The Principle of Inclusion–Exclusion and Möbius Inversion


Suppose you’re given a set $O$ of objects and $P$ of properties (or conditions), where each object satisfies some subset of the set of conditions. As some introductory examples discussed in class have shown: for a given set $S \subseteq P$, it is often easy to find the number $f(S)$ of objects that satisfy all the conditions $S$ and perhaps others but difficult to find the number $g(S)$ of objects that satisfy all the conditions $S$ and no others. The Principle of Inclusion–Exclusion (PIE) is a formula that allows one to find the function $g$ from the function $f$.

The first step in seeing how it works is to pin down the relationship of $f$ to $g$. This is simply that for each $S \subseteq P$,

$$f(S) = \sum_{S \subseteq T \subseteq P} g(T);$$  

(1)

this is because an object is counted by $f(S) \iff$ it is counted by $g(T)$ for some $T \supseteq S$. (Formula (1), of course, is useless in this form: it tells how to compute the easy function $f$ in terms of the hard function $g$. We must somehow “divide by the Σ-sign.”)

The next step is to transform (1) into a vector–matrix equation. I will do this as follows. Make a vector space $V$ of dimension $2^{|P|}$, with one coordinate corresponding to each subset of $P$. Now the functions $f$ and $g$ can be thought of as vectors $\vec{f}$ and $\vec{g}$; the matrix $I$ need to finish the transformation is generally called the “Zeta matrix” $\zeta$:

**Definition.** The zeta matrix for this vector space $V$ is the matrix

$$\zeta(S,T) = \begin{cases} 
1, & \text{if } S \subseteq T; \\
0, & \text{otherwise.}
\end{cases}$$

It is now easy to see that formula (1) is equivalent to the matrix equation

$$\zeta \cdot \vec{g} = \vec{f}.$$  

(1')

The order in which we wrote the coordinates forces $\zeta$ to be upper triangular, and the fact that $\zeta(S,S) = 1$ for all $S \subseteq P$ implies that $\zeta$ has 1’s down the diagonal. Thus det($\zeta$) = 1, so $\zeta$ has an inverse, usually called the “mu matrix” $\mu$. Clearly, then, one can compute $\vec{g}$ from $\vec{f}$ by

$$\mu \cdot \vec{f} = \vec{g},$$  

(2)

so that the information in $\vec{f}$ is sufficient to find (all the coordinates of) $\vec{g}$. The actual PIE formula itself can be thought of as the formula for $\mu(S,T)$, and one of the goals of this handout is to find it.

---

1 The family of subsets of $P$ will be with us, at least in spirit, throughout this discussion. However, we will not need the set of objects (for mathematical purposes) again: all we will need from now on are two functions $f$ and $g$ that satisfy (1).

2 It will also be convenient to arrange the order of the coordinates so that whenever $S \subseteq T$, the coordinate corresponding to $S$ appears to the left of coordinate corresponding to $T$.

3 It is named after Riemann’s zeta function.

4 This one is named after the Möbius function from number theory. In Exercise 7, you will draw the connection.
The formula is not difficult to understand, and there are ways to derive it that are much more direct than the one you will find here: I have chosen to demolish this target by using far more firepower than is needed. My purpose in doing this is to acquaint you with the cannon I will use; that is the other goal of this handout. You should not come away from this handout believing either that the abstract theory below is needed to derive the PIE or that the theory has no other applications than the PIE. Both these beliefs would be completely mistaken: the PIE predates the general theory by centuries; and an important dividend of the theory is that many well-known formulas, the PIE being just one among them, turn out to be different specific examples that can be treated by these same general techniques. (In addition, the theory has generated many new results and has raised previously unasked theoretical questions.)

Since this “sledgehammer–walnut” approach to the PIE involves quite a detour, let me tell you now what the formula will come out to be. As it turns out,

\[
\mu(S,T) = \begin{cases} 
(-1)^{|T|-|S|}, & \text{if } S \subseteq T, \\
0, & \text{otherwise.} 
\end{cases} \tag{3}
\]

From this formula for \( \mu \), the PIE follows immediately:

**Theorem (PIE).** [i]: If \( f \) and \( g \) satisfy (1), then

\[
g(S) = \sum_{S \subseteq T \subseteq P} (-1)^{|T|-|S|} f(T); \tag{4}
\]

[ii]: if in addition \( f(T) \) depends only on \( k = |T| \)—say \( f(T) = \hat{f}(k) \)—then the same holds for \( g \):

\[
g(S) = \hat{g}(\ell),
\]

where

\[
\hat{g}(\ell) = \sum_{k=\ell}^{p} (-1)^{k-\ell} \binom{p-\ell}{k-\ell} \hat{f}(k), \tag{5}
\]

where \( |P| = p \) and \( |S| = \ell \).

**Proof:** Exercise 1: Derive (4) from (3) and (5) from (4).

**II. Partially Ordered Sets and the Incidence Algebra.**

My chosen route to the derivation of (3) begins here. It concerns relationships analogous to (1) (1’) and (2), but the context is much more abstract and general. The central notion I need is that of a “partially ordered set.”

**Definition.** A binary relation \( \preceq \) on a nonempty set \( X \) is a **partial order** on \( X \) iff it has the following properties:

- \( \preceq \) is **reflexive** (\( x \preceq x \) for all \( x \in X \));
- \( \preceq \) is **antisymmetric** (\( x \preceq y \) and \( y \preceq x \implies x = y \)); and
- \( \preceq \) is **transitive**.

\( 5 \) We have previously discussed another type of binary relation, the **equivalence relation**. Compare.
Definition. A set together with a partial order is called a partially ordered set or poset.

Examples. Examples abound. Here is a small sample:
(1): The integers 0 < ⋯ < (n − 1) constitute an n-element poset (denoted “n”) in which every pair of elements is related. Such a poset is called a chain.
(2): At the opposite extreme, an n-element set in which no two different elements are related is called an n-element antichain.
(3): For a set P, let X be the set of all subsets of P—call this set 2P—with S ⪯ T ⇐⇒ S ⊆ T. (It is this example which will link the PIE to this theory.)
(4): Let X = \(D_n\) be the set of all divisors of a positive integer n, with \(d_1 \mid d_2 \implies d_1 < d_2\).
(5): Let X = \(\Pi_n\) be the set of all set partitions of the set \([n]\), with \(\pi_1 \mid \pi_2 \implies \pi_1 \) is a refinement of \(\pi_2\)—this means that every equivalence class of \(\pi_2\) is a union of classes of \(\pi_1\).

Definition. If X is a poset and x and y are elements of X, the interval \([x, y]\) is the set

\[
\{ z \in X : x \leq z \text{ and } z \leq y \}. 
\]

Note that this set must be empty unless \(x \leq y\); note also that despite the term “interval,” \([x, y]\) is not generally a chain.

One often draws a poset as a digraph (called the Hasse diagram) with vertex set X and a directed edge from x to y whenever \(|[x, y]| = 2\); that is, \(x < y\), and there is no z such that \(x < z < y\).

Exercise 2. Draw the Hasse diagrams for \(D_{12}\) and \(\Pi_4\).

Now suppose that X is an n-element poset. I need to associate to X vectors and matrices to X, just as I did to 2P above. Arrange the elements of X in a sequence in such a way that whenever \(x < y\) then \(x\) will appear to the left of \(y\), and let V be the n-dimensional vector space whose coordinates are indexed by the elements of X. A vector \(\vec{f} \in V\) is the same as a function

\[ f : X \to \mathbb{R}, \]

and if two functions \(f\) and \(g\) are from X to \(\mathbb{R}\) are related by

\[ f(x) = \sum_{y \leq x} g(y), \]

then

\[ \vec{g} \cdot \zeta = \vec{f}; \]

as before, \(\zeta\) is upper triangular with 1’s down the diagonal \(\implies \det(\zeta) = 1 \implies\) \(\zeta\) has an inverse \(\mu\)

\[ \vec{f} \cdot \mu = \vec{g}. \]

The first step needed to learn more about \(\mu\) is to embed both \(\zeta\) and \(\mu\) in a larger set of matrices that arises naturally:

Note that the sum in (6) is over \(\{ y : y \leq x \}\) not over \(\{ y : y \geq x \}\); this is why \(\vec{f}\) and \(\vec{g}\) are row vectors here.
Definition. Let $X$ and $V$ be as above. The incidence algebra $A(X)$ of $X$ is the set of all $n \times n$ matrices, with rows and columns indexed by the elements of $X$, with the additional property that a if $M \in A(X)$ then $M(x, y) = 0$ whenever $x \not\preceq y$.

Obviously, $\zeta \in A(X)$, as is $I$ (the identity matrix). It is not so obvious that $\mu \in A(X)$. This is true, however; it is a consequence of the following proposition.

**Proposition 1.** If $M$ and $N$ are any two elements of $A(X)$ then $M + N$ and $M \cdot N$ will also be in $A(X)$.

**Proof:** The “sum” assertion is obvious. To see the “product” assertion, observe that for any $x$ and $y$,

$$ (M \cdot N)(x, y) = \sum_{z \in X} M(x, z)N(z, y) = \sum_{x \preceq z \preceq y} M(x, z)N(z, y), $$

because at least one of the numbers $M(x, z)$ and $N(z, y)$ must be zero unless $x \preceq z$ and $z \preceq y$. However, in order for there to be even one such $z$, it is necessary that $x \preceq y$. Thus $(M \cdot N)(x, y) = 0$ if $x \not\preceq y$. 

**Corollary.** $\mu \in A(X)$.

**Proof:** This follows from Proposition 1 and the fact (from linear algebra) that the inverse of any invertible matrix is a polynomial in that matrix: that is, for some numbers $a_0, a_1, \ldots, a_t$,

$$ M^{-1} = a_0 I + a_1 M + \cdots + a_t M^t. $$

III. Some Theorems about $\mu$.

We have already proved one theorem about $\mu$, namely that $\mu(x, y) = 0$ whenever $x \not\preceq y$—this is the content of the corollary above. In this section, I will prove two more theorems about $\mu$, one easy and one harder; they will be all that I need to prove the PIE.

**Proposition 2.**

$$ \mu(x, y) = \begin{cases} 0, & \text{if } x \not\preceq y; \\ 1, & \text{if } x = y; \text{ and} \\ -\sum_{x \preceq z \prec y} \mu(x, z), & \text{if } x \prec y. \end{cases} \quad (7) $$

**Proof:** We already know that line 1 is correct. To see that lines 2 and 3 are correct, observe first that

$$ (\mu \cdot \zeta)(x, y) = \sum_{z \in X} \mu(x, z)\zeta(z, y) = \sum_{x \preceq z \preceq y} \mu(x, z)\zeta(z, y) = \sum_{x \preceq z \preceq y} \mu(x, z). \quad (8) $$

This means that $\mu$ will be the inverse to $\zeta$ iff for all $x \preceq y$,

$$ \sum_{x \preceq z \preceq y} \mu(x, z) = I(x, y) = \begin{cases} 1, & \text{if } x = y, \text{ and} \\ 0, & \text{if } x \prec y. \end{cases} \quad (9) $$

But line 2 of (7) is equivalent to the first of these conditions, and line 3 of (7) is equivalent to the second. 

4
Proposition 2, which is really a recurrence relation, tells how to compute \( \mu \) iteratively by working your way out from the diagonal. The harder theorem I want to prove tells how to find \( \mu \) recursively in certain cases; the recursion is over posets. More particularly: if the poset \( X \) we’re interested in is constructed from two other posets \( P \) and \( Q \) in a certain nice way (set out below), the theorem tells how to compute \( \mu_x \) from \( \mu_p \) and \( \mu_q \).

**Definition.** Let \( P \) and \( Q \) be finite posets. Their *product* \( P \times Q \) is the set \( P \times Q \) together with the so-called “product order”

\[
(p_1, q_1) \preceq (p_2, q_2) \iff p_1 \preceq p_2 \text{ and } q_1 \preceq q_2.
\]

**Exercise 3.** Show that one way to get the Hasse diagram of \( P \times Q \)—not the only way—is to start with the Hasse diagram of \( P \), replace each vertex with a little copy of \( Q \), and replace each edge with a set of edges connecting corresponding elements of \( Q \). (Hint: What is the covering relation?)

**Example:** The poset \( 2^p \) of all subsets of \( \{1, \ldots, p\} \) is isomorphic to (and so can be thought of as)

\[
\underbrace{2 \times \cdots \times 2}_p,
\]

where \( 2 \) is the two-element chain \( 0 < 1 \).

**Exercise 4.** Prove this by exhibiting an order isomorphism\(^7\) between these two posets.

**Exercise 5.** Show that

\[
\zeta_{P \times Q}((p_1, q_1); (p_2, q_2)) = \zeta_P(p_1, p_2) \cdot \zeta_Q(q_1, q_2).
\]

Exercise 4 says that one can get \( \zeta_{P \times Q} \) by multiplying \( \zeta_P \) and \( \zeta_Q \) in a natural way. The main theorem I will prove in this subject says that one can also get \( \mu \) in this fashion.

**Theorem 1.**

\[
\mu_{P \times Q}((p_1, q_1); (p_2, q_2)) = \mu_P(p_1, p_2) \cdot \mu_Q(q_1, q_2).
\]

**Proof:** It is clear from the definition of the product order that the matrix \( \mu_P(p_1, p_2) \cdot \mu_Q(q_1, q_2) \) is in \( A(P \times Q) \): if \((p_1, q_1) \not\preceq (p_2, q_2)\), then either \(p_1 \not\preceq p_2\) or \(q_1 \not\preceq q_2\), so at least one of the two factors must be zero.

The proof will consist of showing that \( \mu_P(p_1, p_2) \cdot \mu_Q(q_1, q_2) \) satisfies (9); that is, whenever \((p_1, q_1) \preceq (p_2, q_2)\), then

\[
\sum_{(p_1, q_1) \preceq (x, y) \preceq (p_2, q_2)} \mu_P(p_1, x) \cdot \mu_Q(q_1, y) = \begin{cases} 1, & \text{if } (p_1, q_1) = (p_2, q_2), \text{ and} \\ 0, & \text{if } (p_1, q_1) \prec (p_2, q_2). \end{cases} \tag{9 \times 9}
\]

To see this, note first that for any \( p_1, p_2 \in P \) and \( q_1, q_2 \in Q \),

\[
[(p_1, q_1); (p_2, q_2)] = [p_1, p_2] \times [q_1, q_2];
\]

\(^7\) An *order isomorphism* between posets \( X \) and \( Y \) is a bijection \( f : X \to Y \) such that \( x \preceq y \) if and only if \( f(x) \preceq f(y) \).
in other words,

$$(p_1, q_1) \preceq (x, y) \preceq (p_2, q_2) \iff \begin{cases} p_1 \preceq x \preceq p_2 \\ q_1 \preceq y \preceq q_2. \end{cases}$$

This means that

$$\sum_{(p_1, q_1) \preceq (x, y) \preceq (p_2, q_2)} \mu_F(p_1, x)\mu_Q(q_1, y) = \sum_{p_1, x, p_2, q_1, y, q_2} \mu_F(p_1, x)\mu_Q(q_1, y) =$$

$$\left[ \sum_{p_1 \preceq x \preceq p_2} \mu_F(p_1, x) \right] \cdot \left[ \sum_{q_1 \preceq y \preceq q_2} \mu_Q(q_1, y) \right] = I_F(p_1, p_2) \cdot I_Q(q_1, q_2),$$

which clearly equals 1 if $(p_1, q_1) = (p_2, q_2)$ and equals 0 if $(p_1, q_1) \prec (p_2, q_2)$. \(\blacksquare\)

IV. Using Theorem 1 to Prove the PIE.

I will do this by using Exercise 3 and the fact that for the chain \(2 = \{0 < 1\},\)

$$\mu_2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$ 

Exercise 6. For the chain \(n = 0 < 1 < \cdots < (n - 1),\) show that

$$\mu_n = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

the matrix with 1’s down the diagonal, (-1)’s down the superdiagonal, and 0’s elsewhere.

Theorem 2.

$$\mu_{\{2\}}(S, T) = \begin{cases} (-1)^{|T|-|S|}, & \text{if } S \subseteq T, \text{ and} \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Proof: The “0, otherwise” is just the condition that \(\mu \in A(2^P).\) As for the first condition: I will think of this poset as

$$\begin{array}{c} \underbrace{2 \times \cdots \times 2} \end{array}$$

\(p\) times
by letting a set $S \subseteq P$ be represented by the corresponding bit string. Now if

$$S \leftrightarrow (s_1, \ldots, s_p)$$

and

$$T \leftrightarrow (t_1, \ldots, t_p),$$

then $S \subseteq T \iff (\forall i)(s_i \leq t_i)$. Now, since

$$\mu_2(s_i, t_i) = \begin{cases} 1 & \text{if } (s_i, t_i) = (1, 1) \text{ or } (s_i, t_i) = (0, 0); \\
(-1) & \text{if } (s_i, t_i) = (0, 1), \end{cases}$$

we have, for $S \subseteq T$,

$$\mu(S, T) = \prod_{i=1}^{p} \mu_2(s_i, t_i) = \prod_{(s_i, t_i) = (0, 1)} (-1) = (-1)^{|T \cap S|} = (-1)^{|T| - |S|}. \square$$

**Corollary (PIE, dual form).** Let $f$ and $g$ be functions on $2^P$ such that

$$f(S) = \sum_{\emptyset \subseteq T \subseteq S} g(T). \quad (1d)$$

Then

$$g(S) = \sum_{\emptyset \subseteq T \subseteq S} (-1)^{|S| - |T|} f(T). \quad (4d)$$

**Exercise 7.**

[a]: Show that for $n = p_1^{a_1} \cdots p_l^{a_l}$ the poset $D_n$ is isomorphic to $a_1 \times \cdots \times a_l$.

[b]: Show that for $d_1 \leq d_2$ in $D_n$,

$$\mu_{D_n}(d_1, d_2) = \begin{cases} 1, & \text{if } d_1 = d_2; \\
0, & \text{if } \frac{d_2}{d_1} \text{ is divisible by the square of any prime; and} \\
(-1)^t, & \text{if } \frac{d_2}{d_1} = p_1 \cdots p_i. \end{cases}$$

[c]: Show that if $f(n) = \sum_{d|n} g(d)$, then $g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right)f(d)$, where, for any $x \in \mathbb{N}$,

$$\mu(x) = \begin{cases} 1, & \text{if } x = 1; \\
0, & \text{if } x \text{ is divisible by the square of any prime; and} \\
(-1)^t, & \text{if } x = p_1 \cdots p_i. \end{cases}$$

(Formula (10) defines the classical Möbius function mentioned in Footnote 4.)

**V. Computing $\mu_{D_n}$**.

The argument presented in class to prove that
\[ x^n = \sum_{k=1}^{n} \binom{n}{k} (x)_k \]
can be used to compute \( \mu_n(\pi, \sigma) \) for any \( \pi \preceq \sigma \in \Pi_n \). Some additional notation will streamline the presentation.

**Definitions.**

[i], Let \( \pi \in \Pi_n \) be a \( k \)-class partition of \([n]\). The rank \( r(\pi) \) of \( \pi \) is \( k \).

[ii], The minimum element of \( \Pi_n \) will be denoted \( \hat{0} \), and the maximum element will be denoted \( \hat{1} \).

[iii], for \( h: [n] \to [x] \), let the kernel of \( h \), \( \text{Ker}(h) \), be the partition of \([n]\) induced by the equivalence relation \( x \equiv y \iff h(x) = h(y) \).

The computation depends on the following observations. For \( x \geq 1 \), let
\[
g(\pi) = \left| \{ h: [n] \to [x] : \text{Ker}(h) = \pi \} \right|,
\]
and let
\[
f(\pi) = \left| \{ h: [n] \to [x] : \text{Ker}(h) = \sigma \text{ for some } \sigma \succeq \pi \} \right|.
\]
Then:

**Exercise 8.** Show that \( g(\pi) = (x)^{r(\pi)} \) and that \( f(\pi) = x^{r(\pi)} \).

On the other hand, since
\[
f(\pi) = \sum_{\pi \preceq \sigma} g(\sigma),
\]
by Möbius inversion we have
\[
g(\pi) = \sum_{\pi \preceq \sigma} \mu_n(\pi, \sigma) f(\sigma);
\]
in other words (using Exercise 8),
\[
(x)^{r(\pi)} = \sum_{\pi \preceq \sigma} \mu_n(\pi, \sigma) x^{r(\sigma)}. \tag{12}
\]
Since (12) holds for any \( x \geq 1 \) and since both sides of (12) are polynomials, (12) is an equality\(^8\) of polynomials. I will use (12) and Theorem 1 to find a formula for \( \mu_n(\pi, \sigma) \).

[a]: First, let \( \pi = \hat{0} \) in (12). This case of (12) says that
\[
(x)^{r(\hat{0})} = (x)^n = x(x-1) \cdots (x-(n-1)) = \sum_{\sigma \in \Pi_n} \mu_n(\hat{0}, \sigma) x^{r(\sigma)} \equiv x^n = \sum_{k=1}^{n} \binom{n}{k} (x)_k, \tag{13}
\]
\(^8\) More accurately, (12) is a family of equations, one for each \( \pi \in \Pi_n \).
Moreover, the only $\sigma \in \Pi_n$ with $r(\sigma) = 1$ is $\sigma = \hat{1}$. Therefore, equating coefficients of $x$ in (13) gives the formula for $\mu_n(\hat{0}, \hat{1})$:

$$\mu_n(\hat{0}, \hat{1}) = (-1)^{n-1}(n-1)!.$$  (14)

You will see below that because of regularities in the structure of $\Pi_n$, (14) is really all we need to know to find $\mu_n(\pi, \sigma)$.

[b]: Next, consider any interval of the form $[\hat{0}, \sigma]$. Say that $\sigma$ consists of $r(\sigma) = k$ classes of sizes $n_1, \ldots, n_k$, where $\sum_{i=1}^{k} n_i = n$.

**Exercise 9.** Show that, as posets, the interval $[\hat{0}, \sigma]$ and the product $\Pi_n \times \cdots \times \Pi_n$ are isomorphic.

Exercise 9 allows us now to apply Theorem 1 to $[\hat{0}, \sigma]$. Doing this and using (14) produces

$$\mu_n(\hat{0}, \sigma) = \prod_{i=1}^{r(\sigma)} \mu_{n_i}(\hat{0}, \hat{1}) = \prod_{i=1}^{r(\sigma)} (-1)^{n_i-1}(n_i-1)! = (-1)^{n-r(\sigma)} \prod_{i=1}^{r(\sigma)} (n_i - 1)!$$  (15)

We now have a formula for $\mu_n(\hat{0}, \sigma)$.

[c]: Finally, consider any interval $[\pi, \sigma]$, where $\pi \preceq \sigma$. Since any element $x$ in this interval is obtained by merging blocks of $\pi$, these blocks act just like indivisible elements.

**Exercise 10.** Let $\Pi_{r(\pi)}$ be the poset of all partitions of the set whose elements are the blocks of $\pi$; and for each $x \in \Pi_n$ such that $\pi \preceq x$, let $\hat{x} \in \Pi_{r(\pi)}$ be obtained from $x$ by fusing together all the elements of each block of $\pi$—this can be done precisely because $\pi \preceq x$—and treating each fused block as a single element. Show that the map $x \mapsto \hat{x}$ is an isomorphism from $[\pi, \hat{1}]$ onto $\Pi_{r(\pi)}$. Moreover, show that $x \mapsto \hat{x}$ preserves rank: that is, $\hat{x}$ has the same number of classes as $x$.

It follows from Exercise 10 that

$$\mu_n(\pi, \sigma) = \mu_{r(\pi)}(\hat{0}, \hat{\sigma}) = (-1)^{r(\pi)-r(\sigma)} \prod_{i=1}^{r(\sigma)} (\hat{n}_i - 1)!,$$  (16)

where the $i^{th}$ block of $\sigma$ is the union of exactly $\hat{n}_i$ blocks of $\pi$.

**Exercise 11.** Use one of the formulas above to show that

$$\sum_{r(\sigma)=k} \mu_n(\hat{0}, \sigma) = (-1)^{n-k} \binom{n}{k},$$

this gives a sort of combinatorial interpretation of the signed version of Stirling numbers of the first kind.

VI. Concluding Remark.

What this handout contains is just the small tip of a very large iceberg. If you are intrigued, I can spend some class time developing the theory further. In any event, you now know that the theory is there and have some idea what it is about.