

Real-time Control of Angioplasty Balloon Inflation for Beating Heart Model

Akash K Singh, PhD
IBM Corporation Sacramento, USA

Abstract

We report on real-time control of balloon inflation inside porcine arteries. In a first step, experiments were done in a coronary artery of an excised heart. In a second step, experiments were done in a beating heart setup providing conditions very close to *in vivo* conditions without the complications. A programmable syringe pump was used to inflate a compliant balloon in arteries, while intravascular optical coherence tomography (IVOCT) monitoring was performed. In a feedback loop, IVOCT images were processed to provide the balloon diameter values in real-time to control the pump action in order to achieve a target diameter. In different experiments, various flow rates and target diameters were used. In the excised heart experiment, there was good convergence to target diameters resulting in a satisfactory balloon inflation control. In the beating heart experiment, there were oscillations in the diameter values due to cyclic arterial contractions. In these experiments, the control system maintained diameter averages satisfactorily close to predetermined target values. Real-time control of balloon inflation could not only provide a safer outcome for angioplasty procedures but could also provide additional information for diagnostics since it implicitly provides information about the artery response to the inflation process.

Keywords- Angioplasty, control, optical coherence tomography (OCT), real-time systems.

I. INTRODUCTION

Atherosclerosis is a disease in which accumulation of plaque in the walls of the artery restrains the flow of oxygen-rich blood and appropriate feeding of organs [1]. Intravascular imaging techniques, such as intravascular ultrasound (IVUS) [2] and intravascular optical coherence tomography (IVOCT) [3-5], have been applied as techniques with resolutions superior to X-ray fluoroscopy to visualize the artery walls and plaques. Some clinician have proposed the use of IVUS [6, 7] and IVOCT [8] to verify the results of treatments such as angioplasty. Intravascular balloon inflation has been applied in different medical procedures, such as angioplasty. In these procedures, balloon inflation is usually performed

manually. Computerized balloon inflation has been proposed [9-11] with the aim to improve angioplasty results by reducing arterial injury which is linked to undesired phenomena, such as restenosis [12]. Restenosis is the renarrowing of the artery which may happen after the intervention. Previously [13], we proposed a method to control the luminal diameter of arteries during angioplasty balloon inflation. As a first experimental validation, we tested this method using a semi-compliant balloon and an artery phantom [14]. In the proposed method, the balloon inflation was controlled in order to achieve a target luminal diameter for the phantom. Using an edge detection algorithm, the lumen of the phantom was detected in IVOCT images that were continuously acquired during the inflation. The lumen diameter was estimated in real-time and compared with the target diameter. Based on this comparison, a controller sent commands to a programmable pump to deliver or withdraw liquid until the target diameter was achieved. The proposed control method could improve angioplasty results by reducing arterial injury during balloon inflation. It could also reduce the exposure to X-ray, as the guidance is partly provided by IVOCT. Further safety advantages include a more repeatable inflation procedure since conditions are better controlled and a constant visualization of the response of the artery wall to deformation since it is implicitly provided by the IVOCT imaging. In this study, we extend our experimental validation of balloon inflation control to porcine arteries, both in *ex vivo* and close to *in vivo* conditions. Experiments were performed using a compliant balloon. First, controlled inflation was performed in arteries of an excised porcine heart. The goal was to assess the performance of the control system in response to dynamics of the compliant balloons and real porcine arteries. Inflation control in real arteries was a step forward from our preliminary work on phantoms [13]. In a further advance, i.e., a beating heart setup, we simulated two realistic aspects of the *in vivo* condition, namely, the presence of a blood flow and the presence of cyclic arterial contractions during the heart beat. We investigated if the edge detection could be performed in presence of blood. We also investigated if the control system could robustly provide convergence to target diameters in presence of arterial contractions.

II. MATERIALS AND METHOD

A. Optical Coherence Tomography System

A custom-built SS-OCT system was used for imaging [15]. We used a wavelength-swept laser source (Santec, HSL2000), operating with a sweep rate of 30 kHz and a sweep range of over 108 nm around 1.33 μm wavelength to provide a measured axial resolution of about 15 μm in air. The SS-OCT system was configured as a Mach-Zehnder interferometer with balanced detection and was packaged as a mobile unit. SS-OCT imaging was performed at 30 frames per second.

B. IVOCT Imaging through a compliant Balloon

The IVOCT probe was composed of a single-mode fiber enclosed within a spiral metallic tube in the proximal region and in a flexible polymer tube in the distal region. Near the tip of the catheter, the light exiting from the fiber was focused by a gradient-index (GRIN) lens and was redirected by a right-angle prism. The ensemble was enclosed in a stainless steel ferrule with an outer diameter of 0.7 mm. The probe was inserted in a balloon catheter. The balloon catheter was based on a transparent polymer sheath that contained and protected the probe. A compliant balloon was used to deform the artery. The compliant balloon was a silicon membrane which was glued to the balloon catheter. A liquid was used to inflate the balloon through a few holes punched in the polymer sheath. The liquid also facilitated the rotation of the probe and provided a transparent medium for imagery. In our experiments, we used water for balloon inflation. A soft tip was attached to the distal end of the balloon catheter to help navigate the catheter through the artery without causing damage. IVOCT imaging of the vessel walls was performed through the balloon.

C. Porcine heart experiments

Balloon inflation control experiments were performed in porcine coronary arteries in two different setups, one using an excised heart and the other using a beating heart model.

- Excised heart experiment

Three days before the day of the experiment, a frozen porcine heart was allowed to defrost. On the day of the experiment, some excess segments of aorta were trimmed to facilitate access to coronary arteries. Through the aorta, the balloon catheter and the IVOCT probe were inserted in the left anterior descending (LAD) coronary artery for experiments.

- Beating heart experiment

The beating heart experiments were conducted according to regulations laid out by the Canadian Council on Animal Care and were approved by a local Animal Care Committee of the National Research Council of Canada. The details of the

beating heart model have been reported previously [15]. In this paper, the same model was used to investigate the performance of our inflation control method. Commercial swine was acclimatized in the animal facility, one week before experiments. On the day of the experiment, the animal was first anaesthetized. The chest was opened and the heart was arrested. The heart was separated, rinsed and put in a cold saline bath. It was cannulated and prepared for perfusion. It was then hung over a funnel and perfused with body temperature blood and Krebs-Heinseleit solution in equal proportions. Once the heart warmed up and showed signs of activity, it was defibrillated to establish a normal rhythm. The balloon catheter and the IVOCT probe were inserted through in the LAD artery through an introducer. The beating heart setup is illustrated in Fig. 2. The beating heart model was helpful in generating conditions similar to an *in vivo* setup without the associated complications of a full animal preparation. The setup remained functional for several hours. In our experiments, it allowed us to investigate the efficiency of our image analysis and control algorithms in presence of arterial contractions and blood flow.

D. IVOCT Images

In this section, we present a sample image obtained from a beating heart experiment. An IVOCT image in polar coordinates and in Cartesian coordinates, illustrating the balloon inside the artery. The balloon was not yet inflated. The arterial wall was not visible at many angles, where the blood was occluding the view by scattering the light. The polymer sheath and the balloon were represented each by two contours, corresponding to their inner and outer surfaces. The prism surface was represented by a contour, which was detected and used as a reference to register all images in the radial direction, as reported previously [13]. In order to process and analyze IVOCT images, we used the polar coordinates, providing the image in a matrix format. The horizontal and the vertical axes correspond to the probe rotation angle and the radial depth, respectively. The primary axes (left and below) represent pixel values. The values on the secondary horizontal axis (top) represent the angle in degrees. The values on the secondary vertical axis (right) represent the depth in millimeters in optical distance, i.e. the product of geometrical length and the refractive index. Each column in the image matrix represents depth scanning at a particular angle and is called an A-scan. Two sample A-scans, A-scan "A" and A-scan "B" are depicted by red lines. The "x" marks on these A-scans correspond to sample detected balloon nodes, used to characterize the diameter. A-scan A provides a sample depth profile for a segment of the image where visualization of the vessel wall was obstructed by blood. A-scan B provides a sample depth profile for

the segment of the image where the balloon was in the vicinity of the vessel wall and where there was almost no blood trapped between the balloon and the vessel wall. In this segment, better visualization of the vessel wall structure was provided; the vessel wall layers, namely, intima, media and adventitia could be distinguished. In our images, there were about 1000 A-scans per rotation, resulting in a 0.4° angular resolution. The radial pixel size in the presented images was 8 μm in optical distance.

E. Estimation of Luminal Diameter

In our control algorithm, we used the diameter of the balloon in the feedback loop. The balloon diameter was obtained from real-time detection of the contour corresponding to the outer surface of the balloon. Using this algorithm, we first selected a number of A-scans, distributed at equal angles over a full rotation. Each A-scan was averaged with A-scans within a neighborhood. A median filter was then applied to the result, followed by a gradient operator. The balloon contour node corresponded to an edge and was detected as the maximum of the gradient. A more detailed description of this algorithm was reported earlier [13]. In order to obtain the depth of the detected balloon nodes in geometrical distance, the prism surface was also detected as a reference. The prism surface appeared as a peak on each A-scan. Therefore, it was detected using a peak detection technique. Before proceeding to control experiments, we tested our algorithm on images acquired during manual inflation in a previous beating heart experiment, where no control was applied. Balloon detection in a sample image and in an image sequence in the form of a video clip (available at <http://ieeexplore.ieee.org>). In all images, 24 contour nodes were detected represented as red dots. Let j denote the index of a processed A-scan and ij denote the depth of the obtained balloon node on this A-scan in pixels. The depth value of the detected node in geometrical distance, rp is the distance from the center of rotation to the surface of the prism, iR is the depth of the prism surface in pixels, nw is the refractive index of water, and $sradial$ is the radial step size corresponding to each pixel (in our calculations, $rp=0.15$ mm, $nw=1.33$, and $sradial=0.008$ mm). Once the depth value for each contour node was calculated, the average lumen diameter was estimated. A great advantage of the detection algorithm was that it could be performed in real-time. This was a key element that allowed this algorithm to be used in a feedback loop to control the balloon inflation.

F. Control System Architecture

The goal of the control system was to provide convergence to a desired balloon diameter. A commercial syringe pump (PHD 4400, Harvard

Apparatus) was used, as an actuator to inflate the balloon. The syringe pump was composed of a microcontroller and a stepper motor. It could deliver a liquid with a customized flow rate. In the feedback loop, a PC performed real-time IVOCT image analysis to estimate the diameter. If the diameter was smaller than the desired diameter, the PC sent a command to the pump to deliver liquid. If the diameter was larger than the desired diameter, the PC sent a command to the pump to withdraw liquid. If the diameter was in an acceptable proximity of the desired diameter the PC sent a command to stop the pump. In our experiments, inflation and deflation were performed at a constant flow rate value which was determined at the beginning of each experiment. The delivered volume was estimated by integrating the flow rate. A pressure transducer (MLH150PSB01A, Honeywell) was connected to the tubing to monitor and record inflation pressures. More details on the control system have been reported earlier [13]. In angioplasty procedures, first a reference luminal diameter is estimated for the artery at the location of the stenosis. Then an appropriately-sized balloon or stent is deployed to achieve this diameter. In our experiments, the capability of the control system to provide various target diameters was investigated. As was mentioned previously, a compliant balloon was used in our experiments. Unlike a non-compliant balloon that is designed to achieve a fixed nominal diameter, the compliant balloon did not have a nominal diameter and could be inflated to various diameters, continuously, until it reached its burst pressure. This property gave us more flexibility in testing the control system for various values of the target diameter. The speed of the system was tested, using different inflation rates. The inflation rate has been suggested to be an important factor in some outcomes of angioplasty, e.g. restenosis [16, 17]. Therefore, the tests using various flow rates addressed an important aspect of the controlled inflation system.

We consider the following anycast field equations defined over an open bounded piece of network and/or feature space $\Omega \subset R^d$. They describe the dynamics of the mean anycast of each of p node populations.

$$\begin{cases} \left(\frac{d}{dt} + l_i \right) V_i(t, r) = \sum_{j=1}^p \int_{\Omega} J_{ij}(r, \bar{r}) S[(V_j(t - \tau_{ij}(r, \bar{r}), \bar{r}) - h_{ij})] d\bar{r} \\ \quad + I_i^{ext}(r, t), \quad t \geq 0, 1 \leq i \leq p, \\ V_i(t, r) = \phi_i(t, r) \quad t \in [-T, 0] \end{cases} \quad (1)$$

We give an interpretation of the various parameters and functions that appear in (1), Ω is finite piece of nodes and/or feature space and is represented as an open bounded set of R^d . The vector r and \bar{r} represent points in Ω . The

function $S : R \rightarrow (0,1)$ is the normalized sigmoid function:

$$S(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

It describes the relation between the input rate v_i of population i as a function of the packets potential, for example, $V_i = v_i = S[\sigma_i(V_i - h_i)]$. We note V the p -dimensional vector (V_1, \dots, V_p) . The p function $\phi_i, i = 1, \dots, p$, represent the initial conditions, see below. We note ϕ the p -dimensional vector (ϕ_1, \dots, ϕ_p) . The p function $I_i^{ext}, i = 1, \dots, p$, represent external factors from other network areas. We note I^{ext} the p -dimensional vector $(I_1^{ext}, \dots, I_p^{ext})$. The $p \times p$ matrix of functions $J = \{J_{ij}\}_{i,j=1,\dots,p}$ represents the connectivity between populations i and j , see below. The p real values $h_i, i = 1, \dots, p$, determine the threshold of activity for each population, that is, the value of the nodes potential corresponding to 50% of the maximal activity. The p real positive values $\sigma_i, i = 1, \dots, p$, determine the slopes of the sigmoids at the origin. Finally the p real positive values $l_i, i = 1, \dots, p$, determine the speed at which each anycast node potential decreases exponentially toward its real value. We also introduce the function $S : R^p \rightarrow R^p$, defined by $S(x) = [S(\sigma_1(x_1 - h_1)), \dots, S(\sigma_p(x_p - h_p))]$, and the diagonal $p \times p$ matrix $L_0 = \text{diag}(l_1, \dots, l_p)$. Is the intrinsic dynamics of the population given by the linear response of data transfer. $(\frac{d}{dt} + l_i)$ is replaced by $(\frac{d}{dt} + l_i)^2$ to use the alpha function response. We use $(\frac{d}{dt} + l_i)$ for simplicity although our analysis applies to more general intrinsic dynamics. For the sake, of generality, the propagation delays are not assumed to be identical for all populations, hence they are described by a matrix $\tau(r, \bar{r})$ whose element $\tau_{ij}(r, \bar{r})$ is the propagation delay between population j at \bar{r} and population i at r . The reason for this assumption is that it is still unclear from anycast if propagation delays are independent of the populations. We assume for technical reasons

that τ is continuous, that is $\tau \in C^0(\bar{\Omega}^2, R_+^{p \times p})$. Moreover packet data indicate that τ is not a symmetric function i.e., $\tau_{ij}(r, \bar{r}) \neq \tau_{ji}(\bar{r}, r)$, thus no assumption is made about this symmetry unless otherwise stated. In order to compute the righthand side of (1), we need to know the node potential factor V on interval $[-T, 0]$. The value of T is obtained by considering the maximal delay:

$$\tau_m = \max_{i,j(r,r \in \Omega \times \Omega)} \tau_{i,j}(r, \bar{r}) \quad (3)$$

Hence we choose $T = \tau_m$

G. Mathematical Framework

A convenient functional setting for the non-delayed packet field equations is to use the space $F = L^2(\Omega, R^p)$ which is a Hilbert space endowed with the usual inner product:

$$\langle V, U \rangle_F = \sum_{i=1}^p \int_{\Omega} V_i(r) U_i(r) dr \quad (1)$$

To give a meaning to (1), we defined the history space $C = C^0([- \tau_m, 0], F)$ with

$\|\phi\| = \sup_{t \in [- \tau_m, 0]} \|\phi(t)\|_F$, which is the Banach phase space associated with equation (3). Using the notation $V_t(\theta) = V(t + \theta), \theta \in [- \tau_m, 0]$, we write (1) as

$$\begin{cases} V(t) = -L_0 V(t) + L_1 S(V_t) + I^{ext}(t), \\ V_0 = \phi \in C, \end{cases} \quad (2)$$

Where

$$\begin{cases} L_1 : C \rightarrow F, \\ \phi \rightarrow \int_{\Omega} J(\cdot, \bar{r}) \phi(\bar{r}, -\tau(\cdot, \bar{r})) d\bar{r} \end{cases}$$

Is the linear continuous operator satisfying $\|L_1\| \leq \|J\|_{L^2(\Omega^2, R^{p \times p})}$. Notice that most of the papers on this subject assume Ω infinite, hence requiring $\tau_m = \infty$.

Proposition 1.0 If the following assumptions are satisfied.

1. $J \in L^2(\Omega^2, R^{p \times p})$,
2. The external current $I^{ext} \in C^0(R, F)$,
3. $\tau \in C^0(\bar{\Omega}^2, R_+^{p \times p}), \sup_{\bar{\Omega}^2} \tau \leq \tau_m$.

Then for any $\phi \in C$, there exists a unique solution $V \in C^1([0, \infty), F) \cap C^0([- \tau_m, \infty), F)$ to (3)

Notice that this result gives existence on R_+ , finite-time explosion is impossible for this delayed differential equation. Nevertheless, a particular solution could grow indefinitely, we now prove that this cannot happen.

H. Boundedness of Solutions

A valid model of neural networks should only feature bounded packet node potentials.

Theorem 1.0 All the trajectories are ultimately bounded by the same constant R if

$$I \equiv \max_{t \in R^+} \|I^{ext}(t)\|_F < \infty.$$

Proof : Let us defined $f : R \times C \rightarrow R^+$ as

$$f(t, V_t) \stackrel{def}{=} \left\langle -L_0 V_t(0) + L_1 S(V_t) + I^{ext}(t), V(t) \right\rangle_F = \frac{1}{2} \frac{d \|V\|_F^2}{dt}$$

We note $l = \min_{i=1, \dots, p} l_i$

$$f(t, V_t) \leq -l \|V(t)\|_F^2 + (\sqrt{p|\Omega|} \|J\|_F + I) \|V(t)\|_F$$

Thus, if

$$\|V(t)\|_F \geq 2 \frac{\sqrt{p|\Omega|} \|J\|_F + I \stackrel{def}{=} R, f(t, V_t) \leq -\frac{lR^2}{2} \stackrel{def}{=} -\delta < 0$$

Let us show that the open route of F of center 0 and radius R, B_R , is stable under the dynamics of equation. We know that $V(t)$ is defined for all $t \geq 0$ and that $f < 0$ on ∂B_R , the boundary of B_R . We consider three cases for the initial condition V_0 . If $\|V_0\|_C < R$ and set $T = \sup\{t \mid \forall s \in [0, t], V(s) \in \overline{B_R}\}$. Suppose that $T \in R$, then $V(T)$ is defined and belongs to $\overline{B_R}$, the closure of B_R , because $\overline{B_R}$ is closed, in effect to ∂B_R , we also have

$$\frac{d}{dt} \|V\|_F^2 \Big|_{t=T} = f(T, V_T) \leq -\delta < 0 \quad \text{because}$$

$V(T) \in \partial B_R$. Thus we deduce that for $\varepsilon > 0$ and small enough, $V(T + \varepsilon) \in \overline{B_R}$ which contradicts the definition of T . Thus $T \notin R$ and $\overline{B_R}$ is stable.

Because $f < 0$ on $\partial B_R, V(0) \in \partial B_R$ implies that $\forall t > 0, V(t) \in B_R$. Finally we consider the case $V(0) \in \overline{CB_R}$. Suppose that

$\forall t > 0, V(t) \notin \overline{B_R}$, then

$$\forall t > 0, \frac{d}{dt} \|V\|_F^2 \leq -2\delta, \quad \text{thus } \|V(t)\|_F \text{ is}$$

monotonically decreasing and reaches the value of R in finite time when $V(t)$ reaches ∂B_R . This contradicts our assumption. Thus

$$\exists T > 0 \mid V(T) \in B_R.$$

Proposition 1.1 : Let s and t be measured simple functions on X . for $E \in \mathcal{M}$, define

$$\phi(E) = \int_E s d\mu \quad (1)$$

Then ϕ is a measure on M .

$$\int_X (s+t) d\mu = \int_X s d\mu + \int_X t d\mu \quad (2)$$

Proof : If s and if E_1, E_2, \dots are disjoint members of M whose union is E , the countable additivity of μ shows that

$$\begin{aligned} \phi(E) &= \sum_{i=1}^n \alpha_i \mu(A_i \cap E) = \sum_{i=1}^n \alpha_i \sum_{r=1}^{\infty} \mu(A_i \cap E_r) \\ &= \sum_{r=1}^{\infty} \sum_{i=1}^n \alpha_i \mu(A_i \cap E_r) = \sum_{r=1}^{\infty} \phi(E_r) \end{aligned}$$

Also, $\phi(\emptyset) = 0$, so that ϕ is not identically ∞ .

Next, let s be as before, let β_1, \dots, β_m be the distinct values of t , and let $B_j = \{x : t(x) = \beta_j\}$ If $E_{ij} = A_i \cap B_j$, then

$$\int_{E_{ij}} (s+t) d\mu = (\alpha_i + \beta_j) \mu(E_{ij})$$

$$\text{and } \int_{E_{ij}} s d\mu + \int_{E_{ij}} t d\mu = \alpha_i \mu(E_{ij}) + \beta_j \mu(E_{ij})$$

Thus (2) holds with E_{ij} in place of X . Since X is the disjoint union of the sets E_{ij} ($1 \leq i \leq n, 1 \leq j \leq m$), the first half of our proposition implies that (2) holds.

Theorem 1.1: If K is a compact set in the plane whose complement is connected, if f is a continuous complex function on K which is holomorphic in the interior of K , and if $\varepsilon > 0$, then there exists a polynomial P such that $|f(z) - P(z)| < \varepsilon$ for all $z \in K$. If the interior of K is empty, then part of the hypothesis is vacuously satisfied, and the conclusion holds for every $f \in C(K)$. Note that K need to be connected.

Proof: By Tietze's theorem, f can be extended to a continuous function in the plane, with compact support. We fix one such extension and denote it again by f . For any $\delta > 0$, let $\omega(\delta)$ be the supremum of the numbers $|f(z_2) - f(z_1)|$ Where z_1 and z_2 are subject to the condition $|z_2 - z_1| \leq \delta$. Since f is uniformly continuous, we have $\lim_{\delta \rightarrow 0} \omega(\delta) = 0$ (1) From now on,

δ will be fixed. We shall prove that there is a polynomial P such that

$$|f(z) - P(z)| < 10,000 \omega(\delta) \quad (z \in K) \quad (2)$$

By (1), this proves the theorem. Our first objective is the construction of a function $\Phi \in C_c'(R^2)$, such that for all z

$$|f(z) - \Phi(z)| \leq \omega(\delta), \quad (3)$$

$$|(\partial\Phi)(z)| < \frac{2\omega(\delta)}{\delta}, \quad (4)$$

And

$$\Phi(z) = -\frac{1}{\pi} \iint_X \frac{(\partial\Phi)(\zeta)}{\zeta - z} d\zeta d\eta \quad (\zeta = \xi + i\eta), \quad (5)$$

Where X is the set of all points in the support of Φ whose distance from the complement of K does not exceed δ . (Thus X contains no point which is "far within" K .) We construct Φ as the convolution of f with a smoothing function A . Put $a(r) = 0$ if $r > \delta$, put

$$a(r) = \frac{3}{\pi\delta^2} (1 - \frac{r^2}{\delta^2})^2 \quad (0 \leq r \leq \delta), \quad (6)$$

And define

$$A(z) = a(|z|) \quad (7)$$

For all complex z . It is clear that $A \in C_c'(R^2)$. We claim that

$$\iint_{R^2} A = 1, \quad (8)$$

$$\iint_{R^2} \partial A = 0, \quad (9)$$

$$\iint_{R^2} |\partial A| = \frac{24}{15\delta} < \frac{2}{\delta}, \quad (10)$$

The constants are so adjusted in (6) that (8) holds. (Compute the integral in polar coordinates), (9) holds simply because A has compact support. To

compute (10), express ∂A in polar coordinates, and

$$\text{note that } \frac{\partial A}{\partial \theta} = 0,$$

$$\frac{\partial A}{\partial r} = -a',$$

Now define

$$\Phi(z) = \iint_{R^2} f(z - \zeta) A d\xi d\eta = \iint_{R^2} A(z - \zeta) f(\zeta) d\xi d\eta \quad (11)$$

Since f and A have compact support, so does Φ . Since

$$\begin{aligned} & \Phi(z) - f(z) \\ &= \iint_{R^2} [f(z - \zeta) - f(z)] A(\zeta) d\xi d\eta \quad (12) \end{aligned}$$

And $A(\zeta) = 0$ if $|\zeta| > \delta$, (3) follows from (8).

The difference quotients of A converge boundedly to the corresponding partial derivatives, since $A \in C_c'(R^2)$. Hence the last expression in (11) may be differentiated under the integral sign, and we obtain

$$\begin{aligned} (\partial\Phi)(z) &= \iint_{R^2} (\partial A)(z - \zeta) f(\zeta) d\xi d\eta \\ &= \iint_{R^2} f(z - \zeta) (\partial A)(\zeta) d\xi d\eta \\ &= \iint_{R^2} [f(z - \zeta) - f(z)] (\partial A)(\zeta) d\xi d\eta \quad (13) \end{aligned}$$

The last equality depends on (9). Now (10) and (13) give (4). If we write (13) with Φ_x and Φ_y in place of $\partial\Phi$, we see that Φ has continuous partial derivatives, if we can show that $\partial\Phi = 0$ in G , where G is the set of all $z \in K$ whose distance from the complement of K exceeds δ . We shall do this by showing that

$$\Phi(z) = f(z) \quad (z \in G); \quad (14)$$

Note that $\partial f = 0$ in G , since f is holomorphic there. Now if $z \in G$, then $z - \zeta$ is in the interior of K for all ζ with $|\zeta| < \delta$. The mean value property for harmonic functions therefore gives, by the first equation in (11),

$$\begin{aligned} \Phi(z) &= \int_0^\delta a(r) r dr \int_0^{2\pi} f(z - re^{i\theta}) d\theta \\ &= 2\pi f(z) \int_0^\delta a(r) r dr = f(z) \iint_{R^2} A = f(z) \quad (15) \end{aligned}$$

For all $z \in G$, we have now proved (3), (4), and (5) The definition of X shows that X is compact and that X can be covered by finitely many open

discs D_1, \dots, D_n , of radius 2δ , whose centers are not in K . Since $S^2 - K$ is connected, the center of each D_j can be joined to ∞ by a polygonal path in $S^2 - K$. It follows that each D_j contains a compact connected set E_j , of diameter at least 2δ , so that $S^2 - E_j$ is connected and so that $K \cap E_j = \emptyset$. with $r = 2\delta$. There are functions $g_j \in H(S^2 - E_j)$ and constants b_j so that the inequalities.

$$|Q_j(\zeta, z)| < \frac{50}{\delta}, \quad (16)$$

$$\left| Q_j(\zeta, z) - \frac{1}{z - \zeta} \right| < \frac{4,000\delta^2}{|z - \zeta|^2} \quad (17)$$

Hold for $z \notin E_j$ and $\zeta \in D_j$, if

$$Q_j(\zeta, z) = g_j(z) + (\zeta - b_j)g_j^2(z) \quad (18)$$

Let Ω be the complement of $E_1 \cup \dots \cup E_n$. Then

Ω is an open set which contains K . Put $X_1 = X \cap D_1$ and

$X_j = (X \cap D_j) - (X_1 \cup \dots \cup X_{j-1})$, for

$2 \leq j \leq n$,

Define

$$R(\zeta, z) = Q_j(\zeta, z) \quad (\zeta \in X_j, z \in \Omega) \quad (19)$$

And

$$F(z) = \frac{1}{\pi} \iint_X (\partial\Phi)(\zeta) R(\zeta, z) d\zeta d\eta \quad (20)$$

$(z \in \Omega)$

Since,

$$F(z) = \sum_{j=1}^n \frac{1}{\pi} \iint_{X_j} (\partial\Phi)(\zeta) Q_j(\zeta, z) d\zeta d\eta, \quad (21)$$

(18) shows that F is a finite linear combination of the functions g_j and g_j^2 . Hence $F \in H(\Omega)$. By (20), (4), and (5) we have

$$|F(z) - \Phi(z)| < \frac{2\omega(\delta)}{\pi\delta} \iint_X |R(\zeta, z)| d\zeta d\eta \quad (z \in \Omega) \quad (22)$$

$$- \frac{1}{z - \zeta} |d\zeta d\eta| \quad (z \in \Omega) \quad (22)$$

Observe that the inequalities (16) and (17) are valid with R in place of Q_j if $\zeta \in X$ and $z \in \Omega$.

Now fix $z \in \Omega$, put $\zeta = z + \rho e^{i\theta}$, and estimate the integrand in (22) by (16) if $\rho < 4\delta$, by (17) if $4\delta \leq \rho$. The integral in (22) is then seen to be less than the sum of

$$2\pi \int_0^{4\delta} \left(\frac{50}{\delta} + \frac{1}{\rho} \right) \rho d\rho = 808\pi\delta \quad (23)$$

And

$$2\pi \int_{4\delta}^{\infty} \frac{4,000\delta^2}{\rho^2} \rho d\rho = 2,000\pi\delta. \quad (24)$$

Hence (22) yields

$$|F(z) - \Phi(z)| < 6,000\omega(\delta) \quad (z \in \Omega) \quad (25)$$

Since $F \in H(\Omega)$, $K \subset \Omega$, and $S^2 - K$ is connected, Runge's theorem shows that F can be uniformly approximated on K by polynomials. Hence (3) and (25) show that (2) can be satisfied. This completes the proof.

Lemma 1.0 : Suppose $f \in C_c'(R^2)$, the space of all continuously differentiable functions in the plane, with compact support. Put

$$\partial = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \quad (1)$$

Then the following "Cauchy formula" holds:

$$f(z) = -\frac{1}{\pi} \iint_{R^2} \frac{(\partial f)(\zeta)}{\zeta - z} d\xi d\eta$$

$$(\zeta = \xi + i\eta) \quad (2)$$

Proof: This may be deduced from Green's theorem. However, here is a simple direct proof:

Put $\varphi(r, \theta) = f(z + re^{i\theta})$, $r > 0$, θ real

If $\zeta = z + re^{i\theta}$, the chain rule gives

$$(\partial f)(\zeta) = \frac{1}{2} e^{i\theta} \left[\frac{\partial}{\partial r} + \frac{i}{r} \frac{\partial}{\partial \theta} \right] \varphi(r, \theta) \quad (3)$$

The right side of (2) is therefore equal to the limit, as $\varepsilon \rightarrow 0$, of

$$-\frac{1}{2} \int_{\varepsilon}^{\infty} \int_0^{2\pi} \left(\frac{\partial \varphi}{\partial r} + \frac{i}{r} \frac{\partial \varphi}{\partial \theta} \right) d\theta dr \quad (4)$$

For each $r > 0$, φ is periodic in θ , with period 2π . The integral of $\partial\varphi / \partial\theta$ is therefore 0, and (4) becomes

$$-\frac{1}{2\pi} \int_0^{2\pi} d\theta \int_{\varepsilon}^{\infty} \frac{\partial \varphi}{\partial r} dr = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\varepsilon, \theta) d\theta \quad (5)$$

As $\varepsilon \rightarrow 0$, $\varphi(\varepsilon, \theta) \rightarrow f(z)$ uniformly. This gives (2)

If $X^\alpha \in a$ and $X^\beta \in k[X_1, \dots, X_n]$, then $X^\alpha X^\beta = X^{\alpha+\beta} \in a$, and so A satisfies the condition (*). Conversely,

$$\left(\sum_{\alpha \in A} c_\alpha X^\alpha\right) \left(\sum_{\beta \in \square^n} d_\beta X^\beta\right) = \sum_{\alpha, \beta} c_\alpha d_\beta X^{\alpha+\beta} \quad (\text{finite sums}),$$

and so if A satisfies (*), then the subspace generated by the monomials $X^\alpha, \alpha \in a$, is an ideal. The proposition gives a classification of the monomial ideals in $k[X_1, \dots, X_n]$: they are in one to one correspondence with the subsets A of \square^n satisfying (*). For example, the monomial ideals in $k[X]$ are exactly the ideals $(X^n), n \geq 1$, and the zero ideal (corresponding to the empty set A). We write $\langle X^\alpha \mid \alpha \in A \rangle$ for the ideal corresponding to A (subspace generated by the $X^\alpha, \alpha \in a$).

LEMMA 1.1. Let S be a subset of \square^n . The ideal a generated by $X^\alpha, \alpha \in S$ is the monomial ideal corresponding to

$$A \stackrel{df}{=} \left\{ \beta \in \square^n \mid \beta - \alpha \in \square^n, \text{ some } \alpha \in S \right\}$$

Thus, a monomial is in a if and only if it is divisible by one of the $X^\alpha, \alpha \in S$

PROOF. Clearly A satisfies (*), and $a \subset \langle X^\beta \mid \beta \in A \rangle$. Conversely, if $\beta \in A$, then $\beta - \alpha \in \square^n$ for some $\alpha \in S$, and $X^\beta = X^\alpha X^{\beta-\alpha} \in a$. The last statement follows from the fact that $X^\alpha \mid X^\beta \Leftrightarrow \beta - \alpha \in \square^n$. Let $A \subset \square^n$ satisfy (*). From the geometry of A , it is clear that there is a finite set of elements $S = \{\alpha_1, \dots, \alpha_s\}$ of A such that $A = \left\{ \beta \in \square^n \mid \beta - \alpha_i \in \square^2, \text{ some } \alpha_i \in S \right\}$

(The α_i 's are the corners of A) Moreover, $a \stackrel{df}{=} \langle X^\alpha \mid \alpha \in A \rangle$ is generated by the monomials $X^{\alpha_i}, \alpha_i \in S$.

DEFINITION 1.0. For a nonzero ideal a in $k[X_1, \dots, X_n]$, we let $(LT(a))$ be the ideal generated by

$$\{LT(f) \mid f \in a\}$$

LEMMA 1.2 Let a be a nonzero ideal in $k[X_1, \dots, X_n]$; then $(LT(a))$ is a monomial ideal, and it equals $(LT(g_1), \dots, LT(g_n))$ for some $g_1, \dots, g_n \in a$.

PROOF. Since $(LT(a))$ can also be described as the ideal generated by the leading monomials (rather than the leading terms) of elements of a .

THEOREM 1.2. Every ideal a in $k[X_1, \dots, X_n]$ is finitely generated; more precisely, $a = (g_1, \dots, g_s)$ where g_1, \dots, g_s are any elements of a whose leading terms generate $LT(a)$

PROOF. Let $f \in a$. On applying the division algorithm, we find $f = a_1 g_1 + \dots + a_s g_s + r$, $a_i, r \in k[X_1, \dots, X_n]$, where either $r = 0$ or no monomial occurring in it is divisible by any $LT(g_i)$. But $r = f - \sum a_i g_i \in a$, and therefore $LT(r) \in LT(a) = (LT(g_1), \dots, LT(g_s))$, implies that every monomial occurring in r is divisible by one in $LT(g_i)$. Thus $r = 0$, and $g \in (g_1, \dots, g_s)$.

DEFINITION 1.1. A finite subset $S = \{g_1, \dots, g_s\}$ of an ideal a is a standard (Gröbner) bases for a if $(LT(g_1), \dots, LT(g_s)) = LT(a)$. In other words, S is a standard basis if the leading term of every element of a is divisible by at least one of the leading terms of the g_i .

THEOREM 1.3 The ring $k[X_1, \dots, X_n]$ is Noetherian i.e., every ideal is finitely generated.

PROOF. For $n = 1$, $k[X]$ is a principal ideal domain, which means that every ideal is generated by single element. We shall prove the theorem by induction on n . Note that the obvious map

$k[X_1, \dots, X_{n-1}][X_n] \rightarrow k[X_1, \dots, X_n]$ is an isomorphism – this simply says that every polynomial f in n variables X_1, \dots, X_n can be expressed uniquely as a polynomial in X_n with coefficients in $k[X_1, \dots, X_{n-1}]$:

$$f(X_1, \dots, X_n) = a_0(X_1, \dots, X_{n-1})X_n^r + \dots + a_r(X_1, \dots, X_{n-1})$$

Thus the next lemma will complete the proof

LEMMA 1.3. If A is Noetherian, then so also is $A[X]$

PROOF. For a polynomial

$$f(X) = a_0X^r + a_1X^{r-1} + \dots + a_r, \quad a_i \in A, \quad a_0 \neq 0,$$

r is called the degree of f , and a_0 is its leading coefficient. We call 0 the leading coefficient of the polynomial 0 . Let a be an ideal in $A[X]$. The leading coefficients of the polynomials in a form an ideal a' in A , and since A is Noetherian, a' will be finitely generated. Let g_1, \dots, g_m be elements of a whose leading coefficients generate a' , and let r be the maximum degree of g_i . Now let $f \in a$, and suppose f has degree $s > r$, say,

$$f = aX^s + \dots$$

$$a = \sum b_i a_i, \quad b_i \in A,$$

$$a_i = \text{leading coefficient of } g_i$$

Now

$$f - \sum b_i g_i X^{s-r_i}, \quad r_i = \text{deg}(g_i), \text{ has degree } < \text{deg}(f).$$

By continuing in this way, we find that $f \equiv f_t \pmod{(g_1, \dots, g_m)}$ With f_t a polynomial of degree $t < r$. For each $d < r$, let

a_d be the subset of A consisting of 0 and the leading coefficients of all polynomials in a of degree d ; it is again an ideal in A . Let

$g_{d,1}, \dots, g_{d,m_d}$ be polynomials of degree d whose leading coefficients generate a_d . Then the same argument as above shows that any polynomial f_d in

a of degree d can be written $f_d \equiv f_{d-1} \pmod{(g_{d,1}, \dots, g_{d,m_d})}$ With f_{d-1}

of degree $\leq d-1$. On applying this remark repeatedly we find that

$$f_t \in (g_{r-1,1}, \dots, g_{r-1,m_{r-1}}, \dots, g_{0,1}, \dots, g_{0,m_0}) \text{ Hence}$$

$$f_t \in (g_1, \dots, g_m, g_{r-1,1}, \dots, g_{r-1,m_{r-1}}, \dots, g_{0,1}, \dots, g_{0,m_0})$$

and so the polynomials g_1, \dots, g_{0,m_0} generate a

One of the great successes of category theory in computer science has been the development of a “unified theory” of the constructions underlying denotational semantics. In the untyped λ -calculus, any term may appear in the function position of an application. This means that a model D of the λ -calculus must have the property that given a term t whose interpretation is $d \in D$, Also, the interpretation of a functional abstraction like $\lambda x . x$ is most conveniently defined as a function from D to D , which must then be regarded as an element of D . Let $\psi : [D \rightarrow D] \rightarrow D$ be the function that picks out elements of D to represent elements of $[D \rightarrow D]$ and $\phi : D \rightarrow [D \rightarrow D]$ be the function that maps elements of D to functions of D . Since $\psi(f)$ is intended to represent the function f as an element of D , it makes sense to require that $\phi(\psi(f)) = f$, that is, $\psi \circ \phi = id_{[D \rightarrow D]}$ Furthermore, we often want to view every element of D as representing some function from D to D and require that elements representing the same function be equal – that is

$$\psi(\phi(d)) = d$$

or

$$\psi \circ \phi = id_D$$

The latter condition is called extensionality. These conditions together imply that ϕ and ψ are inverses--- that is, D is isomorphic to the space of functions from D to D that can be the interpretations of functional abstractions:

$$D \cong [D \rightarrow D]$$

Let us suppose we are working with the untyped λ -calculus, we need a solution of the equation $D \cong A + [D \rightarrow D]$, where A is some predetermined domain containing interpretations for elements of C . Each element of D corresponds to either an element of A or an element of $[D \rightarrow D]$, with a tag. This equation can be solved by finding least fixed points of the function $F(X) = A + [X \rightarrow X]$ from domains to domains --- that is, finding domains X such that $X \cong A + [X \rightarrow X]$, and such that for any

domain Y also satisfying this equation, there is an embedding of X to Y --- a pair of maps

$$X \begin{array}{c} \xrightarrow{f} \\ \square \\ \xleftarrow{f^R} \end{array} Y$$

Such that

$$f^R \circ f = id_X$$

$$f \circ f^R \subseteq id_Y$$

Where $f \subseteq g$ means that

f approximates g in some ordering representing their information content. The key shift of perspective from the domain-theoretic to the more general category-theoretic approach lies in considering F not as a function on domains, but as a functor on a category of domains. Instead of a least fixed point of the function, F .

Definition 1.3: Let K be a category and $F : K \rightarrow K$ as a functor. A fixed point of F is a pair (A, a) , where A is a **K-object** and $a : F(A) \rightarrow A$ is an isomorphism. A prefixed point of F is a pair (A, a) , where A is a **K-object** and a is any arrow from $F(A)$ to A

Definition 1.4 : An ω -chain in a category K is a diagram of the following form:

$$\Delta = D_0 \xrightarrow{f_0} D_1 \xrightarrow{f_1} D_2 \xrightarrow{f_2} \dots$$

Recall that a cocone μ of an ω -chain Δ is a K -object X and a collection of K -arrows $\{\mu_i : D_i \rightarrow X \mid i \geq 0\}$ such that $\mu_i = \mu_{i+1} \circ f_i$

for all $i \geq 0$. We sometimes write $\mu : \Delta \rightarrow X$ as a reminder of the arrangement of μ 's components. Similarly, a colimit $\mu : \Delta \rightarrow X$ is a cocone with the property that if $\nu : \Delta \rightarrow X'$ is also a cocone then there exists a unique mediating arrow $k : X \rightarrow X'$ such that for all $i \geq 0$, $\nu_i = k \circ \mu_i$.

Colimits of ω -chains are sometimes referred to as ω -colimits. Dually, an ω^{op} -chain in K is a diagram of the following form:

$$\Delta = D_0 \xleftarrow{f_0} D_1 \xleftarrow{f_1} D_2 \xleftarrow{f_2} \dots$$

A cone $\mu : X \rightarrow \Delta$ of an ω^{op} -chain Δ is a K -object X and a collection of K -arrows $\{\mu_i : D_i \rightarrow X \mid i \geq 0\}$ such that for all $i \geq 0$, $\mu_i = f_i \circ \mu_{i+1}$. An ω^{op} -limit of an ω^{op} -chain Δ is a cone $\mu : X \rightarrow \Delta$ with the property that if $\nu : X' \rightarrow \Delta$ is also a cone, then there exists a unique mediating arrow $k : X' \rightarrow X$ such that for all $i \geq 0$, $\mu_i \circ k = \nu_i$.

We write \perp_k (or just \perp) for the distinguished initial object of K , when it has one, and $\perp \rightarrow A$ for the unique arrow from \perp to each K -object A . It is also

convenient to write $\Delta^- = D_1 \xrightarrow{f_1} D_2 \xrightarrow{f_2} \dots$ to denote all of Δ except D_0 and f_0 . By analogy,

μ^- is $\{\mu_i \mid i \geq 1\}$. For the images of Δ and μ

under F we write

$$F(\Delta) = F(D_0) \xrightarrow{F(f_0)} F(D_1) \xrightarrow{F(f_1)} F(D_2) \xrightarrow{F(f_2)} \dots$$

and $F(\mu) = \{F(\mu_i) \mid i \geq 0\}$

We write F^i for the i -fold iterated composition of F that is,

$$F^0(f) = f, F^1(f) = F(f), F^2(f) = F(F(f))$$

, etc. With these definitions we can state that every monotonic function on a complete lattice has a least fixed point:

Lemma 1.4. Let K be a category with initial object \perp and let $F : K \rightarrow K$ be a functor. Define the ω -chain Δ by

$$\Delta = \perp \xrightarrow{! \rightarrow F(\perp)} F(\perp) \xrightarrow{F(! \rightarrow F(\perp))} F^2(\perp) \xrightarrow{F^2(! \rightarrow F(\perp))} \dots$$

If both $\mu : \Delta \rightarrow D$ and $F(\mu) : F(\Delta) \rightarrow F(D)$ are colimits, then (D, d) is an initial F -algebra, where $d : F(D) \rightarrow D$ is the mediating arrow from $F(\mu)$ to the cocone μ^-

Theorem 1.4 Let a DAG G given in which each node is a random variable, and let a discrete conditional probability distribution of each node given values of its parents in G be specified. Then the product of these conditional distributions yields a joint probability distribution P of the variables, and (G, P) satisfies the Markov condition.

Proof. Order the nodes according to an ancestral ordering. Let X_1, X_2, \dots, X_n be the resultant ordering. Next define.

$$P(x_1, x_2, \dots, x_n) = P(x_n \mid pa_n) P(x_{n-1} \mid pa_{n-1}) \dots$$

$$\dots P(x_2 \mid pa_2) P(x_1 \mid pa_1),$$

Where PA_i is the set of parents of X_i of in G and $P(x_i \mid pa_i)$ is the specified conditional probability distribution. First we show this does indeed yield a joint probability distribution. Clearly, $0 \leq P(x_1, x_2, \dots, x_n) \leq 1$ for all values of the variables. Therefore, to show we have a joint distribution, as the variables range through all their

possible values, is equal to one. To that end, Specified conditional distributions are the conditional distributions they notationally represent in the joint distribution. Finally, we show the Markov condition is satisfied. To do this, we need show for $1 \leq k \leq n$ that whenever

$$P(pa_k) \neq 0, \text{ if } P(nd_k | pa_k) \neq 0$$

$$\text{and } P(x_k | pa_k) \neq 0$$

$$\text{then } P(x_k | nd_k, pa_k) = P(x_k | pa_k),$$

Where ND_k is the set of nondescendents of X_k of in G . Since $PA_k \subseteq ND_k$, we need only show $P(x_k | nd_k) = P(x_k | pa_k)$. First for a given k , order the nodes so that all and only nondescendents of X_k precede X_k in the ordering. Note that this ordering depends on k , whereas the ordering in the first part of the proof does not. Clearly then

$$ND_k = \{X_1, X_2, \dots, X_{k-1}\}$$

Let

$$D_k = \{X_{k+1}, X_{k+2}, \dots, X_n\}$$

follows \sum_{d_k}

We define the m^{th} cyclotomic field to be the field $Q[x]/(\Phi_m(x))$ Where $\Phi_m(x)$ is the m^{th} cyclotomic polynomial. $Q[x]/(\Phi_m(x))$ $\Phi_m(x)$ has degree $\varphi(m)$ over Q since $\Phi_m(x)$ has degree $\varphi(m)$. The roots of $\Phi_m(x)$ are just the primitive m^{th} roots of unity, so the complex embeddings of $Q[x]/(\Phi_m(x))$ are simply the $\varphi(m)$ maps

$$\sigma_k : Q[x]/(\Phi_m(x)) \mapsto C,$$

$$1 \leq k < m, (k, m) = 1, \text{ where}$$

$$\sigma_k(x) = \xi_m^k,$$

ξ_m being our fixed choice of primitive m^{th} root of unity. Note that $\xi_m^k \in Q(\xi_m)$ for every k ; it follows that $Q(\xi_m) = Q(\xi_m^k)$ for all k relatively prime to m . In particular, the images of the σ_i coincide, so $Q[x]/(\Phi_m(x))$ is Galois over Q . This means that we can write $Q(\xi_m)$ for $Q[x]/(\Phi_m(x))$ without much fear of ambiguity; we will do so from now on, the identification being

$\xi_m \mapsto x$. One advantage of this is that one can easily talk about cyclotomic fields being extensions of one another, or intersections or compositums; all of these things take place considering them as subfield of C . We now investigate some basic properties of cyclotomic fields. The first issue is whether or not they are all distinct; to determine this, we need to know which roots of unity lie in $Q(\xi_m)$. Note, for example, that if m is odd, then $-\xi_m$ is a $2m^{\text{th}}$ root of unity. We will show that this is the only way in which one can obtain any non- m^{th} roots of unity.

LEMMA 1.5 If m divides n , then $Q(\xi_m)$ is contained in $Q(\xi_n)$

PROOF. Since $\xi_n^{n/m} = \xi_m$, we have $\xi_m \in Q(\xi_n)$, so the result is clear

LEMMA 1.6 If m and n are relatively prime, then

$$Q(\xi_m, \xi_n) = Q(\xi_{mn})$$

and

$$Q(\xi_m) \cap Q(\xi_n) = Q$$

(Recall the $Q(\xi_m, \xi_n)$ is the compositum of $Q(\xi_m)$ and $Q(\xi_n)$)

PROOF. One checks easily that $\xi_m \xi_n$ is a primitive mn^{th} root of unity, so that

$$Q(\xi_{mn}) \subseteq Q(\xi_m, \xi_n)$$

$$[Q(\xi_m, \xi_n) : Q] \leq [Q(\xi_m) : Q][Q(\xi_n) : Q] = \varphi(m)\varphi(n) = \varphi(mn);$$

Since $[Q(\xi_{mn}) : Q] = \varphi(mn)$; this implies that

$Q(\xi_m, \xi_n) = Q(\xi_{mn})$ We know that $Q(\xi_m, \xi_n)$ has degree $\varphi(mn)$ over Q , so we must have

$$[Q(\xi_m, \xi_n) : Q(\xi_m)] = \varphi(n)$$

and

$$[Q(\xi_m, \xi_n) : Q(\xi_n)] = \varphi(m)$$

$$[Q(\xi_m) : Q(\xi_m) \cap Q(\xi_n)] \geq \varphi(m)$$

And thus that $Q(\xi_m) \cap Q(\xi_n) = Q$

PROPOSITION 1.2 For any m and n

$$Q(\xi_m, \xi_n) = Q(\xi_{[m,n]})$$

And

$$Q(\xi_m) \cap Q(\xi_n) = Q(\xi_{(m,n)});$$

here $[m, n]$ and (m, n) denote the least common multiple and the greatest common divisor of m and n , respectively.

PROOF. Write $m = p_1^{e_1} \dots p_k^{e_k}$ and $p_1^{f_1} \dots p_k^{f_k}$ where the p_i are distinct primes. (We allow e_i or f_i to be zero)

$$Q(\xi_m) = Q(\xi_{p_1^{e_1}}) Q(\xi_{p_2^{e_2}}) \dots Q(\xi_{p_k^{e_k}})$$

and

$$Q(\xi_n) = Q(\xi_{p_1^{f_1}}) Q(\xi_{p_2^{f_2}}) \dots Q(\xi_{p_k^{f_k}})$$

Thus

$$\begin{aligned} Q(\xi_m, \xi_n) &= Q(\xi_{p_1^{e_1}}) \dots Q(\xi_{p_2^{e_2}}) Q(\xi_{p_1^{f_1}}) \dots Q(\xi_{p_k^{f_k}}) \\ &= Q(\xi_{p_1^{e_1}}) Q(\xi_{p_1^{f_1}}) \dots Q(\xi_{p_k^{e_k}}) Q(\xi_{p_k^{f_k}}) \\ &= Q(\xi_{p_1^{\max(e_1, f_1)}}) \dots Q(\xi_{p_k^{\max(e_k, f_k)}}) \\ &= Q(\xi_{p_1^{\max(e_1, f_1)} \dots p_k^{\max(e_k, f_k)}}) \\ &= Q(\xi_{[m, n]}); \end{aligned}$$

An entirely similar computation shows that $Q(\xi_m) \cap Q(\xi_n) = Q(\xi_{(m,n)})$

Mutual information measures the information transferred when x_i is sent and y_i is received, and is defined as

$$I(x_i, y_i) = \log_2 \frac{P(x_i/y_i)}{P(x_i)} \text{ bits} \quad (1)$$

In a noise-free channel, each y_i is uniquely connected to the corresponding x_i , and so they constitute an input-output pair (x_i, y_i) for which

$$P(x_i/y_i) = 1 \text{ and } I(x_i, y_i) = \log_2 \frac{1}{P(x_i)} \text{ bits;}$$

that is, the transferred information is equal to the self-information that corresponds to the input x_i . In a very noisy channel, the output y_i and input x_i would be completely uncorrelated, and so $P(x_i/y_i) = P(x_i)$ and also $I(x_i, y_i) = 0$; that is,

there is no transference of information. In general, a given channel will operate between these two extremes. The mutual information is defined between the input and the output of a given channel. An average of the calculation of the mutual

information for all input-output pairs of a given channel is the average mutual information:

$$I(X, Y) = \sum_{i,j} P(x_i, y_j) I(x_i, y_j) = \sum_{i,j} P(x_i, y_j) \log_2 \left[\frac{P(x_i/y_j)}{P(x_i)} \right]$$

bits per symbol. This calculation is done over the input and output alphabets. The average mutual information. The following expressions are useful for modifying the mutual information expression:

$$P(x_i, y_j) = P(x_i/y_j) P(y_j) = P(y_j/x_i) P(x_i)$$

$$P(y_j) = \sum_i P(y_j/x_i) P(x_i)$$

$$P(x_i) = \sum_j P(x_i/y_j) P(y_j)$$

Then

$$\begin{aligned} I(X, Y) &= \sum_{i,j} P(x_i, y_j) \\ &= \sum_{i,j} P(x_i, y_j) \log_2 \left[\frac{1}{P(x_i)} \right] \\ &\quad - \sum_{i,j} P(x_i, y_j) \log_2 \left[\frac{1}{P(x_i/y_j)} \right] \\ &= \sum_{i,j} \left[P(x_i/y_j) P(y_j) \right] \log_2 \frac{1}{P(x_i)} \\ &= \sum_i P(x_i) \log_2 \frac{1}{P(x_i)} = H(X) \end{aligned}$$

$$I(X, Y) = H(X) - H(X/Y)$$

$$\begin{aligned} &= \sum_{i,j} P(x_i, y_j) \log_2 \left[\frac{1}{P(x_i)} \right] \\ &= \sum_i P(x_i) \log_2 \frac{1}{P(x_i)} = H(X) \end{aligned}$$

$$I(X, Y) = H(X) - H(X/Y)$$

$$I(X, Y) = H(X) - H(X/Y)$$

Where

$$H(X/Y) = \sum_{i,j} P(x_i, y_j) \log_2 \frac{1}{P(x_i/y_j)} \text{ is}$$

usually called the equivocation. In a sense, the equivocation can be seen as the information lost in the noisy channel, and is a function of the backward conditional probability. The observation of an output symbol y_j provides $H(X) - H(X/Y)$ bits of information. This difference is the mutual information of the channel. *Mutual Information: Properties* Since

$$P(x_i/y_j) P(y_j) = P(y_j/x_i) P(x_i)$$

The mutual information fits the condition

$$I(X, Y) = I(Y, X)$$

And by interchanging input and output it is also true that

$$I(X, Y) = H(Y) - H(Y/X)$$

Where

$$H(Y) = \sum_j P(y_j) \log_2 \frac{1}{P(y_j)}$$

This last entropy is usually called the noise entropy. Thus, the information transferred through the channel is the difference between the output entropy and the noise entropy. Alternatively, it can be said that the channel mutual information is the difference between the number of bits needed for determining a given input symbol before knowing the corresponding output symbol, and the number of bits needed for determining a given input symbol after knowing the corresponding output symbol

$$I(X, Y) = H(X) - H(X/Y)$$

As the channel mutual information expression is a difference between two quantities, it seems that this parameter can adopt negative values. However, and in spite of the fact that for some y_j , $H(X/y_j)$ can be larger than $H(X)$, this is not possible for the average value calculated over all the outputs:

$$\sum_{i,j} P(x_i, y_j) \log_2 \frac{P(x_i/y_j)}{P(x_i)} = \sum_{i,j} P(x_i, y_j) \log_2 \frac{P(x_i, y_j)}{P(x_i)P(y_j)}$$

Then

$$-I(X, Y) = \sum_{i,j} P(x_i, y_j) \frac{P(x_i)P(y_j)}{P(x_i, y_j)} \leq 0$$

Because this expression is of the form

$$\sum_{i=1}^M P_i \log_2 \left(\frac{Q_i}{P_i} \right) \leq 0$$

The above expression can be applied due to the factor $P(x_i)P(y_j)$, which is the product of two probabilities, so that it behaves as the quantity Q_i , which in this expression is a dummy variable that fits the condition $\sum_i Q_i \leq 1$. It can be concluded that the average mutual information is a non-negative number. It can also be equal to zero, when the input and the output are independent of each other. A related entropy called the joint entropy is defined as

$$\begin{aligned} H(X, Y) &= \sum_{i,j} P(x_i, y_j) \log_2 \frac{1}{P(x_i, y_j)} \\ &= \sum_{i,j} P(x_i, y_j) \log_2 \frac{P(x_i)P(y_j)}{P(x_i, y_j)} \\ &+ \sum_{i,j} P(x_i, y_j) \log_2 \frac{1}{P(x_i)P(y_j)} \end{aligned}$$

Theorem 1.5: Entropies of the binary erasure channel (BEC) The BEC is defined with an alphabet of two inputs and three outputs, with symbol probabilities.

$P(x_1) = \alpha$ and $P(x_2) = 1 - \alpha$, and transition probabilities

$$P(y_3/x_2) = 1 - p \text{ and } P(y_2/x_1) = 0,$$

$$\text{and } P(y_3/x_1) = 0$$

$$\text{and } P(y_1/x_2) = p$$

$$\text{and } P(y_2/x_2) = 1 - p$$

Lemma 1.7. Given an arbitrary restricted time-discrete, amplitude-continuous channel whose restrictions are determined by sets F_n and whose density functions exhibit no dependence on the state s , let n be a fixed positive integer, and $p(x)$ an arbitrary probability density function on Euclidean n -space. $p(y|x)$ for the density $p_n(y_1, \dots, y_n | x_1, \dots, x_n)$ and F for F_n . For any real number a , let

$$A = \left\{ (x, y) : \log \frac{p(y|x)}{p(y)} > a \right\} \quad (1)$$

Then for each positive integer u , there is a code (u, n, λ) such that

$$\lambda \leq ue^{-a} + P\{(X, Y) \notin A\} + P\{X \notin F\} \quad (2)$$

Where

$$P\{(X, Y) \in A\} = \int_A \dots \int p(x, y) dx dy, \quad p(x, y) = p(x)p(y|x)$$

and

$$P\{X \in F\} = \int_F \dots \int p(x) dx$$

Proof: A sequence $x^{(1)} \in F$ such that

$$P\{Y \in A_{x^{(1)}} | X = x^{(1)}\} \geq 1 - \varepsilon$$

where $A_x = \{y : (x, y) \in A\}$;

Choose the decoding set B_1 to be $A_{x^{(1)}}$. Having

chosen $x^{(1)}, \dots, x^{(k-1)}$ and B_1, \dots, B_{k-1} , select

$x^k \in F$ such that

$$P\left\{Y \in A_{x^{(k)}} - \bigcup_{i=1}^{k-1} B_i \mid X = x^{(k)}\right\} \geq 1 - \varepsilon;$$

Set $B_k = A_{x^{(k)}} - \bigcup_{i=1}^{k-1} B_i$. If the process does not terminate in a finite number of steps, then the sequences $x^{(i)}$ and decoding sets $B_i, i = 1, 2, \dots, u$, form the desired code. Thus assume that the process terminates after t steps. (Conceivably $t = 0$). We will show $t \geq u$ by showing that $\varepsilon \leq te^{-a} + P\{(X, Y) \notin A\} + P\{X \notin F\}$. We proceed as follows.

Let

$$B = \bigcup_{j=1}^t B_j. \quad (\text{If } t = 0, \text{ take } B = \phi). \text{ Then}$$

$$P\{(X, Y) \in A\} = \int_{(x, y) \in A} p(x, y) dx dy$$

$$= \int_x p(x) \int_{y \in A_x} p(y | x) dy dx$$

$$= \int_x p(x) \int_{y \in B \cap A_x} p(y | x) dy dx + \int_x p(x)$$

I. Algorithms

Ideals. Let A be a ring. Recall that an *ideal* a in A is a subset such that a is a subgroup of A regarded as a group under addition;

$$a \in a, r \in A \Rightarrow ra \in a$$

The *ideal generated by a subset* S of A is the intersection of all ideals A containing S ---- it is easy to verify that this is in fact an ideal, and that it consist of all finite sums of the form $\sum r_i s_i$ with $r_i \in A, s_i \in S$. When $S = \{s_1, \dots, s_m\}$, we shall write (s_1, \dots, s_m) for the ideal it generates.

Let a and b be ideals in A . The set $\{a+b \mid a \in a, b \in b\}$ is an ideal, denoted by $a+b$. The ideal generated by $\{ab \mid a \in a, b \in b\}$ is denoted by ab . Note that $ab \subset a \cap b$. Clearly ab consists of all finite sums $\sum a_i b_i$ with $a_i \in a$ and $b_i \in b$, and if $a = (a_1, \dots, a_m)$ and $b = (b_1, \dots, b_n)$, then $ab = (a_1 b_1, \dots, a_1 b_n, \dots, a_m b_1, \dots, a_m b_n)$. Let a be an ideal of A . The set of cosets of a in A forms a ring A/a , and $a \mapsto a+a$ is a homomorphism $\phi: A \mapsto A/a$. The map $b \mapsto \phi^{-1}(b)$ is a one to one correspondence between the ideals of A/a and the ideals of A containing a . An ideal p if

prime if $p \neq A$ and $ab \in p \Rightarrow a \in p$ or $b \in p$.

Thus p is prime if and only if A/p is nonzero and has the property that $ab = 0, b \neq 0 \Rightarrow a = 0$, i.e., A/p is an integral domain. An ideal m is *maximal* if $m \neq A$ and there does not exist an ideal n contained strictly between m and A . Thus m is maximal if and only if A/m has no proper nonzero ideals, and so is a field. Note that m maximal $\Rightarrow m$ prime.

The ideals of $A \times B$ are all of the form $a \times b$, with a and b ideals in A and B . To see this, note that if c is an ideal in $A \times B$ and $(a, b) \in c$, then $(a, 0) = (a, b)(1, 0) \in c$ and $(0, b) = (a, b)(0, 1) \in c$. This shows that $c = a \times b$ with

$$a = \{a \mid (a, b) \in c \text{ some } b \in b\}$$

and

$$b = \{b \mid (a, b) \in c \text{ some } a \in a\}$$

Let A be a ring. An A -algebra is a ring B together with a homomorphism $i_B: A \rightarrow B$. A *homomorphism of A -algebra* $B \rightarrow C$ is a homomorphism of rings $\phi: B \rightarrow C$ such that $\phi(i_B(a)) = i_C(a)$ for all $a \in A$. An A -algebra B is said to be *finitely generated* (or of *finite-type* over A) if there exist elements $x_1, \dots, x_n \in B$ such that every element of B can be expressed as a polynomial in the x_i with coefficients in $i(A)$, i.e., such that the homomorphism $A[X_1, \dots, X_n] \rightarrow B$ sending X_i to x_i is surjective. A ring homomorphism $A \rightarrow B$ is *finite*, and B is finitely generated as an A -module. Let k be a field, and let A be a k -algebra. If $1 \neq 0$ in A , then the map $k \rightarrow A$ is injective, we can identify k with its image, i.e., we can regard k as a subring of A . If $1=0$ in a ring R , the R is the zero ring, i.e., $R = \{0\}$.

Polynomial rings. Let k be a field. A *monomial* in X_1, \dots, X_n is an expression of the form $X_1^{a_1} \dots X_n^{a_n}$, $a_j \in N$. The *total degree* of the monomial is $\sum a_i$. We sometimes abbreviate it by X^α , $\alpha = (a_1, \dots, a_n) \in \mathbb{N}^n$. The elements of the polynomial ring $k[X_1, \dots, X_n]$ are finite sums $\sum c_{a_1, \dots, a_n} X_1^{a_1} \dots X_n^{a_n}$, $c_{a_1, \dots, a_n} \in k$, $a_j \in \mathbb{N}$.

With the obvious notions of equality, addition and multiplication. Thus the monomials form basis for $k[X_1, \dots, X_n]$ as a k -vector space. The ring $k[X_1, \dots, X_n]$ is an integral domain, and the only units in it are the nonzero constant polynomials. A polynomial $f(X_1, \dots, X_n)$ is *irreducible* if it is nonconstant and has only the obvious factorizations, i.e., $f = gh \Rightarrow g$ or h is constant. **Division in $k[X]$.** The division algorithm allows us to divide a nonzero polynomial into another: let f and g be polynomials in $k[X]$ with $g \neq 0$; then there exist unique polynomials $q, r \in k[X]$ such that $f = qg + r$ with either $r = 0$ or $\deg r < \deg g$. Moreover, there is an algorithm for deciding whether $f \in (g)$, namely, find r and check whether it is zero. Moreover, the Euclidean algorithm allows to pass from finite set of generators for an ideal in $k[X]$ to a single generator by successively replacing each pair of generators with their greatest common divisor.

(Pure) **lexicographic ordering (lex).** Here monomials are ordered by lexicographic (dictionary) order. More precisely, let $\alpha = (a_1, \dots, a_n)$ and $\beta = (b_1, \dots, b_n)$ be two elements of \square^n ; then $\alpha > \beta$ and $X^\alpha > X^\beta$ (lexicographic ordering) if, in the vector difference $\alpha - \beta \in \square^n$, the left most nonzero entry is positive. For example,

$XY^2 > Y^3Z^4$; $X^3Y^2Z^4 > X^3Y^2Z$. Note that this isn't quite how the dictionary would order them: it would put $XXXXYYYYZZ$ after $XXXXYYYY$. **Graded reverse lexicographic order (grevlex).** Here monomials are ordered by total degree, with ties broken by reverse lexicographic ordering. Thus, $\alpha > \beta$ if $\sum a_i > \sum b_i$, or $\sum a_i = \sum b_i$ and in $\alpha - \beta$ the right most nonzero entry is negative. For example:

$$X^4Y^4Z^7 > X^5Y^5Z^4 \quad (\text{total degree greater})$$

$$XY^5Z^2 > X^4YZ^3, \quad X^5YZ > X^4YZ^2$$

Orderings on $k[X_1, \dots, X_n]$. Fix an ordering on the monomials in $k[X_1, \dots, X_n]$. Then we can write an element f of $k[X_1, \dots, X_n]$ in a canonical fashion, by re-ordering its elements in decreasing order. For example, we would write

$$f = 4XY^2Z + 4Z^2 - 5X^3 + 7X^2Z^2$$

as

$$f = -5X^3 + 7X^2Z^2 + 4XY^2Z + 4Z^2 \quad (\text{lex})$$

or

$$f = 4XY^2Z + 7X^2Z^2 - 5X^3 + 4Z^2 \quad (\text{grevlex})$$

Let $\sum a_\alpha X^\alpha \in k[X_1, \dots, X_n]$, in decreasing order:

$$f = a_{\alpha_0} X^{\alpha_0} + a_{\alpha_1} X^{\alpha_1} + \dots, \quad \alpha_0 > \alpha_1 > \dots, \quad \alpha_0 \neq 0$$

Then we define.

- The *multidegree* of f to be $\text{multdeg}(f) = \alpha_0$;
- The *leading coefficient* of f to be $LC(f) = a_{\alpha_0}$;
- The *leading monomial* of f to be $LM(f) = X^{\alpha_0}$;
- The *leading term* of f to be $LT(f) = a_{\alpha_0} X^{\alpha_0}$

For the polynomial $f = 4XY^2Z + \dots$, the multidegree is (1,2,1), the leading coefficient is 4, the leading monomial is XY^2Z , and the leading term is $4XY^2Z$. **The division algorithm in $k[X_1, \dots, X_n]$.** Fix a monomial ordering in \square^n .

Suppose given a polynomial f and an ordered set (g_1, \dots, g_s) of polynomials; the division algorithm then constructs polynomials a_1, \dots, a_s and r such that $f = a_1g_1 + \dots + a_sg_s + r$ Where either $r = 0$ or no monomial in r is divisible by any of $LT(g_1), \dots, LT(g_s)$ **Step 1:** If $LT(g_1) | LT(f)$, divide g_1 into f to get

$$f = a_1g_1 + h, \quad a_1 = \frac{LT(f)}{LT(g_1)} \in k[X_1, \dots, X_n]$$

If $LT(g_1) | LT(h)$, repeat the process until $f = a_1g_1 + f_1$ (different a_1) with $LT(f_1)$ not divisible by $LT(g_1)$. Now divide g_2 into f_1 , and so on, until $f = a_1g_1 + \dots + a_sg_s + r_1$ With $LT(r_1)$ not divisible by any $LT(g_1), \dots, LT(g_s)$

Step 2: Rewrite $r_1 = LT(r_1) + r_2$, and repeat Step 1 with r_2 for f :

$f = a_1 g_1 + \dots + a_s g_s + LT(r_1) + r_3$ (different a_i 's) **Monomial ideals.** In general, an ideal a will contain a polynomial without containing the individual terms of the polynomial; for example, the ideal $a = (Y^2 - X^3)$ contains $Y^2 - X^3$ but not Y^2 or X^3 .

DEFINITION 1.5. An ideal a is *monomial* if $\sum c_\alpha X^\alpha \in a \Rightarrow X^\alpha \in a$ all α with $c_\alpha \neq 0$.

PROPOSITION 1.3. Let a be a *monomial ideal*, and let $A = \{\alpha \mid X^\alpha \in a\}$. Then A satisfies the condition $\alpha \in A, \beta \in \square^n \Rightarrow \alpha + \beta \in A$ (*) And a is the k -subspace of $k[X_1, \dots, X_n]$ generated by the $X^\alpha, \alpha \in A$. Conversely, if A is a subset of \square^n satisfying (*), then the k -subspace a of $k[X_1, \dots, X_n]$ generated by $\{X^\alpha \mid \alpha \in A\}$ is a monomial ideal.

PROOF. It is clear from its definition that a monomial ideal a is the k -subspace of $k[X_1, \dots, X_n]$ generated by the set of monomials it contains. If $X^\alpha \in a$ and $X^\beta \in k[X_1, \dots, X_n]$.

If a permutation is chosen uniformly and at random from the $n!$ possible permutations in S_n , then the counts $C_j^{(n)}$ of cycles of length j are dependent random variables. The joint distribution of $C^{(n)} = (C_1^{(n)}, \dots, C_n^{(n)})$ follows from Cauchy's formula, and is given by

$$P[C^{(n)} = c] = \frac{1}{n!} N(n, c) = \frac{1}{n!} \left\{ \sum_{j=1}^n j c_j = n \right\} \prod_{j=1}^n \frac{1}{j} \frac{1}{c_j!}, \quad (1.1)$$

for $c \in \square_+^n$.

Lemma 1.7 For nonnegative integers m_1, \dots, m_n ,

$$E \left(\prod_{j=1}^n (C_j^{(n)})^{m_j} \right) = \left(\prod_{j=1}^n \left(\frac{1}{j} \right)^{m_j} \right) \mathbb{1} \left\{ \sum_{j=1}^n j m_j \leq n \right\} \quad (1.4)$$

Proof. This can be established directly by exploiting cancellation of the form $c_j^{[m_j]} / c_j! = 1 / (c_j - m_j)!$ when $c_j \geq m_j$, which occurs between the ingredients in Cauchy's formula

and the falling factorials in the moments. Write $m = \sum j m_j$. Then, with the first sum indexed by $c = (c_1, \dots, c_n) \in \square_+^n$ and the last sum indexed by $d = (d_1, \dots, d_n) \in \square_+^n$ via the correspondence $d_j = c_j - m_j$, we have

$$\begin{aligned} E \left(\prod_{j=1}^n (C_j^{(n)})^{m_j} \right) &= \sum_c P[C^{(n)} = c] \prod_{j=1}^n (c_j)^{m_j} \\ &= \sum_{c: c_j \geq m_j \text{ for all } j} \mathbb{1} \left\{ \sum_{j=1}^n j c_j = n \right\} \prod_{j=1}^n \frac{(c_j)^{m_j}}{j^{c_j} c_j!} \\ &= \prod_{j=1}^n \frac{1}{j^{m_j}} \sum_d \mathbb{1} \left\{ \sum_{j=1}^n j d_j = n - m \right\} \prod_{j=1}^n \frac{1}{j^{d_j} (d_j)!} \end{aligned}$$

This last sum simplifies to the indicator $\mathbb{1}(m \leq n)$, corresponding to the fact that if $n - m \geq 0$, then $d_j = 0$ for $j > n - m$, and a random permutation in S_{n-m} must have some cycle structure (d_1, \dots, d_{n-m}) . The moments of $C_j^{(n)}$ follow immediately as

$$E(C_j^{(n)})^{[r]} = j^{-r} \mathbb{1}\{jr \leq n\} \quad (1.2)$$

We note for future reference that (1.4) can also be written in the form

$$E \left(\prod_{j=1}^n (C_j^{(n)})^{m_j} \right) = E \left(\prod_{j=1}^n Z_j^{m_j} \right) \mathbb{1} \left\{ \sum_{j=1}^n j m_j \leq n \right\}, \quad (1.3)$$

Where the Z_j are independent Poisson-distribution random variables that satisfy $E(Z_j) = 1/j$

The marginal distribution of cycle counts provides a formula for the joint distribution of the cycle counts C_j^n , we find the distribution of C_j^n using a combinatorial approach combined with the inclusion-exclusion formula.

Lemma 1.8. For $1 \leq j \leq n$,

$$P[C_j^{(n)} = k] = \frac{j^{-k}}{k!} \sum_{l=0}^{[n/j]-k} (-1)^l \frac{j^{-l}}{l!} \quad (1.1)$$

Proof. Consider the set I of all possible cycles of length j , formed with elements chosen from $\{1, 2, \dots, n\}$, so that $|I| = n^{[n/j]}$. For each $\alpha \in I$, consider the "property" G_α of having α ; that is, G_α is the set of permutations $\pi \in S_n$ such that α is one of the cycles of π . We then have $|G_\alpha| = (n-j)!$, since the elements of $\{1, 2, \dots, n\}$ not in α must be permuted among themselves. To

use the inclusion-exclusion formula we need to calculate the term S_r , which is the sum of the probabilities of the r -fold intersection of properties, summing over all sets of r distinct properties. There are two cases to consider. If the r properties are indexed by r cycles having no elements in common, then the intersection specifies how rj elements are moved by the permutation, and there are $(n-rj)!(rj \leq n)$ permutations in the intersection. There are $n^{\lfloor rj \rfloor} / (j^r r!)$ such intersections. For the other case, some two distinct properties name some element in common, so no permutation can have both these properties, and the r -fold intersection is empty. Thus

$$S_r = (n-rj)!(rj \leq n) \times \frac{n^{\lfloor rj \rfloor}}{j^r r! n!} = 1(rj \leq n) \frac{1}{j^r r!}$$

Finally, the inclusion-exclusion series for the number of permutations having exactly k properties is

$$\sum_{l \geq 0} (-1)^l \binom{k+l}{l} S_{k+l}$$

Which simplifies to (1.1) Returning to the original hat-check problem, we substitute $j=1$ in (1.1) to obtain the distribution of the number of fixed points of a random permutation. For $k = 0, 1, \dots, n$,

$$P[C_1^{(n)} = k] = \frac{1}{k!} \sum_{l=0}^{n-k} (-1)^l \frac{1}{l!}, \quad (1.2)$$

and the moments of $C_1^{(n)}$ follow from (1.2) with $j=1$. In particular, for $n \geq 2$, the mean and variance of $C_1^{(n)}$ are both equal to 1. The joint distribution of $(C_1^{(n)}, \dots, C_b^{(n)})$ for any $1 \leq b \leq n$ has an expression similar to (1.7); this too can be derived by inclusion-exclusion. For any $c = (c_1, \dots, c_b) \in \square_+^b$ with $m = \sum c_i$,

$$P[(C_1^{(n)}, \dots, C_b^{(n)}) = c] = \left\{ \prod_{i=1}^b \left(\frac{1}{i} \right)^{c_i} \frac{1}{c_i!} \right\} \sum_{\substack{l \geq 0 \text{ with} \\ \sum l_i \leq n-m}} (-1)^{l_1 + \dots + l_b} \prod_{i=1}^b \left(\frac{1}{i} \right)^{l_i} \frac{1}{l_i!}$$

The joint moments of the first b counts $C_1^{(n)}, \dots, C_b^{(n)}$ can be obtained directly from (1.2) and (1.3) by setting $m_{b+1} = \dots = m_n = 0$

The limit distribution of cycle counts

It follows immediately from Lemma 1.2 that for each fixed j , as $n \rightarrow \infty$,

$$P[C_j^{(n)} = k] \rightarrow \frac{j^{-k}}{k!} e^{-1/j}, \quad k = 0, 1, 2, \dots,$$

So that $C_j^{(n)}$ converges in distribution to a random variable Z_j having a Poisson distribution with mean $1/j$; we use the notation $C_j^{(n)} \rightarrow_d Z_j$ where $Z_j \square P_o(1/j)$ to describe this. Infact, the limit random variables are independent.

Theorem 1.6 The process of cycle counts converges in distribution to a Poisson process of \square with intensity j^{-1} . That is, as $n \rightarrow \infty$,

$$(C_1^{(n)}, C_2^{(n)}, \dots) \rightarrow_d (Z_1, Z_2, \dots) \quad (1.1)$$

Where the $Z_j, j = 1, 2, \dots$, are independent Poisson-distributed random variables with

$$E(Z_j) = \frac{1}{j}$$

Proof. To establish the converges in distribution one shows that for each fixed $b \geq 1$, as $n \rightarrow \infty$,

$$P[(C_1^{(n)}, \dots, C_b^{(n)}) = c] \rightarrow P[(Z_1, \dots, Z_b) = c]$$

Error rates

The proof of Theorem says nothing about the rate of convergence. Elementary analysis can be used to estimate this rate when $b=1$. Using properties of alternating series with decreasing terms, for $k = 0, 1, \dots, n$,

$$\frac{1}{k!} \left(\frac{1}{(n-k+1)!} - \frac{1}{(n-k+2)!} \right) \leq |P[C_1^{(n)} = k] - P[Z_1 = k]| \leq \frac{1}{k!(n-k+1)!}$$

It follows that

$$(1.3) \frac{2^{n+1}}{(n+1)! n+2} \leq \sum_{k=0}^n |P[C_1^{(n)} = k] - P[Z_1 = k]| \leq \frac{2^{n+1}-1}{(n+1)!} \quad (1.11)$$

Since

$$P[Z_1 > n] = \frac{e^{-1}}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \dots \right) < \frac{1}{(n+1)!}$$

We see from (1.11) that the total variation distance between the distribution $L(C_1^{(n)})$ of $C_1^{(n)}$ and the distribution $L(Z_1)$ of Z_1

Establish the asymptotics of $P[A_n(C^{(n)})]$ under conditions (A_0) and (B_{01}) , where

$$A_n(C^{(n)}) = \bigcap_{1 \leq i \leq n} \bigcap_{r_i+1 \leq j \leq r_i} \{C_{ij}^{(n)} = 0\},$$

and $\zeta_i = (r_i' / r_{id}) - 1 = O(i^{-g'})$ as $i \rightarrow \infty$, for some $g' > 0$. We start with the expression

$$P[A_n(C^{(n)})] = \frac{P[T_{0m}(Z') = n]}{P[T_{0m}(Z) = n]} \prod_{\substack{1 \leq i \leq n \\ r_i+1 \leq j \leq r_i}} \left\{ 1 - \frac{\theta}{ir_i} (1 + E_{i0}) \right\} \quad (1.1)$$

$$P[T_{0n}(Z') = n] = \frac{\theta d}{n} \exp \left\{ \sum_{i \geq 1} [\log(1 + i^{-1}\theta d) - i^{-1}\theta d] \right\} \left\{ 1 + O(n^{-1}\phi_{\{1,2,7\}}(n)) \right\} \quad (1.2)$$

and

$$P[T_{0n}(Z) = n] = \frac{\theta d}{n} \exp \left\{ \sum_{i \geq 1} [\log(1 + i^{-1}\theta d) - i^{-1}\theta d] \right\} \left\{ 1 + O(n^{-1}\phi_{\{1,2,7\}}(n)) \right\} \quad (1.3)$$

Where $\phi_{\{1,2,7\}}(n)$ refers to the quantity derived from Z' . It thus follows that $P[A_n(C^{(n)})] \square Kn^{-\theta(1-d)}$ for a constant K , depending on Z and the r_i' and computable explicitly from (1.1) – (1.3), if Conditions (A_0) and (B_{01}) are satisfied and if $\zeta_i^* = O(i^{-g'})$ from some $g' > 0$, since, under these circumstances, both $n^{-1}\phi_{\{1,2,7\}}(n)$ and $n^{-1}\phi_{\{1,2,7\}}(n)$ tend to zero as $n \rightarrow \infty$. In particular, for polynomials and square free polynomials, the relative error in this asymptotic approximation is of order n^{-1} if $g' > 1$.

For $0 \leq b \leq n/8$ and $n \geq n_0$, with n_0

$$\begin{aligned} & d_{TV}(L(C[1,b]), L(Z[1,b])) \\ & \leq d_{TV}(L(C[1,b]), L(Z[1,b])) \\ & \leq \varepsilon_{\{7,7\}}(n,b), \end{aligned}$$

Where $\varepsilon_{\{7,7\}}(n,b) = O(b/n)$ under Conditions $(A_0), (D_1)$ and (B_{11}) . Since, by the Conditioning Relation,

$$L(C[1,b] | T_{0b}(C) = l) = L(Z[1,b] | T_{0b}(Z) = l),$$

It follows by direct calculation that

$$\begin{aligned} & d_{TV}(L(C[1,b]), L(Z[1,b])) \\ & = d_{TV}(L(T_{0b}(C)), L(T_{0b}(Z))) \\ & = \max_A \sum_{r \in A} P[T_{0b}(Z) = r] \\ & \left\{ 1 - \frac{P[T_{bn}(Z) = n-r]}{P[T_{0n}(Z) = n]} \right\} \quad (1.4) \end{aligned}$$

Suppressing the argument Z from now on, we thus obtain

$$\begin{aligned} & d_{TV}(L(C[1,b]), L(Z[1,b])) \\ & = \sum_{r \geq 0} P[T_{0b} = r] \left\{ 1 - \frac{P[T_{bn} = n-r]}{P[T_{0n} = n]} \right\}_+ \\ & \leq \sum_{r > n/2} P[T_{0b} = r] + \sum_{r=0}^{[n/2]} \frac{P[T_{0b} = r]}{P[T_{0b} = n]} \\ & \times \left\{ \sum_{s=0}^n P[T_{0b} = s] (P[T_{bn} = n-s] - P[T_{bn} = n-r]) \right\}_+ \\ & \leq \sum_{r > n/2} P[T_{0b} = r] + \sum_{r=0}^{[n/2]} P[T_{0b} = r] \\ & \times \sum_{s=0}^{[n/2]} P[T_{0b} = s] \frac{\{P[T_{bn} = n-s] - P[T_{bn} = n-r]\}}{P[T_{0n} = n]} \\ & + \sum_{s=0}^{[n/2]} P[T_{0b} = r] \sum_{s=[n/2]+1}^n P[T = s] P[T_{bn} = n-s] / P[T_{0n} = n] \end{aligned}$$

The first sum is at most $2n^{-1}ET_{0b}$; the third is bound by

$$\begin{aligned} & (\max_{n/2 < s \leq n} P[T_{0b} = s]) / P[T_{0n} = n] \\ & \leq \frac{2\varepsilon_{\{10.5(1)\}}(n/2, b)}{n} \frac{3n}{\theta P_\theta[0,1]}, \\ & \frac{3n}{\theta P_\theta[0,1]} 4n^{-2}\phi_{\{10.8\}}^*(n) \sum_{r=0}^{[n/2]} P[T_{0b} = r] \sum_{s=0}^{[n/2]} P[T_{0b} = s] \frac{1}{2} |r-s| \\ & \leq \frac{12\phi_{\{10.8\}}^*(n)}{\theta P_\theta[0,1]} \frac{ET_{0b}}{n} \end{aligned}$$

Hence we may take

$$\varepsilon_{\{7,7\}}(n,b) = 2n^{-1}ET_{0b}(Z) \left\{ 1 + \frac{6\phi_{\{10,8\}}^*(n)}{\theta P_\theta[0,1]} \right\} P + \frac{6}{\theta P_\theta[0,1]} \varepsilon_{\{10,5(1)\}}(n/2,b) \quad (1.5)$$

Required order under Conditions $(A_0), (D_1)$ and (B_{11}) , if $S(\infty) < \infty$. If not, $\phi_{\{10,8\}}^*(n)$ can be replaced by $\phi_{\{10,11\}}^*(n)$ in the above, which has the required order, without the restriction on the r_i implied by $S(\infty) < \infty$. Examining the Conditions $(A_0), (D_1)$ and (B_{11}) , it is perhaps surprising to find that (B_{11}) is required instead of just (B_{01}) ; that is, that we should need $\sum_{l \geq 2} l\varepsilon_{il} = O(i^{-a_1})$ to hold for some $a_1 > 1$. A first observation is that a similar problem arises with the rate of decay of ε_{i1}

as well. For this reason, n_1 is replaced by n_1 . This makes it possible to replace condition (A_1) by the weaker pair of conditions (A_0) and (D_1) in the eventual assumptions needed for $\varepsilon_{\{7,7\}}(n,b)$ to be of order $O(b/n)$; the decay rate requirement of order $i^{-1-\gamma}$ is shifted from ε_{i1} itself to its first difference. This is needed to obtain the right approximation error for the random mappings example. However, since all the classical applications make far more stringent assumptions about the $\varepsilon_{il}, l \geq 2$, than are made in (B_{11}) . The critical point of the proof is seen where the initial estimate of the difference $P[T_{bn}^{(m)} = s] - P[T_{bn}^{(m)} = s + 1]$. The factor $\varepsilon_{\{10,10\}}(n)$, which should be small, contains a far

tail element from n_1 of the form $\phi_1^\theta(n) + u_1^*(n)$, which is only small if $a_1 > 1$, being otherwise of order $O(n^{1-a_1+\delta})$ for any $\delta > 0$, since $a_2 > 1$ is in any case assumed. For $s \geq n/2$, this gives rise to a contribution of order $O(n^{-1-a_1+\delta})$ in the estimate of the difference $P[T_{bn} = s] - P[T_{bn} = s + 1]$, which, in the remainder of the proof, is translated into a contribution of order $O(n^{-1-a_1+\delta})$ for differences of the form $P[T_{bn} = s] - P[T_{bn} = s + 1]$, finally

leading to a contribution of order $bn^{-a_1+\delta}$ for any $\delta > 0$ in $\varepsilon_{\{7,7\}}(n,b)$. Some improvement would seem to be possible, defining the function g by $g(w) = 1_{\{w=s\}} - 1_{\{w=s+t\}}$, differences that are of the form $P[T_{bn} = s] - P[T_{bn} = s + t]$ can be directly estimated, at a cost of only a single contribution of the form $\phi_1^\theta(n) + u_1^*(n)$. Then, iterating the cycle, in which one estimate of a difference in point probabilities is improved to an estimate of smaller order, a bound of the form $|P[T_{bn} = s] - P[T_{bn} = s + t]| = O(n^{-2}t + n^{-1-a_1+\delta})$ for any $\delta > 0$ could perhaps be attained, leading to a final error estimate in order $O(bn^{-1} + n^{-a_1+\delta})$ for any $\delta > 0$, to replace $\varepsilon_{\{7,7\}}(n,b)$. This would be of the ideal order $O(b/n)$ for large enough b , but would still be coarser for small b .

With b and n as in the previous section, we wish to show that

$$\left| d_{TV}(L(C[1,b]), L(Z[1,b])) - \frac{1}{2}(n+1)^{-1} |1 - \theta| E|T_{0b} - ET_{0b}| \right| \leq \varepsilon_{\{7,8\}}(n,b),$$

Where $\varepsilon_{\{7,8\}}(n,b) = O(n^{-1}b[n^{-1}b + n^{-\beta_{12}+\delta}])$ for any $\delta > 0$ under Conditions $(A_0), (D_1)$ and (B_{12}) , with β_{12} . The proof uses sharper estimates. As before, we begin with the formula

$$d_{TV}(L(C[1,b]), L(Z[1,b])) = \sum_{r \geq 0} P[T_{0b} = r] \left\{ 1 - \frac{P[T_{bn} = n-r]}{P[T_{0n} = n]} \right\}_+$$

Now we observe that

$$\begin{aligned} & \left| \sum_{r \geq 0} P[T_{0b} = r] \left\{ 1 - \frac{P[T_{bn} = n-r]}{P[T_{0n} = n]} \right\}_+ - \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{P[T_{0b} = r]}{P[T_{0n} = n]} \right| \\ & \times \left| \sum_{s=\lfloor n/2 \rfloor+1}^n P[T_{0b} = s] (P[T_{bn} = n-s] - P[T_{bn} = n-r]) \right| \\ & \leq 4n^{-2}ET_{0b}^2 + (\max_{n/2 < s \leq n} P[T_{0b} = s]) / P[T_{0n} = n] \\ & + P[T_{0b} > n/2] \\ & \leq 8n^{-2}ET_{0b}^2 + \frac{3\varepsilon_{\{10,5(2)\}}(n/2,b)}{\theta P_\theta[0,1]}, \quad (1.1) \end{aligned}$$

We have

$$\begin{aligned}
 & \left| \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{P[T_{0b} = r]}{P[T_{0n} = n]} \right. \\
 & \times \left(\left\{ \sum_{s=0}^{\lfloor n/2 \rfloor} P[T_{0b} = s](P[T_{bn} = n-s] - P[T_{bn} = n-r]) \right\} \right. \\
 & \left. - \left\{ \sum_{s=0}^{\lfloor n/2 \rfloor} P[T_{0b} = s] \frac{(s-r)(1-\theta)}{n+1} P[T_{0n} = n] \right\} \right) \Big| \\
 & \leq \frac{1}{n^2 P[T_{0n} = n]} \sum_{r \geq 0} P[T_{0b} = r] \sum_{s \geq 0} P[T_{0b} = s] |s-r| \\
 & \times \left\{ \varepsilon_{\{10.14\}}(n, b) + 2(r \vee s) |1-\theta| n^{-1} \left\{ K_0 \theta + 4\phi_{\{10.8\}}^*(n) \right\} \right\} \\
 & \leq \frac{6}{\theta n P_\theta[0,1]} ET_{0b} \varepsilon_{\{10.14\}}(n, b) \\
 & + 4 |1-\theta| n^{-2} ET_{0b}^2 \left\{ K_0 \theta + 4\phi_{\{10.8\}}^*(n) \right\} \\
 & \left(\frac{3}{\theta n P_\theta[0,1]} \right) \Big\}, \quad (1.2)
 \end{aligned}$$

The approximation in (1.2) is further simplified by noting that

$$\begin{aligned}
 & \sum_{r=0}^{\lfloor n/2 \rfloor} P[T_{0b} = r] \left| \left\{ \sum_{s=0}^{\lfloor n/2 \rfloor} P[T_{0b} = s] \frac{(s-r)(1-\theta)}{n+1} \right\} \right. \\
 & \left. - \left\{ \sum_{s=0}^{\lfloor n/2 \rfloor} P[T_{0b} = s] \frac{(s-r)(1-\theta)}{n+1} \right\} \right| \\
 & \leq \sum_{r=0}^{\lfloor n/2 \rfloor} P[T_{0b} = r] \sum_{s \geq \lfloor n/2 \rfloor} P[T_{0b} = s] \frac{(s-r) |1-\theta|}{n+1} \\
 & \leq |1-\theta| n^{-1} E(T_{0b} 1_{\{T_{0b} > n/2\}}) \leq 2 |1-\theta| n^{-2} ET_{0b}^2, \quad (1.3)
 \end{aligned}$$

and then by observing that

$$\begin{aligned}
 & \sum_{r > \lfloor n/2 \rfloor} P[T_{0b} = r] \left\{ \sum_{s \geq 0} P[T_{0b} = s] \frac{(s-r)(1-\theta)}{n+1} \right\} \\
 & \leq n^{-1} |1-\theta| (ET_{0b} P[T_{0b} > n/2] + E(T_{0b} 1_{\{T_{0b} > n/2\}})) \\
 & \leq 4 |1-\theta| n^{-2} ET_{0b}^2 \quad (1.4)
 \end{aligned}$$

Combining the contributions of (1.2)–(1.3), we thus find

$$\begin{aligned}
 & \left| d_{TV}(L(\overset{\square}{C}[1, b]), L(\overset{\square}{Z}[1, b])) \right. \\
 & \left. - (n+1)^{-1} \sum_{r \geq 0} P[T_{0b} = r] \left\{ \sum_{s \geq 0} P[T_{0b} = s](s-r)(1-\theta) \right\} \right| \\
 & \leq \varepsilon_{\{7.8\}}(n, b) \\
 & = \frac{3}{\theta P_\theta[0,1]} \left\{ \varepsilon_{\{10.5(2)\}}(n/2, b) + 2n^{-1} ET_{0b} \varepsilon_{\{10.14\}}(n, b) \right\} \\
 & + 2n^{-2} ET_{0b}^2 \left\{ 4 + 3 |1-\theta| + \frac{24 |1-\theta| \phi_{\{10.8\}}^*(n)}{\theta P_\theta[0,1]} \right\} \quad (1.5)
 \end{aligned}$$

The quantity $\varepsilon_{\{7.8\}}(n, b)$ is seen to be of the order claimed under Conditions (A_0) , (D_1) and (B_{12}) , provided that $S(\infty) < \infty$; this supplementary condition can be removed if $\phi_{\{10.8\}}^*(n)$ is replaced by $\phi_{\{10.11\}}^*(n)$ in the definition of $\varepsilon_{\{7.8\}}(n, b)$, has the required order without the restriction on the r_i implied by assuming that $S(\infty) < \infty$. Finally, a direct calculation now shows that

$$\begin{aligned}
 & \sum_{r \geq 0} P[T_{0b} = r] \left\{ \sum_{s \geq 0} P[T_{0b} = s](s-r)(1-\theta) \right\} \\
 & = \frac{1}{2} |1-\theta| |E|T_{0b} - ET_{0b}|
 \end{aligned}$$

Example 1.0. Consider the point $O = (0, \dots, 0) \in \square^n$. For an arbitrary vector r , the coordinates of the point $x = O + r$ are equal to the respective coordinates of the vector $r : x = (x^1, \dots, x^n)$ and $r = (x^1, \dots, x^n)$. The vector r such as in the example is called the position vector or the radius vector of the point x . (Or, in greater detail: r is the radius-vector of x w.r.t an origin O). Points are frequently specified by their radius-vectors. This presupposes the choice of O as the “standard origin”. Let us summarize. We have considered \square^n and interpreted its elements in two ways: as points and as vectors. Hence we may say that we leading with the two copies of $\square^n : \square^n = \{\text{points}\}, \square^n = \{\text{vectors}\}$ Operations with vectors: multiplication by a number, addition. Operations with points and vectors: adding a vector to a point (giving a point), subtracting two points (giving a vector). \square^n treated in this way is called an n -dimensional affine space. (An “abstract” affine space is a pair of sets, the set of points and the set of vectors so that the operations as above are defined axiomatically). Notice that vectors in an affine space are also known as “free vectors”. Intuitively, they are not fixed at points and

“float freely” in space. From \mathbb{R}^n considered as an affine space we can precede in two opposite directions: \mathbb{R}^n as an Euclidean space $\Leftarrow \mathbb{R}^n$ as an affine space $\Rightarrow \mathbb{R}^n$ as a manifold. Going to the left means introducing some extra structure which will make the geometry richer. Going to the right means forgetting about part of the affine structure; going further in this direction will lead us to the so-called “smooth (or differentiable) manifolds”. The theory of differential forms does not require any extra geometry. So our natural direction is to the right. The Euclidean structure, however, is useful for examples and applications. So let us say a few words about it:

Remark 1.0. *Euclidean geometry.* In \mathbb{R}^n considered as an affine space we can already do a good deal of geometry. For example, we can consider lines and planes, and quadric surfaces like an ellipsoid. However, we cannot discuss such things as “lengths”, “angles” or “areas” and “volumes”. To be able to do so, we have to introduce some more definitions, making \mathbb{R}^n a Euclidean space. Namely, we define the length of a vector $a = (a^1, \dots, a^n)$ to be

$$|a| := \sqrt{(a^1)^2 + \dots + (a^n)^2} \quad (1)$$

After that we can also define distances between points as follows:

$$d(A, B) := |\overline{AB}| \quad (2)$$

One can check that the distance so defined possesses natural properties that we expect: is it always non-negative and equals zero only for coinciding points; the distance from A to B is the same as that from B to A (symmetry); also, for three points, A, B and C, we have $d(A, B) \leq d(A, C) + d(C, B)$ (the “triangle inequality”). To define angles, we first introduce the scalar product of two vectors

$$(a, b) := a^1 b^1 + \dots + a^n b^n \quad (3)$$

Thus $|a| = \sqrt{(a, a)}$. The scalar product is also denote by dot: $a \cdot b = (a, b)$, and hence is often referred to as the “dot product”. Now, for nonzero vectors, we define the angle between them by the equality

$$\cos \alpha := \frac{(a, b)}{|a||b|} \quad (4)$$

The angle itself is defined up to an integral multiple of 2π . For this definition to be consistent we have to ensure that the r.h.s. of (4) does not exceed 1 by the absolute value. This follows from the inequality

$$(a, b)^2 \leq |a|^2 |b|^2 \quad (5)$$

known as the Cauchy–Bunyakovsky–Schwarz inequality (various combinations of these three names are applied in different books). One of the ways of proving (5) is to consider the scalar square of the linear combination $a + tb$, where $t \in \mathbb{R}$. As $(a + tb, a + tb) \geq 0$ is a quadratic polynomial in t which is never negative, its discriminant must be less or equal zero. Writing this explicitly yields (5). The triangle inequality for distances also follows from the inequality (5).

Example 1.1. Consider the function $f(x) = x^i$ (the i -th coordinate). The linear function dx^i (the differential of x^i) applied to an arbitrary vector h is simply h^i . From these examples follows that we can rewrite df as

$$df = \frac{\partial f}{\partial x^1} dx^1 + \dots + \frac{\partial f}{\partial x^n} dx^n, \quad (1)$$

which is the standard form. Once again: the partial derivatives in (1) are just the coefficients (depending on x); dx^1, dx^2, \dots are linear functions giving on an arbitrary vector h its coordinates h^1, h^2, \dots , respectively. Hence

$$df(x)(h) = \partial_{hf(x)} = \frac{\partial f}{\partial x^1} h^1 + \dots + \frac{\partial f}{\partial x^n} h^n, \quad (2)$$

Theorem 1.7. Suppose we have a parametrized curve $t \mapsto x(t)$ passing through $x_0 \in \mathbb{R}^n$ at $t = t_0$ and with the velocity vector $x(t_0) = v$. Then $\frac{df(x(t))}{dt}(t_0) = \partial_v f(x_0) = df(x_0)(v)$ (1)

Proof. Indeed, consider a small increment of the parameter $t : t_0 \mapsto t_0 + \Delta t$, Where $\Delta t \mapsto 0$. On the other hand, we have $f(x_0 + h) - f(x_0) = df(x_0)(h) + \beta(h)|h|$ for an arbitrary vector h , where $\beta(h) \rightarrow 0$ when $h \rightarrow 0$. Combining it together, for the increment of $f(x(t))$ we obtain

$$\begin{aligned}
 & f(x(t_0 + \Delta t)) - f(x_0) \\
 &= df(x_0)(v \cdot \Delta t + \alpha(\Delta t)\Delta t) \\
 &+ \beta(v \cdot \Delta t + \alpha(\Delta t)\Delta t) \cdot |v \Delta t + \alpha(\Delta t)\Delta t| \\
 &= df(x_0)(v) \cdot \Delta t + \gamma(\Delta t)\Delta t
 \end{aligned} \tag{3}$$

For a certain $\gamma(\Delta t)$ such that $\gamma(\Delta t) \rightarrow 0$ when $\Delta t \rightarrow 0$ (we used the linearity of $df(x_0)$). By the definition, this means that the derivative of $f(x(t))$ at $t = t_0$ is exactly $df(x_0)(v)$. The statement of the theorem can be expressed by a simple formula:

$$\frac{df(x(t))}{dt} = \frac{\partial f}{\partial x^1} x^1 + \dots + \frac{\partial f}{\partial x^n} x^n \tag{2}$$

To calculate the value Of df at a point x_0 on a given vector v one can take an arbitrary curve passing Through x_0 at t_0 with v as the velocity vector at t_0 and calculate the usual derivative of $f(x(t))$ at $t = t_0$.

Theorem 1.8. For functions $f, g : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$,

$$d(f + g) = df + dg \tag{1}$$

$$d(fg) = df \cdot g + f \cdot dg \tag{2}$$

Proof. Consider an arbitrary point x_0 and an arbitrary vector v stretching from it. Let a curve $x(t)$ be such that $x(t_0) = x_0$ and $x'(t_0) = v$.

Hence

$$d(f + g)(x_0)(v) = \frac{d}{dt}(f(x(t)) + g(x(t)))$$

at $t = t_0$ and

$$d(fg)(x_0)(v) = \frac{d}{dt}(f(x(t))g(x(t)))$$

at $t = t_0$ Formulae (1) and (2) then immediately follow from the corresponding formulae for the usual derivative Now, almost without change the theory generalizes to functions taking values in \mathbb{R}^m instead of \mathbb{R} . The only difference is that now the differential of a map $F : U \rightarrow \mathbb{R}^m$ at a point x will be a linear function taking vectors in \mathbb{R}^n to vectors in \mathbb{R}^m (instead of \mathbb{R}). For an arbitrary vector $h \in \mathbb{R}^n$,

$$F(x+h) = F(x) + dF(x)(h)$$

Where $\beta(h) \rightarrow 0$ when $h \rightarrow 0$. We have

$$dF = (dF^1, \dots, dF^m) \text{ and}$$

$$dF = \frac{\partial F}{\partial x^1} dx^1 + \dots + \frac{\partial F}{\partial x^n} dx^n$$

$$= \begin{pmatrix} \frac{\partial F^1}{\partial x^1} & \dots & \frac{\partial F^1}{\partial x^n} \\ \dots & \dots & \dots \\ \frac{\partial F^m}{\partial x^1} & \dots & \frac{\partial F^m}{\partial x^n} \end{pmatrix} \begin{pmatrix} dx^1 \\ \dots \\ dx^n \end{pmatrix} \tag{4}$$

In this matrix notation we have to write vectors as vector-columns.

Theorem 1.9. For an arbitrary parametrized curve $x(t)$ in \mathbb{R}^n , the differential of a map $F : U \rightarrow \mathbb{R}^m$ (where $U \subset \mathbb{R}^n$) maps the velocity vector $x'(t)$ to the velocity vector of the curve $F(x(t))$ in \mathbb{R}^m :

$$\frac{dF(x(t))}{dt} = dF(x(t))(x'(t)) \tag{1}$$

Proof. By the definition of the velocity vector,

$$x(t + \Delta t) = x(t) + x'(t) \cdot \Delta t + \alpha(\Delta t)\Delta t \tag{2}$$

Where $\alpha(\Delta t) \rightarrow 0$ when $\Delta t \rightarrow 0$. By the definition of the differential,

$$F(x+h) = F(x) + dF(x)(h) + \beta(h)|h| \tag{3}$$

Where $\beta(h) \rightarrow 0$ when $h \rightarrow 0$. we obtain

$$F(x(t + \Delta t)) = F(x + \underbrace{x'(t) \cdot \Delta t + \alpha(\Delta t)\Delta t}_h)$$

$$= F(x) + dF(x)(x'(t)\Delta t + \alpha(\Delta t)\Delta t) +$$

$$\beta(x'(t)\Delta t + \alpha(\Delta t)\Delta t) \cdot |x'(t)\Delta t + \alpha(\Delta t)\Delta t|$$

$$= F(x) + dF(x)(x'(t)\Delta t + \gamma(\Delta t)\Delta t)$$

For some $\gamma(\Delta t) \rightarrow 0$ when $\Delta t \rightarrow 0$. This precisely means that $dF(x)x'(t)$ is the velocity vector of $F(x)$. As every vector attached to a point can be viewed as the velocity vector of some curve passing through this point, this theorem gives a clear geometric picture of dF as a linear map on vectors.

Theorem 1.10 Suppose we have two maps $F:U \rightarrow V$ and $G:V \rightarrow W$, where $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m, W \subset \mathbb{R}^p$ (open domains). Let $F: x \mapsto y = F(x)$. Then the differential of the composite map $GoF:U \rightarrow W$ is the composition of the differentials of F and G :

$$d(GoF)(x) = dG(y)odF(x) \quad (4)$$

Proof. We can use the description of the differential. Consider a curve $x(t)$ in \mathbb{R}^n with the velocity vector \dot{x} . Basically, we need to know to which vector in \mathbb{R}^p it is taken by $d(GoF)$. the curve $(GoF)(x(t)) = G(F(x(t)))$. By the same theorem, it equals the image under dG of the Anycast Flow vector to the curve $F(x(t))$ in \mathbb{R}^m . Applying the theorem once again, we see that the velocity vector to the curve $F(x(t))$ is the image under dF of the vector $\dot{x}(t)$. Hence $d(GoF)(\dot{x}) = dG(dF(\dot{x}))$ for an arbitrary vector \dot{x} .

Corollary 1.0. If we denote coordinates in \mathbb{R}^n by (x^1, \dots, x^n) and in \mathbb{R}^m by (y^1, \dots, y^m) , and write

$$dF = \frac{\partial F}{\partial x^1} dx^1 + \dots + \frac{\partial F}{\partial x^n} dx^n \quad (1)$$

$$dG = \frac{\partial G}{\partial y^1} dy^1 + \dots + \frac{\partial G}{\partial y^m} dy^m, \quad (2)$$

Then the chain rule can be expressed as follows:

$$d(GoF) = \frac{\partial G}{\partial y^1} dF^1 + \dots + \frac{\partial G}{\partial y^m} dF^m, \quad (3)$$

Where dF^i are taken from (1). In other words, to get $d(GoF)$ we have to substitute into (2) the expression for $dy^i = dF^i$ from (3). This can also be expressed by the following matrix formula:

$$d(GoF) = \begin{pmatrix} \frac{\partial G^1}{\partial y^1} & \dots & \frac{\partial G^1}{\partial y^m} \\ \dots & \dots & \dots \\ \frac{\partial G^p}{\partial y^1} & \dots & \frac{\partial G^p}{\partial y^m} \end{pmatrix} \begin{pmatrix} \frac{\partial F^1}{\partial x^1} & \dots & \frac{\partial F^1}{\partial x^n} \\ \dots & \dots & \dots \\ \frac{\partial F^m}{\partial x^1} & \dots & \frac{\partial F^m}{\partial x^n} \end{pmatrix} \begin{pmatrix} dx^1 \\ \dots \\ dx^n \end{pmatrix} \quad (4)$$

i.e., if dG and dF are expressed by matrices of partial derivatives, then $d(GoF)$ is expressed by

the product of these matrices. This is often written as

$$\begin{pmatrix} \frac{\partial z^1}{\partial x^1} & \dots & \frac{\partial z^1}{\partial x^n} \\ \dots & \dots & \dots \\ \frac{\partial z^p}{\partial x^1} & \dots & \frac{\partial z^p}{\partial x^n} \end{pmatrix} = \begin{pmatrix} \frac{\partial z^1}{\partial y^1} & \dots & \frac{\partial z^1}{\partial y^m} \\ \dots & \dots & \dots \\ \frac{\partial z^p}{\partial y^1} & \dots & \frac{\partial z^p}{\partial y^m} \end{pmatrix} \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \dots & \frac{\partial y^1}{\partial x^n} \\ \dots & \dots & \dots \\ \frac{\partial y^m}{\partial x^1} & \dots & \frac{\partial y^m}{\partial x^n} \end{pmatrix}, \quad (5)$$

Or

$$\frac{\partial z^\mu}{\partial x^a} = \sum_{i=1}^m \frac{\partial z^\mu}{\partial y^i} \frac{\partial y^i}{\partial x^a}, \quad (6)$$

Where it is assumed that the dependence of $y \in \mathbb{R}^m$ on $x \in \mathbb{R}^n$ is given by the map F , the dependence of $z \in \mathbb{R}^p$ on $y \in \mathbb{R}^m$ is given by the map G , and the dependence of $z \in \mathbb{R}^p$ on $x \in \mathbb{R}^n$ is given by the composition GoF .

Definition 1.6. Consider an open domain $U \subset \mathbb{R}^n$. Consider also another copy of \mathbb{R}^n , denoted for distinction \mathbb{R}_y^n , with the standard coordinates $(y^1 \dots y^n)$. A system of coordinates in the open domain U is given by a map $F:V \rightarrow U$, where $V \subset \mathbb{R}_y^n$ is an open domain of \mathbb{R}_y^n , such that the following three conditions are satisfied:

- (1) F is smooth;
- (2) F is invertible;
- (3) $F^{-1}:U \rightarrow V$ is also smooth

The coordinates of a point $x \in U$ in this system are the standard coordinates of $F^{-1}(x) \in \mathbb{R}_y^n$.

In other words,

$$F:(y^1, \dots, y^n) \mapsto x = x(y^1, \dots, y^n) \quad (1)$$

Here the variables (y^1, \dots, y^n) are the “new” coordinates of the point x .

Example 1.2. Consider a curve in \mathbb{R}^2 specified in polar coordinates as

$$x(t): r = r(t), \varphi = \varphi(t) \quad (1)$$

We can simply use the chain rule. The map $t \mapsto x(t)$ can be considered as the composition of the maps $t \mapsto (r(t), \varphi(t)), (r, \varphi) \mapsto x(r, \varphi)$.

Then, by the chain rule, we have

$$\dot{x} = \frac{dx}{dt} = \frac{\partial x}{\partial r} \frac{dr}{dt} + \frac{\partial x}{\partial \varphi} \frac{d\varphi}{dt} = \frac{\partial x}{\partial r} \dot{r} + \frac{\partial x}{\partial \varphi} \dot{\varphi} \quad (2)$$

Here \dot{r} and $\dot{\varphi}$ are scalar coefficients depending on t , whence the partial derivatives $\frac{\partial x}{\partial r}, \frac{\partial x}{\partial \varphi}$ are vectors depending on point in \square^2 . We can compare this with the formula in the “standard” coordinates:

$\dot{x} = e_1 \dot{x} + e_2 \dot{y}$. Consider the vectors $\frac{\partial x}{\partial r}, \frac{\partial x}{\partial \varphi}$. Explicitly we have

$$\frac{\partial x}{\partial r} = (\cos \varphi, \sin \varphi) \quad (3)$$

$$\frac{\partial x}{\partial \varphi} = (-r \sin \varphi, r \cos \varphi) \quad (4)$$

From where it follows that these vectors make a basis at all points except for the origin (where $r=0$). It is instructive to sketch a picture, drawing vectors corresponding to a point as starting from that point. Notice that $\frac{\partial x}{\partial r}, \frac{\partial x}{\partial \varphi}$ are, respectively, the velocity vectors for the curves $r \mapsto x(r, \varphi)$ ($\varphi = \varphi_0$ fixed) and $\varphi \mapsto x(r, \varphi)$ ($r = r_0$ fixed). We can conclude that for an arbitrary curve given in polar coordinates the velocity vector will have components $(\dot{r}, \dot{\varphi})$ if as a basis we take $e_r := \frac{\partial x}{\partial r}, e_\varphi := \frac{\partial x}{\partial \varphi}$:

$$\dot{x} = e_r \dot{r} + e_\varphi \dot{\varphi} \quad (5)$$

A characteristic feature of the basis e_r, e_φ is that it is not “constant” but depends on point. Vectors “stuck to points” when we consider curvilinear coordinates.

Proposition 1.3. The velocity vector has the same appearance in all coordinate systems.

Proof. Follows directly from the chain rule and the transformation law for the basis e_i . In particular, the elements of the basis $e_i = \frac{\partial x}{\partial x^i}$ (originally, a formal notation) can be understood directly as the velocity vectors of the coordinate lines $x^i \mapsto x(x^1, \dots, x^n)$ (all coordinates but x^i are fixed). Since we now know how to handle velocities

in arbitrary coordinates, the best way to treat the differential of a map $F: \square^n \rightarrow \square^m$ is by its action on the velocity vectors. By definition, we set

$$dF(x_0): \frac{dx(t)}{dt}(t_0) \mapsto \frac{dF(x(t))}{dt}(t_0) \quad (1)$$

Now $dF(x_0)$ is a linear map that takes vectors attached to a point $x_0 \in \square^n$ to vectors attached to the point $F(x) \in \square^m$

$$dF = \frac{\partial F}{\partial x^1} dx^1 + \dots + \frac{\partial F}{\partial x^n} dx^n$$

$$(e_1, \dots, e_m) \begin{pmatrix} \frac{\partial F^1}{\partial x^1} & \dots & \frac{\partial F^1}{\partial x^n} \\ \dots & \dots & \dots \\ \frac{\partial F^m}{\partial x^1} & \dots & \frac{\partial F^m}{\partial x^n} \end{pmatrix} \begin{pmatrix} dx^1 \\ \dots \\ dx^n \end{pmatrix}, \quad (2)$$

In particular, for the differential of a function we always have

$$df = \frac{\partial f}{\partial x^1} dx^1 + \dots + \frac{\partial f}{\partial x^n} dx^n, \quad (3)$$

Where x^i are arbitrary coordinates. The form of the differential does not change when we perform a change of coordinates.

Example 1.3 Consider a 1-form in \square^2 given in the standard coordinates:

$A = -ydx + xdy$ In the polar coordinates we will have $x = r \cos \varphi, y = r \sin \varphi$, hence

$$dx = \cos \varphi dr - r \sin \varphi d\varphi$$

$$dy = \sin \varphi dr + r \cos \varphi d\varphi$$

Substituting into A , we get

$$A = -r \sin \varphi (\cos \varphi dr - r \sin \varphi d\varphi)$$

$$+ r \cos \varphi (\sin \varphi dr + r \cos \varphi d\varphi)$$

$$= r^2 (\sin^2 \varphi + \cos^2 \varphi) d\varphi = r^2 d\varphi$$

Hence $A = r^2 d\varphi$ is the formula for A in the polar coordinates. In particular, we see that this is again a 1-form, a linear combination of the differentials of coordinates with functions as coefficients. Secondly, in a more conceptual way, we can define a 1-form in a domain U as a linear function on vectors at every point of U :

$$\omega(v) = \omega_1 v^1 + \dots + \omega_n v^n, \quad (1)$$

If $v = \sum e_i v^i$, where $e_i = \frac{\partial x}{\partial x^i}$. Recall that the differentials of functions were defined as linear

functions on vectors (at every point), and $dx^i(e_j) = dx^i\left(\frac{\partial x}{\partial x^j}\right) = \delta_j^i$ at every point x .

Theorem 1.9. For arbitrary 1-form ω and path γ , the integral $\int_{\gamma} \omega$ does not change if we change parametrization of γ provide the orientation remains the same.

Proof: Consider $\left\langle \omega(x(t)), \frac{dx}{dt} \right\rangle$ and

$$\left\langle \omega(x(t(t'))), \frac{dx}{dt} \right\rangle \text{ As}$$

$$\left\langle \omega(x(t(t'))), \frac{dx}{dt} \right\rangle = \left\langle \omega(x(t(t'))), \frac{dx}{dt} \right\rangle \cdot \frac{dt}{dt},$$

Let p be a rational prime and let $K = \mathbb{Q}(\zeta_p)$. We write ζ for ζ_p or this section. Recall that K has degree $\varphi(p) = p-1$ over \mathbb{Q} . We wish to show that $O_K = \mathbb{Z}[\zeta]$. Note that ζ is a root of $x^p - 1$, and thus is an algebraic integer; since O_K is a ring we have that $\mathbb{Z}[\zeta] \subseteq O_K$. We give a proof without assuming unique factorization of ideals. We begin with some norm and trace computations. Let j be an integer. If j is not divisible by p , then ζ^j is a primitive p^{th} root of unity, and thus its conjugates are $\zeta, \zeta^2, \dots, \zeta^{p-1}$. Therefore

$$Tr_{K/\mathbb{Q}}(\zeta^j) = \zeta + \zeta^2 + \dots + \zeta^{p-1} = \Phi_p(\zeta) - 1 = -1$$

If p does divide j , then $\zeta^j = 1$, so it has only the one conjugate 1, and $Tr_{K/\mathbb{Q}}(\zeta^j) = p-1$ By linearity of the trace, we find that

$$Tr_{K/\mathbb{Q}}(1-\zeta) = Tr_{K/\mathbb{Q}}(1-\zeta^2) = \dots$$

$$= Tr_{K/\mathbb{Q}}(1-\zeta^{p-1}) = p$$

We also need to compute the norm of $1-\zeta$. For this, we use the factorization

$$x^{p-1} + x^{p-2} + \dots + 1 = \Phi_p(x)$$

$$= (x-\zeta)(x-\zeta^2)\dots(x-\zeta^{p-1});$$

Plugging in $x=1$ shows that

$$p = (1-\zeta)(1-\zeta^2)\dots(1-\zeta^{p-1})$$

Since the $(1-\zeta^j)$ are the conjugates of $(1-\zeta)$, this shows that $N_{K/\mathbb{Q}}(1-\zeta) = p$ The key result for determining the ring of integers O_K is the following.

LEMMA 1.9

$$(1-\zeta)O_K \cap \mathbb{Z} = p\mathbb{Z}$$

Proof. We saw above that p is a multiple of $(1-\zeta)$ in O_K , so the inclusion $(1-\zeta)O_K \cap \mathbb{Z} \supseteq p\mathbb{Z}$ is immediate. Suppose now that the inclusion is strict. Since $(1-\zeta)O_K \cap \mathbb{Z}$ is an ideal of \mathbb{Z} containing $p\mathbb{Z}$ and $p\mathbb{Z}$ is a maximal ideal of \mathbb{Z} , we must have $(1-\zeta)O_K \cap \mathbb{Z} = \mathbb{Z}$ Thus we can write $1 = \alpha(1-\zeta)$

For some $\alpha \in O_K$. That is, $1-\zeta$ is a unit in O_K .

COROLLARY 1.1 For any $\alpha \in O_K$,

$$Tr_{K/\mathbb{Q}}((1-\zeta)\alpha) \in p\mathbb{Z}$$

PROOF. We have

$$Tr_{K/\mathbb{Q}}((1-\zeta)\alpha) = \sigma_1((1-\zeta)\alpha) + \dots + \sigma_{p-1}((1-\zeta)\alpha)$$

$$= \sigma_1(1-\zeta)\sigma_1(\alpha) + \dots + \sigma_{p-1}(1-\zeta)\sigma_{p-1}(\alpha)$$

$$= (1-\zeta)\sigma_1(\alpha) + \dots + (1-\zeta^{p-1})\sigma_{p-1}(\alpha)$$

Where the σ_i are the complex embeddings of K (which we are really viewing as automorphisms of K) with the usual ordering. Furthermore, $1-\zeta^j$ is a multiple of $1-\zeta$ in O_K for every $j \neq 0$. Thus $Tr_{K/\mathbb{Q}}(\alpha(1-\zeta)) \in (1-\zeta)O_K$ Since the trace is also a rational integer.

PROPOSITION 1.4 Let p be a prime number and let $K = \mathbb{Q}(\zeta_p)$ be the p^{th} cyclotomic field. Then $O_K = \mathbb{Z}[\zeta_p] \cong \mathbb{Z}[x]/(\Phi_p(x))$; Thus $1, \zeta_p, \dots, \zeta_p^{p-2}$ is an integral basis for O_K .

PROOF. Let $\alpha \in O_K$ and write

$$\alpha = a_0 + a_1\zeta + \dots + a_{p-2}\zeta^{p-2} \quad \text{With } a_i \in \mathbb{Z}.$$

Then

$$\alpha(1-\zeta) = a_0(1-\zeta) + a_1(\zeta - \zeta^2) + \dots$$

$$+ a_{p-2}(\zeta^{p-2} - \zeta^{p-1})$$

By the linearity of the trace and our above calculations we find that $Tr_{K/\mathbb{Q}}(\alpha(1-\zeta)) = pa_0$

We also have

$Tr_{K/\mathbb{Q}}(\alpha(1-\zeta)) \in p\mathbb{Z}$, so $a_0 \in \mathbb{Z}$ Next

consider the algebraic integer

$(\alpha - a_0)\zeta^{-1} = a_1 + a_2\zeta + \dots + a_{p-2}\zeta^{p-3}$; This is

an algebraic integer since $\zeta^{-1} = \zeta^{p-1}$ is. The same

argument as above shows that $a_1 \in \mathbb{Z}$, and

continuing in this way we find that all of the a_i are

in \mathbb{Z} . This completes the proof.

Example 1.4 Let $K = \mathbb{Q}(\zeta)$, then the local ring

$\mathbb{Z}_{(p)}$ is simply the subring of \mathbb{Q} of rational numbers with denominator relatively prime to p .

Note that this ring $\mathbb{Z}_{(p)}$ is not the ring \mathbb{Z}_p of p -

adic integers; to get \mathbb{Z}_p one must complete $\mathbb{Z}_{(p)}$.

The usefulness of $O_{K,p}$ comes from the fact that it

has a particularly simple ideal structure. Let a be

any proper ideal of $O_{K,p}$ and consider the ideal

$a \cap O_K$ of O_K . We claim that

$a = (a \cap O_K)O_{K,p}$; That is, that a is generated

by the elements of a in $a \cap O_K$. It is clear from

the definition of an ideal that $a \supseteq (a \cap O_K)O_{K,p}$.

To prove the other inclusion, let α be any element

of a . Then we can write $\alpha = \beta/\gamma$ where

$\beta \in O_K$ and $\gamma \notin p$. In particular, $\beta \in a$ (since

$\beta/\gamma \in a$ and a is an ideal), so $\beta \in O_K$ and

$\gamma \notin p$. so $\beta \in a \cap O_K$. Since $1/\gamma \in O_{K,p}$, this

implies that $\alpha = \beta/\gamma \in (a \cap O_K)O_{K,p}$, as

claimed. We can use this fact to determine all of the

ideals of $O_{K,p}$. Let a be any ideal of $O_{K,p}$ and

consider the ideal factorization of $a \cap O_K$ in O_K .

write it as $a \cap O_K = p^n b$ For some n and some

ideal b , relatively prime to p . we claim first that

$bO_{K,p} = O_{K,p}$. We now find that

$a = (a \cap O_K)O_{K,p} = p^n bO_{K,p} = p^n O_{K,p}$

Since $bO_{K,p} = O_{K,p}$. Thus every ideal of $O_{K,p}$ has the

form $p^n O_{K,p}$ for some n ; it follows immediately

that $O_{K,p}$ is noetherian. It is also now clear that

$p^n O_{K,p}$ is the unique non-zero prime ideal in $O_{K,p}$

. Furthermore, the inclusion $O_K \hookrightarrow O_{K,p} / pO_{K,p}$

Since $pO_{K,p} \cap O_K = p$, this map is also

surjection, since the residue class of $\alpha/\beta \in O_{K,p}$

(with $\alpha \in O_K$ and $\beta \notin p$) is the image of $\alpha\beta^{-1}$

in $O_{K/p}$, which makes sense since β is invertible

in $O_{K/p}$. Thus the map is an isomorphism. In

particular, it is now abundantly clear that every non-

zero prime ideal of $O_{K,p}$ is maximal. To

show that $O_{K,p}$ is a Dedekind domain, it remains to

show that it is integrally closed in K . So let

$\gamma \in K$ be a root of a polynomial with coefficients

in $O_{K,p}$; write this polynomial as

$$x^m + \frac{\alpha_{m-1}}{\beta_{m-1}}x^{m-1} + \dots + \frac{\alpha_0}{\beta_0}$$

With $\alpha_i \in O_K$ and

$\beta_i \in O_{K-p}$. Set $\beta = \beta_0\beta_1\dots\beta_{m-1}$. Multiplying by

β^m we find that $\beta\gamma$ is the root of a monic

polynomial with coefficients in O_K . Thus

$\beta\gamma \in O_K$; since $\beta \notin p$, we have

$\beta\gamma/\beta = \gamma \in O_{K,p}$. Thus $O_{K,p}$ is integrally close

in K .

COROLLARY 1.2. Let K be a number field of

degree n and let α be in O_K then

$$N'_{K/\mathbb{Q}}(\alpha O_K) = |N_{K/\mathbb{Q}}(\alpha)|$$

PROOF. We assume a bit more Galois theory than

usual for this proof. Assume first that K/\mathbb{Q} is

Galois. Let σ be an element of $Gal(K/\mathbb{Q})$. It is

clear that $\sigma(O_K)/\sigma(\alpha) \cong O_{K/\alpha}$; since

$\sigma(O_K) = O_K$, this shows that

$$N'_{K/\mathbb{Q}}(\sigma(\alpha)O_K) = N'_{K/\mathbb{Q}}(\alpha O_K)$$

Taking the product over all $\sigma \in Gal(K/\mathbb{Q})$, we have

$$N'_{K/\mathbb{Q}}(N_{K/\mathbb{Q}}(\alpha)O_K) = N'_{K/\mathbb{Q}}(\alpha O_K)^n$$

Since $N_{K/\mathbb{Q}}(\alpha)$ is a rational integer and O_K is a free \mathbb{Z} -

module of rank n ,

$$O_K / N_{K/\mathbb{Q}}(\alpha)O_K \text{ Will have order } N_{K/\mathbb{Q}}(\alpha)^n;$$

therefore

$$N'_{K/\mathbb{Q}}(N_{K/\mathbb{Q}}(\alpha)O_K) = N_{K/\mathbb{Q}}(\alpha O_K)^n$$

This completes the proof. In the general case, let L

be the Galois closure of K and set $[L:\mathbb{Q}] = m$.

III. RESULTS

In this section, we provide results of controlled inflation in arteries of an excised and a beating heart. An investigation of the performance of the control system is presented for various experimental scenarios.

A. Inflation control in an excised heart

Fig. 5 demonstrates the results of an experiment in which two target diameters of 2.3 and 2.7 mm were sequentially targeted. IVOCT images at different diameter milestones, i.e. at initial diameter and at each of the two target diameters. The real-time values of the balloon diameter, the delivered volume, and the inflation pressure. The ones to follow, all diameter values are presented in geometrical distances. The results obtained from this experiment proved the ability of the control system to provide a good convergence to various target diameters. Therefore, the results of this experiment extended the validation of the performance of the control system to compliant balloons and real arteries. The described procedure using an excised heart, could be used with various target diameter sequences and various durations to maintain each target diameter to study the responses of the vessel wall and plaques, e.g. deformation and rupture, under repeatable and controlled conditions. Satisfactory results in this excised heart experiment, where the control system was employed, as well as successful balloon detection in beating heart images, where the control system was not employed, set the table for control experiments in a beating heart setup which provides conditions very close to *in vivo* conditions.

B. Inflation control in a beating heart

In this section, we present the results from three experimental scenarios. The results from the first scenario correspond to controlled inflation for various target diameters. The results from the second scenario correspond to three experiments with various flow rates. The results from the third scenario correspond to controlled inflation to achieve a sequence of various target diameters.

- **Various target diameters**

We performed two inflation control experiments using the same flow rate but different target diameters. At both target diameters, the blood was not yet fully displaced and the balloon had not yet deformed the artery. The goal was to verify if the balloon diameter could be controlled in the initial phase of the inflation when it was still surrounded by blood. The results, including the real-time values of the diameter, the volume, and the pressure, obtained from these two control experiments. In both experiments, the flow rate to deliver or withdraw the liquid was 250 mL/h. The initial diameter was approximately 1.7 mm. The

target diameters were 2.1 mm and 2.3 mm. In order to characterize and compare the system response, in transient and steady state, we defined different performance characteristics, based on control engineering conventions and notations [18]:

1) Rise time (T_r) determined the response time of the control system to achieve a new target diameter. The desired change in the diameter was the difference between the new target diameter and the current diameter. Rise time was defined as the difference between the times that 10 % and 90 % of the desired change in diameter were achieved.

2) Percentage overshoot (PO) represented a criterion to determine the excessive growth of the diameter beyond the target value. It was calculated as:

where D_0 was the initial diameter, D_t was the target diameter, and D_{max} was the maximum diameter achieved.

3) Settling time (T_s) was defined as a parameter that determined convergence. It was defined as the time, after which the convergence error was smaller than 25 % of the difference between the initial and the target diameters:

The applied error margins for settling time are normally less than or equal to 5 %. The reason for choosing a larger value was that, in our experiments, artery contractions constantly disturbed the diameter value. Therefore, the diameter values did not settle within frequently used margins of the target diameters.

4) Steady state diameter (D_{ss}) was defined to determine the degree of convergence. After the settling time, the system was considered to be in steady state. We calculated the average and the standard deviation of the diameter in steady state. Before calculating the performance characteristics, the diameter response versus time was filtered with a low-pass filter to reduce noise. In both cases, although fluctuations were caused by arterial contractions, the difference between the target diameter and the steady-state average diameter was not larger than 20 μm . The rise time and the settling time were increased by 2 s and 1 s, respectively, when the target diameter was increased from 2.1 mm to 2.3 mm. The percentage overshoot decreased when the target diameter was increased.

- **Various flow rates**

We performed three control experiments with the same target diameter but using various flow rates. The results, including the real-time values of the diameter, the volume and the pressure obtained from these control experiments. The initial diameter was approximately 1.7 mm. The target diameter was 2.6 mm. To control the diameter in these experiments, delivery and withdrawal of liquid was performed at flow rates of 100, 250 and 400 mL/h. At the target diameter, the blood was displaced and the artery was deformed by the balloon. The difference between the target diameter and the

steady-state average diameter was on the order of 10 μm . Increasing the flow rate improved the rise time. The rise time corresponding to the flow rate of 100 mL/h was significantly larger than it was in the other experiments, resulting in a much slower convergence. Although increasing the flow rate could improve the rise time, it was expected to increase the percentage overshoot or oscillations in delivery/withdrawal stages as was also observed in experiments with phantoms [13]. Such effect was confirmed in the calculated percentage overshoots and standard deviations of the steady-state diameter. The obtained performance characteristics may be affected by the superimposed effect of the oscillatory contractions of the artery. Nonetheless, they provide a preliminary assessment of the control system performance.

IV. CONCLUSION

In this study, we extended the validity of a previously proposed methodology to control balloon inflation inside porcine arteries. The experiments were performed in excised and beating hearts. In the excised heart experiment, the goal was to investigate if the response time of the control method was sufficient for inflation of a compliant balloon inside porcine arteries. In the beating heart experiment, the first goal was to investigate if the edge detection algorithm was functional when blood was present in IVOCT images. The second goal was to investigate if the speed of the control system was sufficient to provide desired target diameters in presence of arterial contractions. The control system provided good convergence to target diameters in the excised heart experiment. This experiment provided us with a smooth transition to beating heart experiments. In the beating heart experiments, the balloon detection algorithm was successfully applied to estimate the balloon diameter in real-time. Although artery contractions constantly caused disturbances in the actual diameter values, the control system provided average steady-state diameters close to the target diameters. In this study, we worked on sound arteries. Future work should aim at performing a similar study on diseased arteries. Diseased arteries may bring in additional challenges since structures like calcifications can resist balloon inflation. For diseased arteries, more parameters, like the balloon shape, might be needed from the OCT monitoring to provide safe and efficient controlled balloon inflation. Future work should also aim at assessing of the clinical value of our proposed methodology. We have successfully demonstrated that controlled inflation based on IVOCT monitoring is possible. The true usefulness of our methodology lies in the demonstration that it can lead to safer procedures in the clinical world. For example, our methodology opens the path to standard methods of performing multiple step inflations to provide safer artery extension before stenting. It also opens the door to

new approaches to stent deployment. The validation of our methodology as a relevant clinical tool sets the table for exciting new research.

A. Authors and Affiliations

Dr Akash Singh is working with IBM Corporation as an IT Architect and has been designing Mission Critical System and Service Solutions; He has published papers in IEEE and other International Conferences and Journals.

He joined IBM in Jul 2003 as a IT Architect which conducts research and design of High Performance Smart Grid Services and Systems and design mission critical architecture for High Performance Computing Platform and Computational Intelligence and High Speed Communication systems. He is a member of IEEE (Institute for Electrical and Electronics Engineers), the AAAI (Association for the Advancement of Artificial Intelligence) and the AACR (American Association for Cancer Research). He is the recipient of numerous awards from World Congress in Computer Science, Computer Engineering and Applied Computing 2010, 2011, and IP Multimedia System 2008 and Billing and Roaming 2008. He is active research in the field of Artificial Intelligence and advancement in Medical Systems. He is in Industry for 18 Years where he performed various role to provide the Leadership in Information Technology and Cutting edge Technology.

REFERENCES

- [1] D. P. Libby, "Atherosclerosis: The new view," *Scientific American*, vol. 286, no. 5, pp. 46, 2002.
- [2] C. Briguori, J. Tobis, T. Nishida, M. Vaghetti, R. Albiero, C. Di Mario, and A. Colombo, "Discrepancy between angiography and intravascular ultrasound when analysing small coronary arteries," *European Heart Journal*, vol. 23, p. 247, 2002.
- [3] E. Regar, K. A. H. van Leeuwen, P. W. Serruys, and T. G. van Leeuwen, *Optical coherence tomography in cardiovascular research: Informa HealthCare*, 2007.
- [4] T. Kume, T. Akasaka, T. Kawamoto, N. Watanabe, E. Toyota, Y. Neishi, R. Sukmawan, Y. Sadahira, and K. Yoshida, "Assessment of coronary intima-media thickness by optical coherence tomography: comparison with intravascular ultrasound," *Circulation*, vol. 69, p. 903, 2005.
- [5] I. K. Jang, G. J. Tearney, B. MacNeill, M. Takano, F. Moselewski, N. Iftima, M. Shishkov, S. Houser, H. T. Aretz, and E. F. Halpern, "In vivo characterization of

- coronary atherosclerotic plaque by use of optical coherence tomography," *Circulation*, vol. 111, p. 1551, 2005.
- [6] A. Colombo, P. Hall, S. Nakamura, Y. Almagor, L. Maiello, G. Martini, A. Gaglione, S. L. Goldberg, and J. M. Tobis, "Intracoronary stenting without anticoagulation accomplished with intravascular ultrasound guidance," *Circulation*, vol. 91, p. 1676, 1995.
- [7] R. Blasini, F. J. Neumann, C. Schmitt, J. Bökenkamp, A. Schömig, and J. M. Tobis, "Comparison of angiography and intravascular ultrasound for the assessment of lumen size after coronary stent placement: Impact of dilation pressures. Commentary," *Catheterization and cardiovascular diagnosis*, vol. 42, pp. 113-120, 1997.
- [8] B. E. Bouma, G. J. Tearney, H. Yabushita, M. Shishkov, C. R. Kauffman, D. DeJoseph Gauthier, B. D. MacNeill, S. L. Houser, H. T. Aretz, and E. F. Halpern, "Evaluation of intracoronary stenting by intravascular optical coherence tomography," *British Medical Journal*, vol. 89, p. 317, 2003.
- [9] M. Unverdorben, P. Glaeser, R. Degenhardt, G. Rippin, H. K. Berthold, and C. Vallbracht, "Controlled balloon inflation reduces long-term restenosis after percutaneous transluminal coronary angioplasty," *The Journal of invasive cardiology*, vol. 13, no. 12, pp. 774, 2001.
- [10] D. Leibowitz, M. Abu-Gazala, I. Katz, H. Danenberg, H. Nassar, L. Boguslavsky, M. Mosseri, B. Varshitzsky, C. Lotan, A. T. Weiss, Computerized gradual balloon inflation: a novel strategy of coronary angioplasty superior to a standard manual approach, *Cardiovascular Revascularization Medicine*, vol. 10, pp. 45-48, 2009.
- [11] T. Olbrich, D. O. Williams, J. C. Doig and A. Murray, *In vivo* assessment of coronary artery angioplasty and stent deployment from balloon pressure-volume data, *Physiological Measurement*, vol. 27, pp. 213-223, 2006.
- [12] D. P. Faxon, *Restenosis, a Guide to Therapy*, London: Martin Dunitz, 2001.
- [13] H. Azarnoush, S. Vergnole, B. Boulet, and G. Lamouche, "Real-time Control of Angioplasty Balloon Inflation Based on Feedback from Intravascular Optical Coherence Tomography: Preliminary Study on an Artery Phantom," *IEEE Transactions on Biomedical Engineering*, Accepted.
- [14] C. É. Bisaiillon, M. L. Dufour, and G. Lamouche, "Artery phantoms for intravascular optical coherence tomography: healthy arteries," *Biomedical Optics Express*, vol. 2, pp. 2599-2613, 2011.
- [15] G. Lamouche, M. Dufour, M. Hewko, S. Vergnole, B. Gauthier, C.-E. Bisaiillon, J.-P. Monchalain, and M. G. Sowa, "Intravascular optical coherence tomography on a beating heart model," *Journal of Biomedical Optics*, 15, 046023-046027, 2010.
- [16] R. Iliia, H. Cabin, S. McConnell, M. Cleman, and M. Remetz, "Coronary angioplasty with gradual versus rapid balloon inflation: initial results and complications," *Catheterization and cardiovascular diagnosis*, vol. 29, no. 3, pp. 199-202, 1993.
- [17] E. M. Ohman, J. Marquis, D. R. Ricci, R. Brown, M. L. Knudtson, D. J. Kereiakes, J. K. Samaha, J. R. Margolis, A. L. Niederman, and L. S. Dean, "A randomized comparison of the effects of gradual prolonged versus standard primary balloon inflation on early and late outcome. Results of a multicenter clinical trial. Perfusion Balloon Catheter Study Group," *Circulation*, vol. 89, no. 3, pp. 1118, 1994.
- [18] R. C. Dorf, R. H. Bishop, "Modern Control Systems," 11 ed: Prentice Hall, 2008.