results that led to this conclusion. If explosive particle production does indeed occur, it could induce a large backreaction, i.e. an *instability of de Sitter space*, and a breakdown of semiclassical techniques. The hope is that this may lead to a rapid screening of the cosmological constant, which initially takes on a 'naturally high' value, down to the low 'effective' value that we observe today. Due to formidable technical difficulties, nobody has yet succeeded to work these subsequent steps out in a quantitative fashion.

In what will follow we have chosen to consider only results that apply to the expanding Poincare patch, although interesting calculations can and indeed have been carried out for the CPP and global de Sitter space as well (e.g. [8, 13]). It is clear that our universe can be modeled quite well by the expanding section of de Sitter space, but the same cannot be said about the CPP or even global de Sitter space. Therefore, the physical signicance of results pertaining only to the CPP and global de Sitter space is currently unclear, so results which apply to the EPP appear to be of greater relevance.

A point that is commonly raised by those who argue that there are no interesting IR effects in the EPP is that, although some particle production may occur, this can never happen at a rate fast enough to overcome the dilution due to the exponential expansion of space. However, Akhmedov suggests that nonlinear 'explosive' particle production processes may become relevant once the comoving particle density, which is unaffected by the spatial expansion, becomes large enough [20]. This may 'outrace' the expansion, giving rise to a large backreaction. In section 6.4 we will see the justication for this argument.

6.1 SELF-INTERACTION AND FEYNMAN RULES

In this chapter, we will study the simplest interacting theory, obtained by adding a cubic self-interaction $\mathcal{L}_{\text{int}} = -\frac{\lambda}{3!}\phi^3$ to the free theory, following the approach of [7, 20]. We will disregard the well-known instability of this potential, because this will not affect our results. Similar calculations as the ones presented here have been carried out using Φ^4 theory, leading to similar conclusions [14]. The total action is

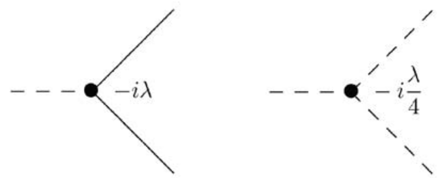
$$S = -\int d^N x\sqrt{|g|}\left[\partial_\mu\phi_+\partial^\mu\phi_+ - \partial_\mu\phi_-\partial^\mu\phi_- + m^2(\phi_+^2 - \phi_-^2) + \frac{\lambda}{3!}(\phi_+^3 - \phi_-^3)\right] \quad (6.1)$$

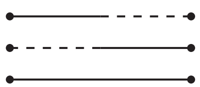
or, after the Keldysh rotation,

$$S = -\int d^N x\sqrt{|g|}\left[\partial_\mu\phi_{\rm cl}\partial^\mu\phi_{\rm q} + m^2\phi_{\rm cl}\phi_{\rm q} + \frac{\lambda}{3!}\left(3\phi_{\rm cl}^2\phi_{\rm q} + \frac{1}{4}\phi_{\rm q}^3\right)\right] \quad (6.2)$$

The Feynman rules in position space (after the Keldysh rotation) are derived for a general class of self-interactions in Appendix A of [48]. The Feynman rules for our field theory can be summarized as follows:

- Each field is assigned a type of line:
  $\Phi_{\rm cl}$:
  $\Phi_{\rm q}$:
  
- There are two distinct vertices, each with an associated factor:



- There are three ways to connect points. Each of them represents a propagator:
  $-iG_R(x, y)$
  $-iG_A(x, y)$
  $G_K(x, y)$
  
- Vertices must be integrated over the entire spacetime.

6.2 DE SITTER INVARIANCE OF THE EXACT BD VACUUM

We will begin by investigating the de Sitter invariance of the propagators (and hence of the vacuum states) at loop level. For this, we will use the original basis where the fields and propagators carry + and – labels. It is clear that each of the α-vacua is de Sitter-invariant at tree level. However, we will show that only the BD vacuum is de Sitter-invariant when loop corrections are taken into account. This argument is due to Polyakov [8].

Forgetting for a moment about the specific theory outlined in the previous section, consider a vertex in a general Keldysh-Schwinger diagram, located at the point Y in terms of the coordinates of the ambient Minkowski space. It is multiplied by a number of propagators, $\Pi_S G(Z(X_S, Y))$, which 'connect' it to the rest of the diagram. We must now integrate the position of the vertex over the entire Poincare patch. Let us define $Y^\pm = Y^0 \pm Y^D$. The measure of integration is then $[dY] = d^{N+1}Y\,\delta(Y^\mu Y_\mu - 1)\theta(Y^-)$, where the delta function confines us to de Sitter space while the Heaviside step function further restricts to the EPP. Generally, this step function seems to break the de Sitter symmetry. Let us investigate what happens when we perform an $SO(N, 1)$ transformation which affects the step function's argument. For instance, consider a small rotation around Y^0 in the Y^1Y^D-plane, so that $\delta Y^D = -\varphi Y^1$. Using the fact that the delta function is the distributional derivative of the step function, we find the variation of the integral

$$\begin{aligned}\delta_\varphi \int [dY] \prod_S G\big(Z(X_S, Y)\big)\ldots \quad (6.3)\\ = \int d^{N+1}Y\,\delta(Y^\mu Y_\mu - 1)\delta(Y^-)\varphi Y^1 \prod_S G\big(Z(X_S, Y)\big)\ldots\\ = \int dY^+ d^{N-1}Y\,\delta(Y^\mu Y_\mu - 1)\varphi Y^1 \prod_S G\big(Z(X_S, Y)\big)\big|_{Y^-=0}\cdots\end{aligned}$$

In the last step, we performed the integration over Y^-. The integrand depends only on the geodesic distance $Z(X_S, Y)$. Using $Y^\pm$, Z can be explicitly written out:

$$Z(X_S, Y) = -\frac{1}{2}(X_S^+ Y^- + X_S^- Y^+) + X^j Y^j \qquad j = 1, 2, \ldots, N-1 \quad (6.4)$$

In our case $Y^- = 0$ because of the appearance of $\delta(Y^-)$ in the integral. Furthermore, $X^-_S\,S \geq 0$ since we are in the EPP. Let us now try to perform the integration over Y^+. We will need the pole prescription for the propagators. Combining (4.22) and (5.16) yields

$$\begin{aligned}G^{++}(Z) &\longrightarrow G^{++}(Z + i\epsilon)\\ G^{+-}(Z) &\longrightarrow G^{+-}(Z - i\epsilon\,\mathrm{sgn}(\Delta\eta))\\ G^{-+}(Z) &\longrightarrow G^{-+}(Z + i\epsilon\,\mathrm{sgn}(\Delta\eta))\\ G^{--}(Z) &\longrightarrow G^{--}(Z - i\epsilon)\end{aligned} \quad (6.5)$$

In fact, the condition $Y^- = 0$ implies $\eta_Y = 0$, i.e. there is a definite pole prescription for all propagators. When we view the propagators as functions of Y^+ + (i.e. $G[Y^+]$ $G[Z[Y^+]]$) they have the same properties of analyticity that they have in the complex Z– plane. Now, let us consider the BD state. Then, the integrand of Y^+ is an analytic function with a single branch cut, shifted up or down slightly depending on whether the vertex is located on the forward or backward time branch.

We see from (6.4) that $|Z| \to \infty$ as $Y^+ \to \infty$. The Wightman function associated with the BD vacuum then displays power-like decay, as can be seen from (4.17). Therefore, we can close the contour using a semicircle at infinity around either the upper or lower half of the complex plane. Inside this contour, the function is analytic so the integral is zero by Cauchy's theorem: The variation due to the rotation vanishes! This argument can easily be extended to rotations in other directions as well as boosts in the ambient Minkowski space. Since we considered a general vertex we conclude that the BD vacuum is de Sitter-invariant even in the presence of an interacting scalar field. The above argument fails for any other propagators than the ones associated with the BD vacuum because the presence of another branch from $Z = -1$ to infinity in the complex Z–plane and the different ϵ – prescription associated with this spoil the properties of analyticity that the argument relies on. Therefore, the BD vacuum is the only de Sitter-invariant state.

Before moving on, some remarks are in order. Firstly, we have ignored the UV divergences associated with loop diagrams. As we argued earlier, the BD vacuum state reproduces the correct behavior in this limit, where everything should behave as in Minkowski space. In all of the calculations of this chapter, we will assume some UV renormalization procedure that adequately deals with these divergences (keeping in mind that this is not a trivial requirement!). It is also important to note that the above conclusions only hold when there are no IR divergences in the loops. We will show in the next section that there are large IR contributions but no divergences in the EPP.

6.3 CALCULATION OF ONE-LOOP INFRARED CONTRIBUTIONS

In this section we will describe a one-loop calculation of the leading corrections to the Keldysh propagator in the IR limit, closely following the description in [20]. This propagator receives large IR contributions. We will also briefly comment on the IR corrections to the advanced and retarded propagator as well as the vertex functions, all of which do not receive big IR contributions.

The first thing to do is set up notation. It is conventional to perform a Fourier transform along all spatial directions:

$$G_X(p;\eta_1,\eta_2) = \int d^{N-1}\vec{x}\ G_X(\vec{x},\eta_1;0,\eta_2)e^{-i\vec{p}\cdot\vec{x}}$$

where $X \in \{K, R, A\}$. Integrating the vertices now takes on a slightly different meaning. We must integrate over the times corresponding to them, while we should also integrate over internal momenta: $\int d^{N-1}\vec{p}/(2\pi)^{N-1}$. Recalling (4.21) and (5.20), the Keldysh propagator is now expressed as

$$G^0_K(p;\eta_1,\eta_2) = (\eta_1\eta_2)^{\frac{N-1}{2}}\mathrm{Re}\big[f(p\eta_1)f^*(p\eta_2)\big] \tag{6.6}$$

while the advanced and retarded propagators are

$$G^0_R(p;\eta_1,\eta_2) = G^0_A(p;\eta_2,\eta_1) = 2\theta(\eta_2-\eta_1)(\eta_1\eta_2)^{\frac{N-1}{2}}\mathrm{Im}\big[f(p\eta_1)f^*(p\eta_2)\big] \tag{6.7}$$

We introduced the superscript 0 to denote a tree level propagator. Similarly, we will use a superscript 1 to denote one-loop quantities.

6.3.1 THE CORRECTION TO THE KELDYSH PROPAGATOR

The next thing to do is identify the one-loop diagrams that contribute to the Keldysh propagator. They are depicted in fig. 6.1.

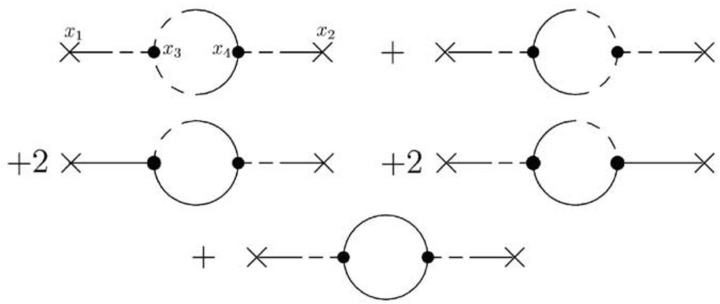


Figure 6.1: *The diagrams that contribute to the one-loop correction to the Keldysh propagator, including symmetry factors. We have not included one-loop tadpole diagrams as momentum conservation at the vertices dictates that these do not contain any interesting momentum-dependent effects.*

Turning on the interactions at some time $\eta_0 \to \infty$ far in the past and applying the Feynman rules we can write out these diagrams:

$$\begin{aligned}
G^1_K(p;\eta_1,\eta_2) = (-i\lambda)^2 \iint \frac{d^{N-1}\vec{q_1}}{(2\pi)^{N-1}}\frac{d^{N-1}\vec{q_2}}{(2\pi)^{N-1}} \iint_{\eta_0}^{0} \frac{d\eta_3\ d\eta_3}{(\eta_3\eta_4)^N}\delta(\vec{p}-\vec{q_1}-\vec{q_2}) \\
\times \Big[\frac{1}{4}G^0_R(p;\eta_1,\eta_3)G^0_A(q_1;\eta_3,\eta_4)G^0_A(q_2;\eta_3,\eta_4)G^0_A(p;\eta_4,\eta_2) \\
+\frac{1}{4}G^0_R(p;\eta_1,\eta_3)G^0_R(q_1;\eta_3,\eta_4)G^0_R(q_2;\eta_3,\eta_4)G^0_A(p;\eta_4,\eta_2) \\
-2G^0_K(p;\eta_1,\eta_3)G^0_K(q_1;\eta_3,\eta_4)G^0_A(q_2;\eta_3,\eta_4)G^0_A(p;\eta_4,\eta_2) \\
-2G^0_R(p;\eta_1,\eta_3)G^0_R(q_1;\eta_3,\eta_4)G^0_K(q_2;\eta_3,\eta_4)G^0_K(p;\eta_4,\eta_2) \\
-\ G^0_R(p;\eta_1,\eta_3)G^0_K(q_1;\eta_3,\eta_4)G^0_K(q_2;\eta_3,\eta_4)G^0_A(p;\eta_4,\eta_2)\Big]
\end{aligned} \tag{6.8}$$

The relative signs arise because of the appearance of two or four Retarded/Advanced propagators, which come with factors of $-i$. In evaluating this integral, we will take an IR limit such that

$$p\eta \to 0, \text{ where } \eta \equiv \sqrt{\eta_1\eta_2}, \text{ while } \frac{\eta_1}{\eta_2} = \text{constant} \tag{6.9}$$

When we remember the definition of η in the EPP, $\eta = e^{-\tau}$, we see that this limit corresponds to taking $\eta_{1,2}$ to future infinity while keeping the time between them constant. Therefore, a large correction in this limit signals a contribution that grows with time.

We begin by substituting in the expressions for the propagators given above. We can take out a common part that appears in the expression for each diagram, followed by a clutter of various $f(p_i\eta_j)$, multiplied by constant pre-factors and step functions that depend on the diagrams:

$$\begin{aligned}
G^1_K(p;\eta_1,\eta_2) = -\frac{\lambda^2}{(2\pi)^{2(N-1)}}(\eta_1\eta_2)^{\frac{N-1}{2}} \\
\times \iint d^{N-1}\vec{q_1}d^{N-1}\vec{q_2} \iint_{\eta_0}^{0} \frac{d\eta_3\ d\eta_3}{(\eta_3\eta_4)^N}\delta(\vec{p}-\vec{q_1}-\vec{q_2})(\eta_3\eta_4)^{\frac{N-3}{2}}\Big[\dots\Big]
\end{aligned} \tag{6.10}$$

The step functions inside [...] will change the limits of integration over $\eta_{3,4}$ for each term. One thing to note is that $\eta_{1,2}$ only appear as second argument of the step functions, so that they become upper limits of the integration over $\eta_{3,4}$. In all but the last diagram, there are also step functions restricting $\eta_3 \to \eta_4$ or vice versa. Since we take the limit where η_1 and η_2 go to the infinite future (with a constant distance between the two), we can simplify by substituting the `average conformal time' $\eta = e^{-(\tau_1+\tau_2)/2}$ instead of $\eta_{1,2}$ in the upper limits of integration. This means that we are implicitly discarding some sub-leading terms. Akhmedov [20] claims that this expression simplifies, after a considerable amount of algebra, to the following expression for the one-loop correction to the Keldysh propagator (the present author was unable to reproduce this step in the derivation):

$$G^1_K(p;\eta_1,\eta_2) = (\eta_1\eta_2)^{\frac{N-1}{2}}\left[f(p\eta_1)f^*(p\eta_2)n^1_p(\eta) + f(p\eta_1)f(p\eta_2)\kappa^1_p(\eta) + \text{complex conj.}\right]$$

$$n^1_p(\eta) \equiv -\frac{\lambda^2}{(2\pi)^{2(N-1)}}\int d^{N-1}\vec{q}_1 d^{N-1}\vec{q}_2 \iint_{\eta_0}^{\eta} d\eta_3 d\eta_4\, (\eta_3\eta_4)^{\frac{N-3}{2}}\delta(\vec{p}-\vec{q}_1-\vec{q}_2)$$
$$\times\ f^*(p\eta_3)f(p\eta_4)f^*(q_1\eta_3)f(q_1\eta_4)f^*(q_2\eta_3)f(q_2\eta_4)$$
$$\kappa^1_p(\eta) \equiv \frac{\lambda^2}{(2\pi)^{2(N-1)}}\int d^{N-1}\vec{q}_1 d^{N-1}\vec{q}_2 \int_{\eta_0}^{\eta} d\eta_3 \int_{\eta_0}^{\eta_3} d\eta_4\, (\eta_3\eta_4)^{\frac{N-3}{2}}\delta(\vec{p}-\vec{q}_1-\vec{q}_2)$$
$$\times\ f^*(p\eta_3)f^*(p\eta_4)f^*(q_1\eta_3)f(q_1\eta_4)f^*(q_2\eta_3)f(q_2\eta_4) \qquad (6.11)$$

This expression can be explicitly evaluated by choosing a set of mode functions. To understand what it represents physically, let us see what would happen if we were in flat space: We substitute $\eta \to t$, $\sqrt{|9|} \to 1$ and replace the mode functions $\eta_i^{\frac{N-1}{2}} f(p\eta_i)$ by the plane wave solutions $e^{-i\omega(p)t}/\sqrt{2\omega(p)}$. We also take the limits of integration over $\eta_{3,4}$ to $t = \pm\infty$. Then, these integrals simply become delta functions expressing energy conservation: $\delta(\omega(p) + \omega(q_1) + \omega(q_2))$. The argument of these is never zero, so performing the integrals over $\vec{q}_1$ and $\vec{q}_2$ yields zero: Both n^1_p and κ^1_p vanish.

Now, evaluating the one-loop IR contribution to G_K is reduced to estimating n^1_p and κ^1_p, which seem to capture the essential physics that differs between the EPP and Minkowski space. The fact that n^1_p and κ^1_p vanish due to energy conservation in flat space can be interpreted as a clue pointing towards particle production when they are nonzero. Furthermore, this tells us that contributions to the integral from the far past, where $p\eta$, $q_{1,2}\eta \gg \mu$ and the mode functions behave like plane waves with a pre-factor, are negligible. This can be traced back to the blueshift when moving towards the boundary with the CPP, and implies we can safely take $\eta_0 \to \infty$ without encountering an IR divergence. In the CPP, it is not possible to take 0 to past infinity: One encounters explicit IR divergences.

We use the delta function to perform the integration over $\vec{q}_2$, and relabel $\vec{q}_1 \to \vec{q}_2$. A number of $f(|\vec{p}-\vec{q}|\eta_i)$ now appear in (6.11). Noting that the largest contribution comes from the regime $|\vec{q}| \gg |\vec{p}|$, we simplify further by ignoring the $\vec{p}$ in the argument of those harmonics. Furthermore, we can change the lower limit of integration over $\eta_{3,4}$ to μ/p, since contributions from the infinite past are negligible.† After further manipulations which do not contain any new physics and have therefore been omitted, we arrive at an estimate of the leading terms of n^1_p and κ^1_p. A detailed description of the necessary changes of variables etc. can be found in section IV of [20].

Our main interest is, of course, in estimating the IR correction (6.11) in the Bunch-Davies state. There are no IR divergences in the EPP, but we do find a one-loop-contribution that is large as $p\eta \to 0$, signaling a large IR contribution (in calculations, 1 and 2 are nite, so no real divergence occurs). The leading contributions are:

$$n^1_p \sim \lambda^2 \log\left(\frac{\mu}{p\eta}\right)$$
$$\kappa^1_p \sim -\lambda^2 \log\left(\frac{\mu}{p\eta}\right) \qquad (6.12)$$

A similar result was derived by Polyakov and Krotov in [7]. As we see, n^1_p and κ^1_p give a non-negligible contribution, in stark contrast with the flat space case where they remain zero forever. The most difficult and perhaps most important question is how to interpret this result. Akhmedov has argued [11, 20] that κ^1_p and n^1_p should be interpreted as the comoving particle density and measure of the backreaction due to quantum fluctuations respectively, which are initially zero (if we start in an exact vacuum state) but grow with time. We will return to this point in the upcoming sections.

6.3.2 CORRECTIONS TO OTHER QUANTITIES

Consider the one-loop correction to the Retarded propagator. As it turns out, there is only a single diagram (shown in fig. 6.2) that contributes. The same diagram, but mirrored left-to-right, represents the one-loop correction to the Advanced propagator. We can quickly see that this correction will not be large. The Retarded propagators that appear in the diagram impose $\eta_2 \to \eta_4 \to \eta_3 \to \eta_1$. Because we are taking a limit such that the ratio between η_1 and η_2 is constant, this integral never becomes large. An similar argument can be made for the correction to the Advanced propagator, since the contributing diagram is in essence the same. Finally, we should mention that it can be shown that the vertex function does not receive any large one-loop IR corrections either. This 'calculation', which amounts to nothing more than observing that none of the contributing terms diverge if we simply set all physical momenta to zero, is shown explicitly in [20].

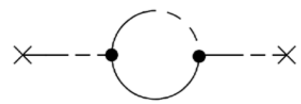


Figure 6.2: *The only diagram that contributes to the one-loop correction of the Retarded propagator.*

6.4 SUMMING THE INFRARED CONTRIBUTIONS

Even though we found that there are large IR loop corrections to the Keldysh propagator, one might wonder why the above calculations are relevant: Did we not just show that the BD vacuum is de Sitter invariant at loop level? Does this not preclude any explosive particle production or similar effects which could cause a large backreaction?

It is true that there does not seem to be an instability of the expanding patch of de Sitter space if the initial state is the exact BD vacuum. However, the large IR contributions to the Keldysh propagator seem to indicate that there is some interesting IR physics going on. To better understand these effects, we can investigate what happens if the initial state is a (non-symmetric) perturbation of the exact BD state. Although one could argue that this is a somewhat 'artificial' problem (the de Sitter invariance is broken 'by hand' at the level of the initial conditions), it might be interesting to track the evolution of such a perturbation as the large IR contributions are generated. The question then becomes whether the BD vacuum is stable under such perturbations: Does time evolution drive us back to a de Sitter invariant state, or is there 'runaway behavior'?

To investigate this, Akhmedov and others have attempted to sum the leading logarithmic IR contributions from all loops. This only needs to be done for the Keldysh propagator: We saw that the other quantities have no large one-loop contributions, while higher-loop corrections will be suppressed by additional powers of λ. Therefore, one can approximate them by using the tree level expressions when solving for the exact Keldysh propagator. To achieve this, we need to set up and solve the so-called Dyson-Schwinger equation, which provides a way of finding the exact propagator directly. We will not reproduce this lengthy calculation here, providing a qualitative discussion of the main points instead. For the mathematical details, we once again refer the reader to [20].

Currently, solving the Dyson-Schwinger equation for the Keldysh propagator has only been achieved by making an ansatz that induces a considerable simplication in the IR limit (6.9). This ansatz once again defines the Keldysh propagator in terms of two quantities n_p and κ_p in a fashion that is similar to the one-loop calculation. For technical reasons, this simplication is only possible if one considers fields in the principal series ($m \to (N-1)/2$). From now on, we will therefore restrict ourselves to this class of fields. It is not known how to find a solution for fields that belongs to the complementary series.

From the Dyson-Schwinger equation, it is then possible to derive approximate integro-dierential equations describing the evolution of n_p and κ_p. In order to form a physical picture of what they represent it is instructive to once again examine the at space limit. Then $\frac{dn_p}{dt}$ and $\frac{d\kappa_p}{dt}$ can be related to a number of terms which can each be interpreted as a concrete particle production and decay process. Based on these equations, Akhmedov concludes that n_p should be interpreted as the comoving particle density while κ_p is thought to represent a measure of the backreaction of the quantum field on the gravitational background. Similar arguments are presented by Polyakov in [8]. Like in the previous section, energy conservation imposes that the expressions representing these processes vanish in flat space, so that a nonzero initial n_p and κ_p would not grow in time.

We now want to find a solution for n_p and κ_p such that $\kappa_p(\eta)$ is (almost) zero. This is a consistency requirement: If the backreaction were large one needs to also include other things such as the graviton in all calculations. As it turns out, it is possible to find a solution such that $\kappa_p(\eta) \to 0$ as time progresses if one uses the out-harmonics. This result holds even if κ_p is initially small but nonzero. Thus, the setup is as follows: We start with some perturbation of the exact BD state and use a Bogolyubov transformation to the out-harmonics, in terms of which we solve the integro-differential equations for n_p and κ_p. The solution for n_p depends strongly on the initial conditions. If we start with a small perturbation, n_p will remain constant. If we accept the interpretation as a comoving particle density, this means that the de Sitter invariance, which was broken by the initial conditions, is asymptotically restored as the particles are diluted away.

However, for a strong enough perturbation of the exact BD state, there is a different solution for n_p:

$$n_p(\eta) \sim \frac{1}{\log(\eta/\eta_\star)}, \qquad \eta_\star \equiv \frac{\mu}{p} e^{-C}, \quad \frac{\mu}{p} > \eta > \eta_\star \tag{6.13}$$

6.5 CONCLUSIONS

In conclusion, we have obtained the following results: Firstly, the exact Bunch-Davies vacuum is de Sitter invariant at loop level in a general interacting theory. However, when we consider the one-loop contribution to the Keldysh propagator, it turns out that there are large IR contributions. These IR contributions vanish in the flat space limit, indicating that they are connected to the presence of a gravitational background field. The fact that they vanish because of conservation of energy in Minkowski space hints at an interpretation in terms of interactions between the quantum field and the gravitational background, which allows for a number of particle decay and production processes that would otherwise be impossible.

Similarly, when summing the leading logarithms the flat space limit seems to invite an interpretation in terms of particle processes. It seems clear that n_p and κ_p capture some new physics that does not play a role in flat space. However, it is notoriously dangerous to interpret results from QFT with a gravitational background in terms of particles. The flat space limit can be particularly deceptive in this respect. More work should be done to elucidate the physical meaning of the variables n_p and κ_p before one can have full confidence about the conclusions that follow from this interpretation.

The last, and perhaps most surprising, result that we discussed is that it is possible to construct a state that shows such strong IR effects that they outrace the rapid expansion of space. When we start with a small perturbation of the exact Bunch-Davies vacuum that is put in *by hand* to break the de Sitter isometry, there is a tendency to return to the exact Bunch-Davies vacuum. However, when this perturbation is large enough, one can obtain a runaway solution that signals an unstable situation.

The physical signicance of this infrared surprise is yet unclear. On the one hand, we can ask how surprising it really is that one can induce an instability by assuming a strongly perturbed initial state. On the other hand, it may be very well possible that there was a large local perturbation somewhere in our large (perhaps even infinite) universe. If this leads to a small unstable patch, could the instability spread and ultimately affect the entire spacetime? To answer such questions reliably, we need to go beyond the regime where semiclassical methods are appropriate because of the large backreaction that is to be expected. This would clearly be a long-term goal.

At a more 'accessible' level, there are a number of things that remain to be done as well. As mentioned before, it would be interesting to find out with more certainty how to interpret the infrared physics captured by the variables n_p and κ_p. Summing the leading contribution for fields that belong to the complementary series is another open question that is worth pursuing. Despite the many unresolved problems, some have already attempted to incorporate the idea of a decaying cosmological constant due to particle production into quantitative models [49].

Finally, it should be noted that Polyakov has focused particularly on deriving results for global de Sitter space, arguing that the full space must be understood before considering only the expanding section [7, 8]. In global de Sitter space, stronger results have been achieved than in the expanding Poincare patch alone [8, 13]. Intuitively, this could be expected because of the blueshift associated with the contracting section of de Sitter space. In the absence of strong evidence that global de Sitter space is indeed the right way to describe our universe, it remains unclear what the physical relevance of these results is.

All in all, the results seem inconclusive. There are some interesting infrared effects and it is too early to rule out the viability of this dynamical screening mechanism of the cosmological constant, but proclaiming that de Sitter space is fundamentally unstable seems premature to say the least. As so often in physics, much seems to depend on the initial state, in this case that of our universe. To paraphrase Hawking: It's initial conditions all the way down![‡]

‡ In Hawking's popular book `A Brief History of Time' he opens with a humorous anecdote about a theory of the universe which postulates that everything rests on the back of a cosmic turtle, which itself rests on another, larger turtle, etc., leading to the conclusion that 'it's turtles all the way down!'

APPENDIX A

DERIVATION OF SEMICLASSICAL FIELD EQUATIONS

Despite tremendous effort by many physicists throughout the past decades, it has thus far proven impossible to formulate a satisfactory theory of quantum gravity. In order to nonetheless obtain results that are beyond pure GR, we therefore have to resort to an approximate or effective description. The study of a quantum field interacting with a classical yet dynamical general relativistic background, known as semiclassical gravity, is the first step in this direction. The techniques of semiclassical gravity are sufficiently powerful to yield some important theoretical results such as Hawking's famous discovery of radiation coming from black holes. We can use the same methods to study e.g. the Maxwell equations in the presence of a quantum field. Here, we will present a brief introduction to semiclassical eld theory. This appendix should ideally be read after chapter 5, as it draws upon several results from earlier chapters.

When considering the interaction between the quantum field Φ and a classical background field J, we have to account for the influence that both fields exert on each other. In fact, we are already familiar with the way in which the path integral formulation incorporates the influence of a classical 'source' J on the expectation value of quantum observables from regular QFT. In the special case where the external classical background field vanishes (i.e. the quantum field is in the vacuum state with no classical source), it is given by (5.6) or, in the non-equilibrium case, (5.15). If one does not set $J = 0$, the formula immediately generalizes to nonzero background fields.

Note that it is always assumed that Φ is initially in the vacuum state. Understanding the influence of the quantum field on the classical background, known as the backreaction, is slightly less simple. We will now present two different ways of deriving the backreaction, following the treatment by Mukhanov and Winitzki [19].

A.1 HEURISTIC DERIVATION OF THE BACKREACTION

In this section, we will present a heuristic derivation of the backreaction, with little regard for rigor. A more rigorous derivation is presented in the next section. In order to understand the influence of the quantum field on the classical background we must of course treat this field as dynamically. There is a corresponding 'background action' $S_{BG}[J]$. In the absence of any interaction with the quantum field the equation of motion is simply

$$\frac{\delta S_{\mathrm{BG}}[J]}{\delta J(x)} = 0$$

When we consider a full, interacting system, the total action is

$$S_{\mathrm{total}} = S_{\mathrm{BG}}[J] + S[\phi, J] \tag{A.1}$$

When we quantize only the field, the dynamics of J are characterized by an eective action which only depends on J, yet incorporates the influence of the quantum field. If the -eld were not quantized, we could simply use the classical solution, but for a quantum field other congurations contribute as well. This motivates us to perform a path integral over the quantum field Φ. We define the effective action of the full system as follows:

$$\exp\left[iS_{\mathrm{eff}}[J]\right] = \int \mathcal{D}\phi \exp\left[i\left(S_{\mathrm{BG}}[J] + S[\phi, J]\right)\right] = \exp\left(iS_{\mathrm{BG}}[J] + i\Gamma[J]\right) \tag{A.2}$$

where we have introduced the quantum effective action $\Gamma[J]$, defined by

$$\exp\left(i\Gamma[J]\right) = \int \mathcal{D}\phi \exp\left(iS[\phi, J]\right)$$

The equations of motion become

$$\frac{\delta S_{\mathrm{eff}}}{\delta J(x)} = \frac{\delta S_{\mathrm{BG}}[J]}{\delta J(x)} + \frac{\delta \Gamma[J]}{\delta J(x)} = 0 \tag{A.3}$$

There is one more thing that should be noted: Remembering the identity (5.5), we see that the functional derivative of the quantum effective action contains the Feynman propagator. However, this propagator is not causal. To impose causality, we have to substitute the Feynman propagator with the retarded propagator. Our final result is:

$$\frac{\delta S_{\mathrm{eff}}}{\delta J(x)} = \frac{\delta S_{\mathrm{BG}}[J]}{\delta J(x)} + \left.\frac{\delta \Gamma[J]}{\delta J(x)}\right|_{G_F \to G_R} = 0 \tag{A.4}$$

This procedure seems reasonable, but one needs to keep in mind that it does not really make any sense to `mix' classical and quantum degrees of freedom. Next, we will present a more rigorous way of deriving the same resul.

A.2 DERIVATION FROM A FULLY QUANTIZED THEORY

To obtain a consistent semiclassical description we need to start from a fully quantized theory, subsequently making an appropriate approximation where one of the degrees of freedom can be treated classically. We will now present a sketch of such a derivation based on [50].

Consider a quantum field on a classical background J, each with a single degree of freedom, in a Heisenberg (i.e. stationary) state $|\psi\rangle$ where the background variable J is approximately classical with small quantum fluctuations around its expectation value, while is approximately in the vacuum conguration. Formally, this is represented by:

$$\left\langle \psi \middle| \hat{J} \middle| \psi \right\rangle = J, \quad \hat{J} = J + \hat{j}, \qquad \left\langle \psi \middle| \hat{\phi} \middle| \psi \right\rangle \approx 0 \tag{A.5}$$

Here, $\hat{J}$ is small compared to the expectation value J. To avoid confusion with expectation values, we have introduced the hat-notation for operators. We now expand the equation of motion up to second order around the classical situation with Φ in the classical vacuum state: $\hat{J} = J$, and $\hat{\Phi} = 0$

$$\frac{\delta S[\hat{\phi},\hat{J}]}{\delta \hat{J}(x)} = 0$$

$$\begin{aligned} &= \frac{\delta S[0,J]}{\delta J(x)} + \int d^N y \left[\frac{\delta^2 S[0,J]}{\delta J(x)\delta J(y)}\hat{j}(y) + \frac{\delta^2 S[0,J]}{\delta J(x)\delta\phi(y)}\hat{\phi}(y)\right] \\ &\quad + \frac{1}{2}\int d^N y\, d^N z \left[\frac{\delta^3 S[0,J]}{\delta J(x)\delta J(y)\delta J(z)}\hat{j}(y)\hat{j}(z)\right. \\ &\qquad + \frac{\delta^3 S[0,J]}{\delta J(x)\delta J(y)\delta\phi(z)}\hat{j}(y)\hat{\phi}(z) \\ &\qquad \left. + \frac{\delta^3 S[0,J]}{\delta J(x)\delta\phi(y)\delta\phi(z)}\hat{\phi}(y)\hat{\phi}(z)\right] + \dots \end{aligned} \quad \text{(A.6)}$$

Now, we take the expectation value of this equation in the ground state of the quantum field: $|\psi\rangle = |0\rangle$. The linear terms immediately drop out by virtue of (A.5). Furthermore, we disregard the second order terms containing $\hat{j}$ since we are interested in the limit where fluctuations of $\hat{j}$ are negligible. We are then left with

$$\left\langle 0 \left| \frac{\delta S[\hat{\phi},\hat{J}]}{\delta \hat{J}(x)} \right| 0 \right\rangle \approx \frac{\delta S[0,J]}{\delta J(x)} + \frac{1}{2}\int d^N y\, d^N z\, \frac{\delta^3 S[0,J]}{\delta J(x)\delta\phi(y)\delta\phi(z)} \left\langle 0 \left| \hat{\phi}(y)\hat{\phi}(z) \right| 0 \right\rangle = 0 \quad \text{(A.7)}$$

From this procedure it is clear that the semiclassical equations incorporate only the rst order quantum corrections to the dynamics, since we are ignoring all terms of higher order. The first term in the above equation clearly corresponds to the rst term in (A.4). We just need to show that the second term equals $\frac{\delta\Gamma}{\delta J}$ in order to recover the same result. We can start by writing the two-point (Wightman) function in terms of the Feynman and retarded propagators:

$$\left\langle 0 \left| \hat{\phi}(y)\hat{\phi}(z) \right| 0 \right\rangle = \frac{1}{i}G_F(z,y) - \frac{1}{i}G_R(z,y) \quad \text{(A.8)}$$

Assuming that we are starting from an action that respects locality, the second term will contain a factor $\delta^N(y-z)$. Recalling (5.20), we see that $G_R(y,y) = 0$ so that we can drop this term, leaving us with

$$\left\langle 0 \left| \frac{\delta S[\hat{\phi},\hat{J}]}{\delta \hat{J}(x)} \right| 0 \right\rangle \approx \frac{\delta S_{\mathrm{BG}}[J]}{\delta J(x)} + \frac{1}{2i}\int d^N y\, d^N z\, \frac{\delta^3 S[0,J]}{\delta J(x)\delta\phi(y)\delta\phi(z)} G_F(z,y) = 0 \quad \text{(A.9)}$$

In order to show that this second term is equal to the functional derivative of $\Gamma[J]$ it is convenient to view $\Gamma[J]$ as a functional determinant, a concept which we will now introduce.

A.2.1 QUANTUM EFFECTIVE ACTION AS A FUNCTIONAL DETERMINANT

It is easier to carry out the following calculations by switching to Euclidean signature. We analytically continue the time variable to imaginary values $t = i\tau$ (this is also known as the Wick rotation to Euclidean time τ), and define the Euclidean path integral and quantum effective action

$$\int \mathcal{D}\phi\, \exp(-S_E[\phi,J]) = \exp(-\Gamma_E[J]), \qquad S_E[\phi,J] = -iS[\phi,J]\Big|_{t=i\tau} \quad \text{(A.10)}$$

The transition to Euclidean signature is not completely trivial, as we need to switch to Euclidean field variables (e.g. Φ^E, J^E or other fields like $g^E_{\mu\nu}$, A^E_μ in terms of $x^{E,\mu}$. However, these technical details will not be important to us; we refer the interested reader to [19]. We will drop the superscript denoting Euclidean variables in the rest of this section.

It is important to note that the measure $\mathcal{D}\Phi$, in the way it is usually defined in deriving the path integral in QFT, is not generally covariant. In order to cast it in a generally covariant form, we can use an expansion in terms of eigenfunctions of a covariant operator. The action can then be interpreted as a determinant of this dierential operator. Recall that the action of a free field and the corresponding equation of motion can be written as

$$S_E[\phi] = \frac{1}{2}\int d^N x\, \phi(x)\, \hat{A}\, \phi(x) \qquad \hat{A}\, \phi(x) = 0 \quad \text{(A.11)}$$

Where

$$\hat{A} = \frac{\delta^2 S_E[\phi]}{\delta\phi(x)\delta\phi(y)}$$

is a differential operator. Now, let's consider the following eigenvalue problem:

$$\hat{A}\, \phi(x) = \lambda_n \phi(x) \quad \text{(A.12)}$$

We impose boundary zero conditions in a nite box for convenience, as this results in a discrete spectrum. Under reasonable assumptions λ_n is bounded from below and we have a complete orthonormal basis of eigenfunctions in the infinite dimensional space of solutions [19]:

$$(\phi_m, \phi_n) \equiv \int d^N x \sqrt{|g|}\, \phi_m(x)\phi_n(x) = \delta_{mn}$$

Here, (.,.) denotes the inner product. We can use the completeness of the space to express any function in terms of the eigenfunctions:

$$\phi(x) = \sum_{n=0}^{\infty} C_n \phi_n(x), \qquad C_n \equiv (\phi(x), \phi_n(x))$$

This allows us to rewrite the action:

$$S_E[\phi] = \frac{1}{2}\int d^N x \sqrt{|g|} \sum_{m,\, n} C_m\phi_m C_n \lambda_n \phi_n = \frac{1}{2}\sum_n C_n^2 \lambda_n \quad \text{(A.13)}$$

This expression is coordinate independent because the C_n are given by covariant integralswhile the λn are eigenvalues of a covariant operator. We can now define the measure in terms of the coffecients C_n:

$$\mathcal{D}\phi = \prod_n f(C_n) dC_n$$

The simplest choice is $f(C_n)$ = constant, and because of the Gaussian integrals, $f(C_n) = (2\pi)^{-1/2}$ gives the most elegant expression for the Euclidean path integral:

$$\int \mathcal{D}\phi \exp\bigl(-S_E[\phi]\bigr) = \int \prod_n \frac{dC_n}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\, C_n^2\, \lambda_n\right] = \left[\prod_n \lambda_n\right]^{-\frac{1}{2}} \tag{A.14}$$

In terms of the Euclidean effective action we have

$$Z_E[J] = e^{-\Gamma_E[J]} \longrightarrow \Gamma_E[J] = \frac{1}{2} \ln\left[\prod_n \lambda_n\right] \tag{A.15}$$

In a finite dimensional space, it is clear that the product of eigenvalues of a generally covariant operator is equal to its determinant. Working under the assumption that we can make an appropriate generalization to infinite-dimensional operators, we can therefore rewrite Γ_E as the functional determinant of a differential operator:

$$\Gamma_E[J] = \frac{1}{2} \ln \det \hat{A} = \frac{1}{2} \ln \det\left[\frac{\delta^2 S_E[\phi, J]}{\delta\phi(x)\delta\phi(y)}\right] \tag{A.16}$$

It should be noted that these formal manipulations should technically be put on a more rigorous basis. For instance, we need to use a regularization scheme like zeta function regularization to properly define the functional determinant (see, for instance, [51]). However, these technical problems are outside the scope of this work, so we push on. After analytically continuing back to Lorentzian signature using the relation (A.10) we obtain

$$\Gamma[J] = \frac{1}{2i} \ln \det\left[\frac{\delta^2 S[\phi, J]}{\delta\phi(x)\delta\phi(y)}\right] \tag{A.17}$$

Now, we just need to calculate the functional derivative with respect to J. For this, we will need the general identity

$$\delta(\det \hat{A}) = \det \hat{A}\ \mathrm{Tr}(\hat{A}^{-1}\delta\hat{A}) \tag{A.18}$$

This identity is quite simply derived:

$$\delta(\det \hat{A}) = \det\bigl(\hat{A} + \delta\hat{A}\bigr) - \det \hat{A} = \det \hat{A}\ \ \det\bigl(\mathbb{1} + \hat{A}^{-1}\delta\hat{A}\bigr) - \det \hat{A}$$

The next, crucial step is to realize that

$$\det\bigl(\mathbb{1}_{ij} + \epsilon_{ij}\bigr) = 1 + \sum_i \epsilon_{ii} + \mathcal{O}(\epsilon_{ij}^2) \approx 1 + \mathrm{Tr}(\epsilon_{ij})$$

where we have assumed $\epsilon_{ij} \ll 1$ so that second and higher order terms can be neglected. The sought-after identity (A.18) then follows immediately. Recalling that $\hat{A}^{-1}$ is the Feynman propagator, we determine the functional derivative to be

$$\frac{\delta\Gamma[J]}{\delta J(x)} = \frac{\det \hat{A}}{2i \det \hat{A}}\ \mathrm{Tr}\left[G_F \frac{\delta \hat{A}}{\delta J(x)}\right] = \frac{1}{2i} \int d^N y\ d^N z\ \frac{\delta^3 S[\phi, J]}{\delta J(x)\delta\phi(y)\delta\phi(z)} G_F(y, z) \tag{A.19}$$

Which, after once again imposing causality, yields our main result:

$$\frac{\delta S_{\mathrm{BG}}[J]}{\delta J(x)} + \left.\frac{\delta\Gamma[J]}{\delta J(x)}\right|_{G_F \to G_R} = 0$$

Although this alternative derivation is more satisfactory than the heuristic justication from the previous section, we should keep in mind that, in the case of a gravitational background field, the full quantum theory is unknown. Strictly speaking, we therefore have to resort to the heuristic justification in this case.

A.3 EXAMPLES OF SEMICLASSICAL FIELD EQUATIONS

Although we worked out the above derivations working with a scalar background field $J(x)$, the results hold just as well for a vector or tensor field. We will now give two important examples of semiclassical field equations.

A.3.1 SEMICLASSICAL MAXWELL'S EQUATIONS

Classically, Maxwell's equations are derived from the total Lagrangian

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{EM}} + \mathcal{L}_{\mathrm{matter}}, \qquad \mathcal{L}_{\mathrm{EM}} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \quad \mathcal{L}^{\mathrm{cl}}_{\mathrm{matter}} = -j^\mu A_\mu \tag{A.20}$$

where $F_{\mu w} = \partial_\mu A_\nu - \partial_\nu A_\mu$ – is the electromagnetic field strength tensor. We see that

$$\frac{\delta S_{\mathrm{EM}}}{\delta A_\mu} = -\partial_\nu F^{\mu\nu} \qquad \frac{\delta S^{\mathrm{cl}}_{\mathrm{mat}}}{\delta A_\mu} = -j^\mu \tag{A.21}$$

When the matter action is generated by a complex (i.e. charged) scalar quantum field rather than classical sources, requiring local $U(1)$ gauge invariance leads to the following matter Lagrangian [25]:

$$\mathcal{L}^{\mathrm{q}}_{\mathrm{matter}} = D_\mu\psi D^\mu\psi^* + m^2\psi\psi^*, \quad D_\mu\psi \equiv (\partial_\mu + ieA_\mu)\psi,\ D_\mu\psi^* \equiv (\partial_\mu - ieA_\mu)\psi^* \tag{A.22}$$

The conserved Noether current corresponding to this symmetry is

$$\mathcal{J}^\mu = i\bigl(\psi^* D^\mu\psi - \psi D^\mu\psi^*\bigr), \qquad \partial_\mu \mathcal{J}^\mu = 0$$

This is the covariant quantum analog of the classical current j^μ. Inspired by (A.21) we now define

$$-\langle 0\,|\mathcal{J}^\mu|\,0\rangle = \frac{\delta\Gamma[A_\nu]}{\delta A_\mu} \tag{A.23}$$

Which yields the semiclassical Maxwell's equations:

$$\partial_\nu F^{\mu\nu} = -\langle 0\,|\mathcal{J}^\mu|\,0\rangle \tag{A.24}$$

A.3.2 SEMICLASSICAL GRAVITY

In the case of a real scalar field on a gravitational background, the inverse metric $g^{\mu\nu}$ takes on the role of J. The usual gravitational equations of motion are

$$\frac{\delta S_{\mathrm{grav}}[g_{\kappa\lambda}]}{\delta g^{\mu\nu}} + \frac{\delta S_{\mathrm{matter}}[g_{\kappa\lambda}, \ldots]}{\delta g^{\mu\nu}} = 0 \tag{A.25}$$

The gravitational component is the usual Einstein-Hilbert action

$$S_{\mathrm{grav}}[g_{\kappa\lambda}] = \int d^N x \sqrt{|g|}\left[\frac{1}{2}R - \Lambda\right] \longrightarrow \frac{\delta S_{\mathrm{grav}}[g_{\kappa\lambda}]}{\delta g^{\mu\nu}} = \frac{\sqrt{|g|}}{2}\left[R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} + \Lambda g_{\mu\nu}\right] \tag{A.26}$$

while we define the energy-momentum tensor in terms of the matter action [21]:

$$T_{\mu\nu} \equiv -\frac{2}{\sqrt{|g|}}\frac{\delta S_{\text{matter}}[g_{\kappa\lambda}, \ldots]}{\delta g^{\mu\nu}} \tag{A.27}$$

We obtain the Einstein equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = T_{\mu\nu} \tag{A.28}$$

If we use the action of the real scalar field as the matter component we take the same approach as in the previous example. We define the quantum energy-momentum tensor to satisy

$$\langle 0 | T_{\mu\nu} | 0 \rangle = -\frac{2}{\sqrt{|g|}}\frac{\delta\Gamma[g_{\kappa\lambda}]}{\delta g^{\mu\nu}} \tag{A.29}$$

so that the semiclassical Einstein equations read

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \langle 0 | T_{\mu\nu} | 0 \rangle \tag{A.30}$$

As can be seen from these examples, the general recipe to obtain semiclassical equations of motion is to simply replace classical source terms with vacuum expectation values of the corresponding tensors generated by the quantum matter field(s). Much of semiclassical field theory is devoted to finding good ways to estimate these expectation values.

REFERENCES

[1] S. Perlmutter et al. Measurements of Omega and Lambda from 42 high redshift supernovae. *ApJ*, 517:565{586, 1999. arXiv:astro-ph/9812133.

[2] P.A.R. Ade et al. BICEP2 I: Detection Of B-mode Polarization at Degree Angular Scales. 2014. arXiv:astro-ph/1403.3985.

[3] S. Weinberg. Anthropic bound on the cosmological constant. *Phys. Rev. Lett.*, 59:2607{2610, 1987.

[4] I. Antoniadis, P. O. Mazur, and E. Mottola. Cosmological dark energy: Prospects for a dynamical theory. *New J. Phys.*, 9:11, 2007. arXiv:gr-qc/0612068.

[5] A. M. Polyakov. De Sitter space and eternity. *Nucl. Phys. B*, 797(12):199 { 217, 2008. arXiv:hep-th/0709.2899.

[6] A. M. Polyakov. Decay of vacuum energy. *Nucl. Phys. B*, 834(12):316 { 329, 2010. arXiv:hep-th/0912.5503.

[7] D. Krotov and A. M. Polyakov. Infrared sensitivity of unstable vacua. *Nucl. Phys. B*, 849(2):410 { 432, 2011. arXiv:hep-th/1012.2107.

[8] A. M. Polyakov. Infrared instability of the de Sitter space. 2012. arXiv:hepth/1209.4135.

[9] P. R. Anderson and E. Mottola. On the Instability of Global de Sitter Space to Particle Creation. 2013. arXiv:gr-qc/1310.0030.

[10] P. R. Anderson and E. Mottola. Quantum Vacuum Instability of 'Eternal' de Sitter Space. 2013. arXiv:gr-qc/1310.1963.

[11] E. T. Akhmedov. IR divergences and kinetic equation in de Sitter space. (Poincare patch; Principal series). *JHEP*, 1201:066, 2012. arXiv:hep-th/1110.2257.

[12] E. T. Akhmedov and Ph. Burda. Solution of the Dyson{Schwinger equation on de Sitter background in IR limit. *Phys. Rev. D*, 86:044031, 2012. arXiv:hepth/1202.1202.

[13] E. T. Akhmedov. Physical meaning and consequences of the loop infrared divergences in global de Sitter space. *Phys. Rev.* D, 87:044049, 2013. arXiv:hepth/1209.4448.

[14] E. T. Akhmedov, F. K. Popov, and V. M. Slepukhin. Infrared dynamics of the massive '4 theory on de Sitter space. *Phys. Rev. D*, 88(2):024021, 2013. arXiv:hep-th/1303.1068.

[15] D. Marolf and I. A. Morrison. The IR stability of de Sitter QFT: results at all orders. *Phys. Rev. D*, 84:044040. arXiv:hep-th/1010.5327.

[16] D. Marolf and I. A. Morrison. The IR stability of de Sitter: Loop corrections to scalar propagators. *Phys. Rev. D*, 82:105032. arXiv:hep-th/1006.0035.

[17] D. Marolf and I. A. Morrison. The IR stability of de Sitter QFT: Physical initial conditions. *Gen. Relat. Gravit.*, 43(12):3497{3530. arXiv:hep-th/1104.4343.

[18] D. Marolf, I. A. Morrison, and M. Srednicki. Perturbative S-matrix for massive scalar fields in global de Sitter space. *Classical Quant. Grav.*, 30(15):155023. arXiv:hep-th/1209.6039.

[19] V. F. Mukhanov and S. Winitzki. *Introduction to quantum effects in gravity*. Cambridge University Press, 2007.

[20] E.T. Akhmedov. Lecture notes on interacting quantum fields in de Sitter space. *Int. J. Mod. Phys. D*, 23(1), 2014. arXiv:hep-th/1309.2557.

[21] S. M. Carroll. *Spacetime and Geometry: An introduction to General Relativity.* Addison Wesley, 2004.

[22] R. M. Wald. *General Relativity*. University of Chicago Press, 1984.

[23] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. Gravitation. W. H. Freeman, 1973.

[24] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1973.

[25] L. H. Ryder. *Quantum Field Theory*. Cambridge University Press, 1996.

[26] M. Srednicki. *Quantum Field Theory.* Cambridge University Press, 2007.

[27] S.Weinberg. *The Quantum Theory of Fields*: Foundations. Cambridge University Press, 1995.

[28] N. D. Birrell and P. C. W. Davies. *Quantum fields in curved space*. Cambridge University Press, 1982.

[29] R. G. Cai and S. P. Kim. First law of thermodynamics and Friedmann equations of Friedmann-Robertson-Walker universe. *JHEP*, 0502:050, 2005. arXiv:hepth/0501055.

[30] M. Spradlin, A. Strominger, and A. Volovich. Les Houches Lectures on De Sitter Space. 2001. arXiv:hep-th/0110007.

[31] T. S. Bunch and P. C. W. Davies. Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting. *Proc. Roy. Soc. Lond. A*, 360:117{134, 1978.

[32] Conformal vacua and entropy in de Sitter space. *Phys. Rev. D*, 65:104039, 2002. arXiv:hep-th/0112218.

[33] Tagirov E. A. Chernikov, N. A. Quantum theory of scalar fields in de Sitter space-time. *Ann. Poincare Phys. Theor. A*, 9:109, 1968.

[34] J. B. Hartle and S. W. Hawking. Wave function of the universe. *Phys. Rev. D*, 28:2960{2975, 1983.

[35] P. Candelas and D. J. Raine. General Relativistic Quantum Field Theory - An Exactly Soluble Model. *Phys. Rev. D*, 12:965{974, 1975.

[36] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1964.

[37] B. Allen. Vacuum states in de Sitter space. *Phys. Rev. D*, 32:3136{3149, 1985.

[38] I. S Gradshtein, I. M Ryzhik, A. Jerey, and D. Zwillinger. *Table of integrals, series and products*. Academic Press, 2007.

[39] J. S. Schwinger. Brownian motion of a quantum oscillator. *J. Math. Phys.*, 2:407{432, 1961.

[40] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Zh. Eksp. Teor.* Fiz., 47:1515{1527, 1964.

[41] J. Rammer. *Quantum field theory of non-equilibrium states*. Cambridge University Press, 2010.

[42] E. Calzetta and B. L. Hu. Nonequilibrium Quantum Fields: Closed Time Path Effective Action, Wigner Function and Boltzmann Equation. *Phys. Rev. D*, 37:2878, 1988.

[43] S. J. N. Mooij. Effective theories in cosmology. Unpublished doctoral dissertation,
University of Amsterdam, 2013.

[44] A. Kamenev. Many-body theory of non-equilibrium systems. 2004. arXiv:condmat/0412296.

[45] A. Kamenev and A. Levchenko. Keldysh technique and non-linear sigma-model: Basic principles and applications. *Adv. Phys.*, 58(197), 2009. arXiv:condmat/0901.3586.

[46] M. van der Meulen and J. Smit. Classical approximation to quantum cosmological correlations. *JCAP*, 0711:023, 2007. arXiv:hep-th/0707.0842.

[47] J. S. Schwinger. On gauge invariance and vacuum polarization. *Phys. Rev.*, 82:664{679, 1951.

[48] L. Leblond and E. Pajer. Resonant Trispectrum and a Dozen More Primordial N-point functions. *JCAP*, 1101:035, 2011. arXiv:hep-th/1010.4565.

[49] F. R. Klinkhamer. On vacuum-energy decay from particle production. *Mod. Phys. Lett. A*, 27(25):1250150. arXiv:hep-th/1205.7072.

[50] A. O. Barvinsky and D. V. Nesterov. Effective equations in quantum cosmology. *Nucl. Phys. B*, 608:333{374, 2001. arXiv:gr-qc/0008062.

[51] A.A. Bytsenko, G. Cognola, E. Elizalde, V. Moretti, and S. Zerbini. Analytic *Aspects of Quantum Fields*. World Scientific, 2003.