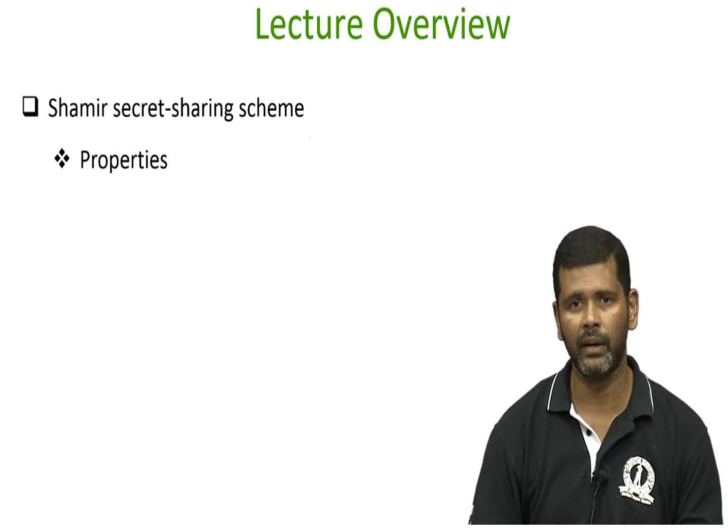
Secure Computation - Part I

Prof. Ashish Choudhury
Department of Computer Science
International Institute of Information Technology, Bangalore

Module - 1
Lecture - 11
Shamir Secret-Sharing

(Refer Slide Time: 00:37)

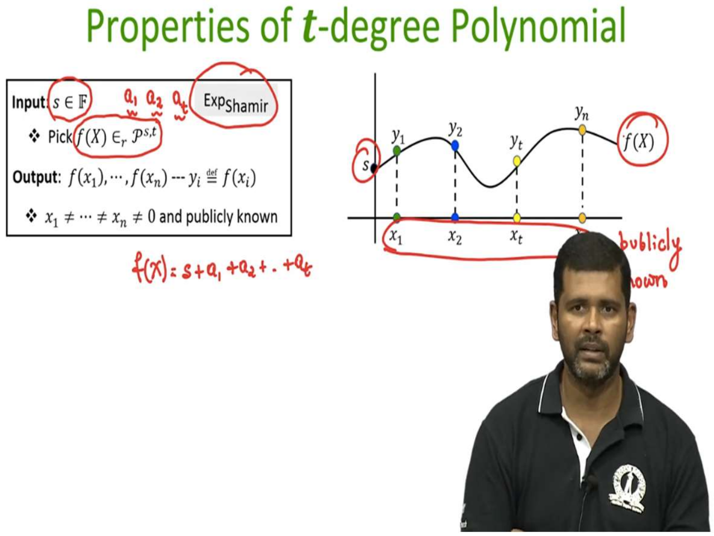


Lecture Overview

- ❑ Shamir secret-sharing scheme
  - ❖ Properties

Hello everyone. Welcome to this lecture. So, now, we will see a very nice secret-sharing scheme called as Shamir secret-sharing, which is an efficient secret-sharing scheme for the threshold setting or threshold adversary for any threshold t strictly less than n.

(Refer Slide Time: 00:48)



Properties of $\boldsymbol{t}$-degree Polynomial

Input: $s \in \mathbb{F}$ a_1 a_2 a_t $\text{Exp}_{\text{Shamir}}$

❖ Pick $f(X) \in_r \mathcal{P}^{s,t}$

Output: $f(x_1), \cdots, f(x_n)$ --- $y_i \overset{\text{def}}{=} f(x_i)$

❖ $x_1 \neq \cdots \neq x_n \neq 0$ and publicly known

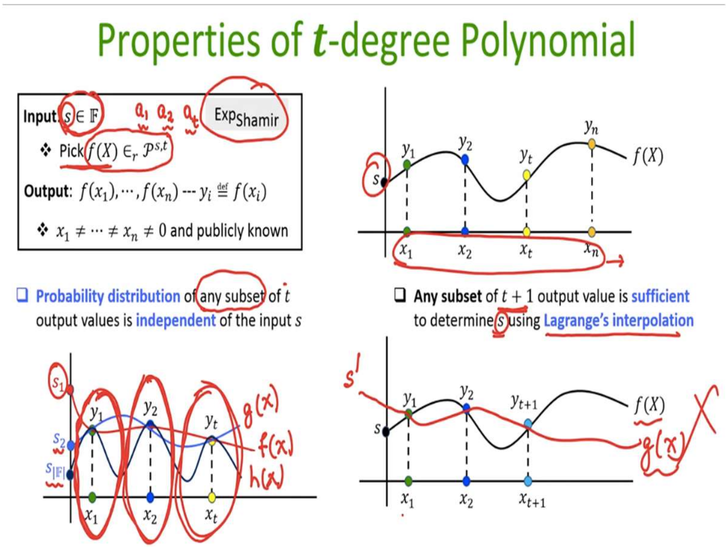
$f(X) = s + a_1 + a_2 + . + a_t$

And this Shamir secret-sharing scheme is based on properties of t-degree polynomials which we have seen extensively in the last lecture. So, let me recall the experiment Shamir which I had discussed in the last lecture in this experiment. Say I am a person who runs this experiment with some private input. My private input will be an element from the field. And in the experiment, what I do is, I randomly pick a t-degree polynomial whose constant term is this value s.

How can I pick that polynomial? Well, I have to just randomly pick the coefficient a_1; I have to randomly pick the coefficient a_2; I have to randomly pick the coefficient a_t. And randomly picking these coefficients means randomly picking field elements as possible coefficients. And then I can say that $f(X) = s + a_1X + a_2X^2 + \cdots + a_tX^t$. That is equivalent to saying that I am picking a polynomial $f(X)$ randomly from this bigger set of all possible t-degree polynomials whose constant term is s.

And output of the experiment is the value of this polynomial at some n publicly known distinct x values. So, I stress that the x components here, at which the polynomial is going to be evaluated will be publicly known. It is only the s and the $f(X)$ polynomial which are the secret part in the experiment.

(Refer Slide Time: 02:32)



Properties of *t*-degree Polynomial

Input: $s \in \mathbb{F}$ a_1 a_2 a_t Exp$_{\text{Shamir}}$

❖ Pick $f(X) \in_r \mathcal{P}^{s,t}$

Output: $f(x_1), \cdots, f(x_n)$ --- $y_i \overset{\text{def}}{=} f(x_i)$

❖ $x_1 \neq \cdots \neq x_n \neq 0$ and publicly known

❑ Probability distribution of any subset of t output values is independent of the input s

❑ Any subset of $t+1$ output value is sufficient to determine s using Lagrange's interpolation

And in the last lecture, we had extensively proved that if I do not give you the entire set of n output values which I have generated in the experiment, but give you any subset of t output values; again I stress, any subset. It need not be the first t output values or the last t output values or the middle t output values. You ask me and I will give you any subset of t output values corresponding to the indices that you have demanded.

From your viewpoint, those t output values could be the evaluation of any possible polynomial of degree-t, whose constant term could be s, or whose constant term could be s', or whose constant term could be s'' and so on. You cannot pinpoint whether what polynomial or what was the input s I had used in the experiment. That is what we had proved rigorously in the last lecture.

Again, pictorially, if I give, if you have say for instance asked for the first t output values; so, basically, asking for the t output values means, I have given you t points on an unknown curve $f(X)$. Why unknown curve? Because $f(X)$ was secretly chosen by me, the person who has run the experiment. You are only getting the output values $y_1, \dots, y_t$, because you asked for the first t output values.

So, first t output values means you now have these t distinct points. What we had proved in the last lecture is, from your viewpoint through these t distinct points, every possible t-degree curve can pass. It could be the case that the curve that I had chosen was this curve, say $f(X)$ whose constant term is s_1, because $(0, s_1)$ along with this t distinct point constitutes 1 possible t-degree curve.

Or it could be the case that, it is the blue curve which I have used in my experiment, and the constant term of this blue curve is s_2. And that means, s_2 was my input in the experiment. That means, both the probability of the occurrence, the probability that I have run the experiment with this $f(X)$ curve or the probability that I have run the experiment with $g(X)$ curve is equi-likely from your viewpoint.

Or it could be the case that this is a black curve which I have usen, say $h(X)$ curve whose constant term is another possible element from the field. That means, just based on the knowledge of this t points, t values, you cannot pinpoint that what was my input, whether it was s_1 or s_2 or s_3 or any arbitrary element from the field. On the other hand, we also know the following:

If instead of t output values, I gave you t + 1 output values. That means, I run the experiment with some private input s; and now, I give you t + 1 output values. It could be again, any t + 1 output values. It could be the first set of t + 1 output values, or it could be the last subset of t + 1 output values, or it could be any subset of t + 1 output values. Then you can completely determine what was my input s in the experiment.

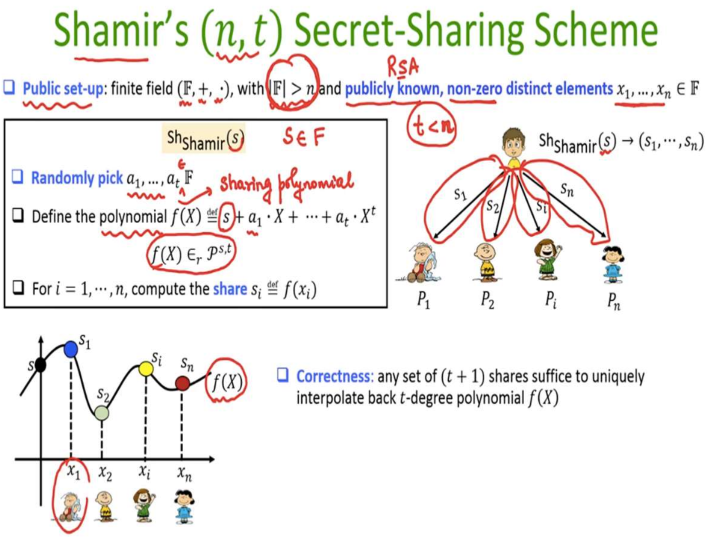
Why so? Because those t + 1 output values actually constitute t + 1 distinct points on the $f(X)$ curve which I have picked in this step. And we know that through Lagrange's interpolation, you can find out the unique t-degree curve passing through t + 1 distinct points. So, basically, in this experiment, what I have done is, I have chosen a random t-degree polynomial whose constant term was my input.

And my claim is that, if you are restricted, if you are given the access only to t of the possible n output values, then you cannot tell what was my input. But if you are given access to t + 1 output values that I have generated in the experiment, then you can completely find out what was the unique curve. Now you cannot say that through this t + 1 curve, it could be this black curve $f(X)$ whose constant term is s, or it could be this red curve $g(X)$ whose constant term could be s'.

No, that cannot be possible, because, now you are given t + 1 points, and 2 different t-degree curves cannot have more than t common values. That is not possible. There can be 1 and only 1 unique t-degree curve passing through the given t + 1 output values, if you are given access to t + 1 output values of this experiment. So, now, based on these 2 properties; what are the 2 properties?

Access up to t output values of the experiment, leaks nothing about the secret of the experiment. Access to t + 1 or more output values of the experiment give full knowledge of the input of this experiment. These are the 2 properties.

(Refer Slide Time: 08:03)



Shamir's (n, t) Secret-Sharing Scheme

□ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \ldots, x_n \in \mathbb{F}$

RSA

$t < n$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$ $s \in \mathbb{F}$

□ Randomly pick $a_1, \ldots, a_t \in \mathbb{F}$ sharing polynomial

□ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

□ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

$\mathsf{Sh}_{\mathsf{Shamir}}(s) \rightarrow (s_1, \cdots, s_n)$

□ Correctness: any set of $(t+1)$ shares suffice to uniquely interpolate back t-degree polynomial $f(X)$

And this is precisely what we need to design an n, t secret-sharing scheme. So, remember, for an n, t secret-sharing scheme, we need a sharing mechanism, namely, we want a mechanism to generate n shares from the secret, in such a way that any subset, any collection of t or less number of shareholders should not learn anything about the secret. Whereas any collection of t + 1 or more number of shareholders, they will have a mechanism to get back the secret uniquely.

And that is precisely what we have discussed till now. This experiment Shamir is nothing but your secret-sharing algorithm. A very beautiful scheme attributed to Adi Shamir. So, this Shamir is basically the S of your famous RSA public-key algorithm. And what is the public setup that we need to instantiate this n, t Shamir's secret-sharing scheme? By public setup I mean that everyone will know that this setup is going to be used.

The setup will be the description of some finite field with some abstract plus operation and abstract dot operation, whose cardinality is at least n. We will see why we want to have this restriction. And we will have the knowledge of, public knowledge of n distinct elements from the field which are non-zero. And this is possible because you are choosing a field whose cardinality is at least n, not at least n, greater than n.

So, that means, apart from the 0 elements, you need at least n distinct non-zero elements also to be present in the field. And that is why we need this restriction. We will discuss later why we need this restriction. And this will be publicly known, I stress. Now, what is the sharing algorithm that will be used to generate the shares for the secret? So, if s is the secret which needs to be shared and s has to be an element from the field.

So, that is why, in this secret-sharing scheme, unlike the secret-sharing schemes that we had seen till now, where all the operations were performed over a group, where the secret was an element of the group, shares were also elements of the group; now, we need a more powerful algebraic structure, namely, we now require a field where we should have the mechanism of both performing the plus operation as well as the dot operation.

So, my secret has to be an element of the field and my shares also will be elements of the field. So, what will be the sharing algorithm? I randomly pick the coefficients $a_1, \dots, a_t$ belonging to the field. So, sorry for the typo, this should be belonging to $\mathbb{F}$. I randomly pick t elements from the field. And then I define what I call as the sharing polynomials. I call this $f(X)$ polynomial as the sharing polynomial, where the coefficients are $a_1, a_2, \dots, a_t$.

And the constant coefficient is fixed. That cannot be any random element from the field. The constant coefficient here is the value which we want to share, namely your secret. So, randomly picking the coefficients $a_1, \dots, a_t$. As I said in the earlier lecture, it is equivalent to randomly picking the polynomial $f(X)$ from the bigger set of all possible t-degree polynomials whose constant term is the secret s.

That means, if I run this sharing algorithm again and again and again, my $f(X)$ polynomial will take different values. It will be different possible polynomials with different probabilities, depending upon what are the values of $a_1, \dots, a_t$. And $a_1, \dots, a_t$ are not deterministically picked, they are randomly chosen. This is because, remember, we want the sharing algorithms to be randomised algorithms.

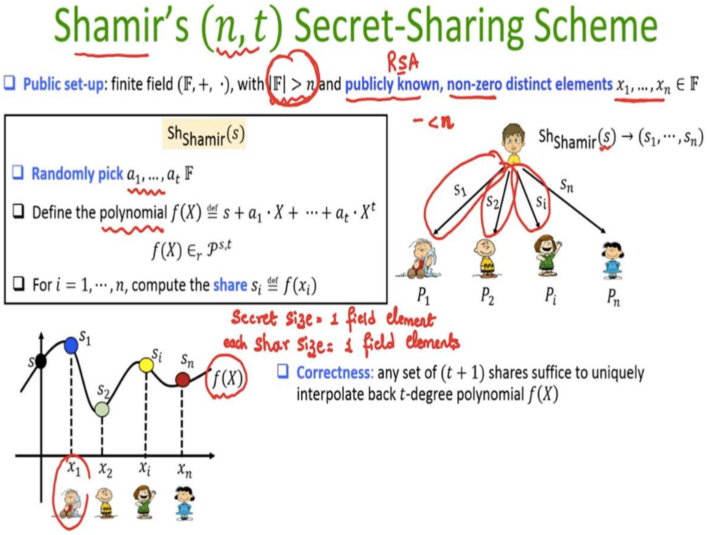
And what are the shares? The ith share will be the evaluation of this polynomial at $X = x_i$. So, the first share will be the value of the polynomial at $X = x_1$, the second share will be the value of the polynomial at $X = x_2$ and so on. That means, imagine if n is equal to, say 4. If $n = 4$ and if dealer has some secret s which it wants to share, then it runs the secret-sharing algorithm with whatever is the value of t that is given to him or not given to him, which is publicly known.

So, remember, t and n, their values will be publicly known. So, the dealer will pick a random t-degree polynomial whose constant term is s. And then, to the first party, he will give the evaluation of this polynomial that he has chosen at x_1; to the second party, he will give the evaluation of this polynomial at x_2; to ith party, it will give the evaluation of this polynomial at x_i; and to the nth party, it will give the value of the polynomial at x_n.

And everyone will know that the party P_1 is getting the value of dealer's polynomial, sharing polynomial at $X = x_1$. That will be a public knowledge. But what is the polynomial dealer has used? That is a private input. That is known only to the dealer. And remember, when dealer is communicating the shares to the respective parties, it is happening over a private channel between the dealer and that specific party; that is an assumption.

So, remember, we are making the assumption that there is a secure channel between the dealer and every shareholder. So, now let us see whether this simple secret-sharing scheme satisfies the correctness and the privacy property. But before going into that, let us verify whether this secret-sharing scheme is efficient or not. So, unlike your secret-sharing scheme by Benaloh et al. and Ito et al. which we had seen in earlier lectures for the n, t threshold setting, where the number of shares could be exponentially large; here, each party is just getting 1 field element.

(Refer Slide Time: 14:51)



Shamir's (n, t) Secret-Sharing Scheme

RSA

❑ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \dots, x_n \in \mathbb{F}$

$- < n$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

❑ Randomly pick $a_1, \dots, a_t$ $\mathbb{F}$

❑ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \dots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

$\mathsf{Sh}_{\mathsf{Shamir}}(s) \to (s_1, \cdots, s_n)$

Secret size = 1 field element

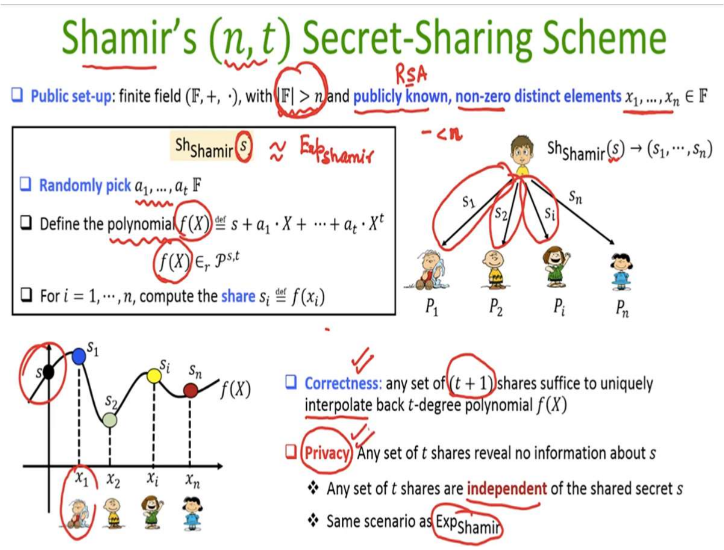
each shar size = 1 field elements

❑ Correctness: any set of $(t + 1)$ shares suffice to uniquely interpolate back t-degree polynomial $f(X)$

So, secret size here is 1 field element. And share size, each share size, namely, each party gets just 1 field element as a share. Unlike your two secret-sharing schemes which we had discussed earlier, where depending upon the number of authorised subsets or the number of forbidden subsets where a party can belong, a party may end up getting exponentially many number of elements as potential shares.

But here, it does not matter. We are first of all, not dividing the set of n shareholders into groups. We are just doing 1 step action; take a polynomial, evaluate at n values, distribute, that is all. And that is why this scheme is a very efficient scheme. It does not require the dealer or the parties to perform exponential amount of computation, communication, or it does not require exponential amount of storage. So, now, the efficiency part is very easy here to verify. Now, let us see whether the correctness and the privacy requirements of secret-sharing are satisfied.

(Refer Slide Time: 16:11)



Shamir's (n, t) Secret-Sharing Scheme

RSA

❑ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \ldots, x_n \in \mathbb{F}$

- < n

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$ ≈ $\mathsf{Exp}_{\mathsf{Shamir}}$

❑ Randomly pick $a_1, \ldots, a_t$ $\mathbb{F}$

❑ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

$\mathsf{Sh}_{\mathsf{Shamir}}(s) \rightarrow (s_1, \cdots, s_n)$

❑ Correctness: any set of $(t+1)$ shares suffice to uniquely interpolate back t-degree polynomial $f(X)$ ✓

❑ Privacy ✓ Any set of t shares reveal no information about s

❖ Any set of t shares are independent of the shared secret s

❖ Same scenario as $\mathsf{Exp}_{\mathsf{Shamir}}$

For correctness, we have to prove that, if a group of t + 1 shareholders come together, then whether they will be able to get back the secret or not. And the answer is yes. Because, if any subset of any group of t + 1 shareholders come together, and they tell each other their respective shares that they have got from the dealer, then by applying the Lagrange's interpolation, they can uniquely get back the polynomial, entire polynomial $f(X)$ that dealer has chosen.

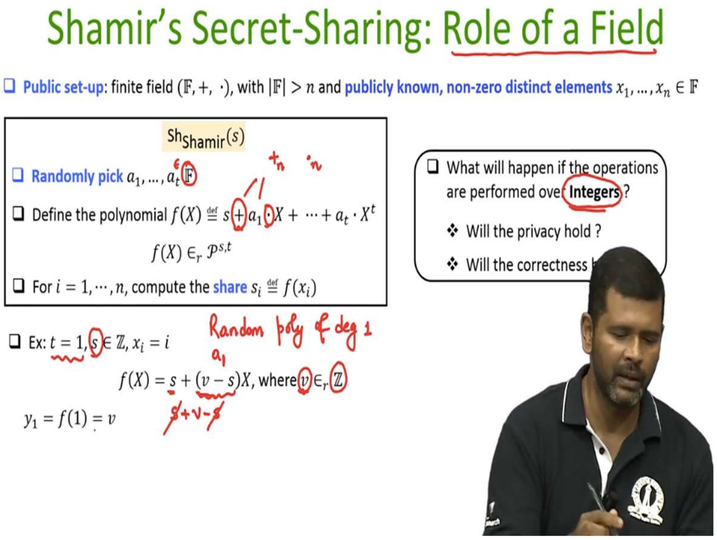
And once $f(X)$ is reconstructed back, they just have to take the constant coefficient of that $f(X)$ polynomial that is the secret s. So, correctness property is satisfied. Now, let us see whether the privacy property is satisfied. For privacy, we have to argue that if you take any group of t or less number of shareholders, they learn absolutely nothing about the secret. And we have rigorously formalised; what do we mean by learning absolutely nothing about a secret?

That means that from the viewpoint of those t shareholders, the probability distribution of the shares that they will see or they will receive from the dealer is should be independent of the dealer's secret. And this is what we had precisely proved as part of the properties of the experiment Shamir. This Shamir’s secret-sharing algorithm is equivalent, is precisely your experiment Shamir which we had rigorously analysed in the last lecture.

And in the last lecture, we had formally proved that, if someone runs this experiment Shamir with some private input s, and just gives any subset of t outputs generated in the experiment, then the probability distribution of those t output, those could, will be independent of what exactly is the value of s with which the experiment has been executed. And that is precisely is what we want to argue as part of the privacy property of your Shamir’s secret-sharing.

That is why; and this holds, even if the group of those t shareholders who want to learn the secret are computationally unbounded, even if they are given unbounded time resources, from their viewpoint, it could be any t-degree curve with which dealer has run the sharing algorithm, and that t-degree polynomial when evaluated at the x-coordinates corresponding to the t parties, gives the corresponding y value. They cannot pinpoint what exactly was the t-degree x t-degree $f(X)$ curve with which D participated in the experiment. So, that proves the privacy property as well.

(Refer Slide Time: 19:17)



Shamir’s Secret-Sharing: Role of a Field

❑ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \dots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

❑ Randomly pick $a_1, \dots, a_t \in \mathbb{F}$

❑ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$+_n$ $\cdot_n$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

❑ What will happen if the operations are performed over Integers ?

❖ Will the privacy hold ?

❖ Will the correctness h

❑ Ex: $t = 1$, $s \in \mathbb{Z}$, $x_i = i$

Random poly of deg 1

$f(X) = s + \overset{a_1}{(v - s)}X$, where $v \in_r \mathbb{Z}$

$y_1 = f(1) = v$

$\cancel{s} + v - \cancel{s}$

So, it is a very simple looking, but a very elegant construction. And this is one of the most popular cryptographic primitives used in several real-world scenario. So, now, you might be wondering that why we want a field here. So, remember, all the operations in the Shamir’s secret-sharing protocol algorithm requires operations to be performed over the field. The Secret has to be an element of the field, the shares are the elements of the field, the plus operation and the dot operations, all of them are performed over a field, unlike your previous secret-sharing schemes where just a group was sufficient.

Now, you might be wondering that why it matters whether I perform a operation on the ring or a field or a group. The thing is, if I want to deploy, implement this scheme inside my computer, then I have to implement this field plus operation and field dot operation on top of the computer plus and computer dot operation, because this plus and this dot operation, the field plus and the field dot operation is not exactly the integer plus an integer dot operation, or the way computers perform the plus operation or the dot operation.

So, for instance, this plus and dot operation could be addition modulo n and multiplication modulo n, where after every addition, you have to do a modulo. And that will be the overall plus operation. So, that is why, I would prefer a secret-sharing protocol where the plus operation, the dot operations are actually same as my computer plus and computer dot operation, because, then I do not have to worry about separately implementing the plus and the dot operation of the corresponding algebraic structure which I am using in my secret-sharing algorithm.

So, for instance, if I could have performed Shamir Shamir's secret-sharing algorithm over a ring, modulo 2 power l, where l is some parameter, then that is precisely performing computer operations over l bit registers. So, I do not have to separately worry about the plus and the dot operation of the ring. So, that is why, the choice of the underlying algebraic structure which is used inside your algorithm matters a lot.

So, now, for Shamir's secret-sharing, we need a field. The question is, why cannot I perform the operations over the integers? By the way, the integer plus operation and integer dot operation, the set of integers along with plus and dot do not constitute a field, because, we do not have multiplicative inverse for several elements and so on. So, the question is, if instead of performing the operations over the field, if I perform the operations over the integers, where the plus and the dot operations are the integer plus and the integer dot operation, will I still get the privacy guarantees? will I still get the correctness guarantees?

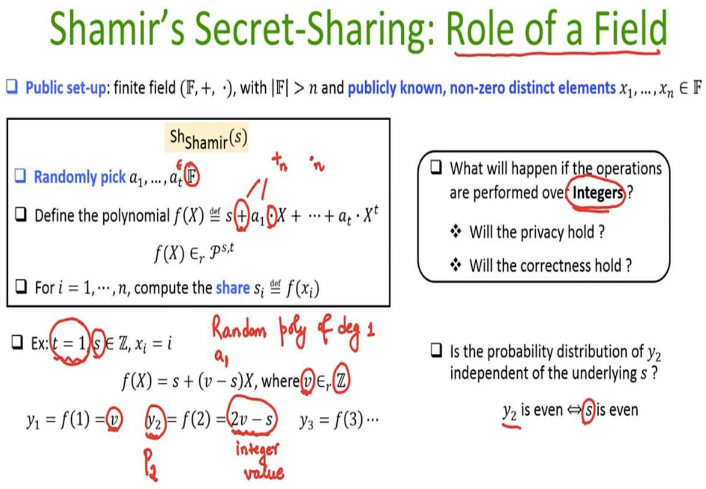
And the answer for both these questions is no. So, let us first try to understand that why the privacy will break if instead of performing operations over the field, I pick, I perform the operations over the integers. So, I denote a set of integers as $\mathbb{Z}$. So, to demonstrate the weakness, let us take a very simple case. Suppose my $t = 1$, n could be anything. And my $x_1, x_2, \dots, x_n$ are 1, 2, 3, 4, 5, 6.

So, if there are n parties, the ith party will get the value of dealer's polynomial at $X = i$. And now let us try to mimic Shamir's secret-sharing algorithm assuming all the operations are performed over the integers with t = 1. And imagine that dealer's secret is now an integer value, it is no longer a field element, it is an element from the set of integers. It could be negative, it could be positive, anything.

So, now, what the dealer has to do? If it is mimicking Shamir's secret-sharing algorithm with $t = 1$, it has to pick a polynomial, random polynomial of degree-1. That is what it has to do. Whose constant term should be the integer s? So, suppose it picks the polynomial $f(X)$, which is s plus the coefficient of X^1 being $v - s$ were v is a randomly chosen integer. So, the requirement would have been, he has to pick the coefficient a_1 randomly, the dealer.

And indeed, the coefficient a_1 which is $v - s$ will be random, if v is randomly chosen. If v is a randomly chosen integer, then irrespective of what is the value of s, the value $v - s$ will also be a random integer. So, now he computes the shares. So, the share for the first party will be the value of this polynomial at $x = 1$, with no mod operation. It is a usual integer plus and integer multiplication. So, f of 1 will be $s + v - s$. s and s cancels out; so, that is why the first share will be the value v.

(Refer Slide Time: 24:51)



Shamir's Secret-Sharing: Role of a Field

□ **Public set-up**: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and **publicly known, non-zero distinct elements** $x_1, \ldots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

□ **Randomly pick** $a_1, \ldots, a_t \overset{\in}{} \mathbb{F}$

$+_n$ $\cdot_n$ //

□ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

□ For $i = 1, \cdots, n$, compute the **share** $s_i \overset{\text{def}}{=} f(x_i)$

□ What will happen if the operations are performed over **Integers**?

❖ Will the privacy hold ?

❖ Will the correctness hold ?

□ Ex: $t = 1$ $s \in \mathbb{Z}, x_i = i$

Random poly of deg 1

a_1

$f(X) = s + (v - s)X$, where $v \in_r \mathbb{Z}$

$y_1 = f(1) = v$ $\quad \underset{P_2}{y_2} = f(2) = \underset{\text{integer value}}{2v - s}$ $\quad y_3 = f(3) \cdots$

□ Is the probability distribution of y_2 independent of the underlying s ?

y_2 is even $\Leftrightarrow$ s is even

The value of the second share will be the value of this f polynomial at $x = 2$, which will be $2 \times v - s$ and so on. Now, what is the value of t? $t = 1$ here. So, this secret-sharing instantiation should have the property that, if I take any single share, say the share of the first party or the share of the second party or the share of the ith party, that should be independent, that should have nothing to do with what is the value of the secret s which dealer has shared.

That is a privacy property. Namely, just 1 shareholder, he alone should not be able to learn anything about the secret, because I have instantiated this protocol with t = 1. Whereas, any set of 2 or more number of shareholders, they should be able to come together and get back the secret. That is the correctness requirement. Now, if I consider y_1, the value of the first share is v, and v is a randomly chosen integer, independent of the value of the secret, so, it does not leak anything about the secret; that is fine.

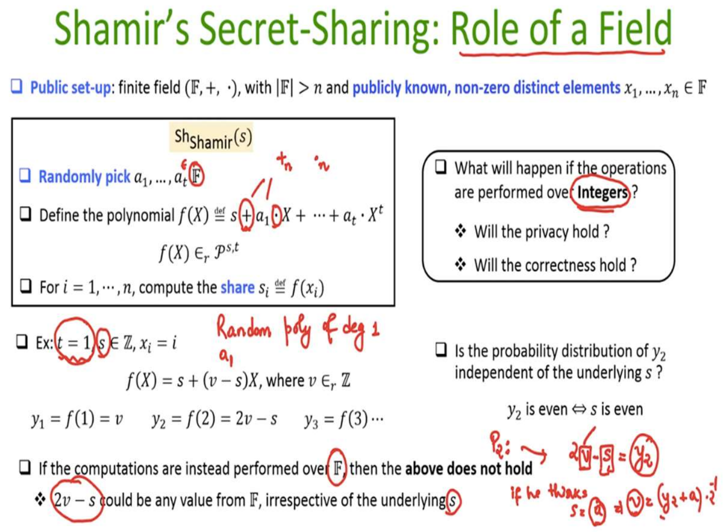
But what about y_2? Suppose, if P_2 is very curious and it is try to learning something about the underlying secret, the secret-sharing algorithm should have ensured that if P_2 is curious and try to learn anything about dealer's secret, he should not learn anything. For him, it could have been any integer which dealer has shared; and he got this share; from the viewpoint of the second shareholder, it could have been any integer value, which dealer has shared and the second shareholder got this value y_2.

But if you see here closely, the value of the second share is the integer value, I stress, integer value, $2v - s$, which will be an even value if and only if dealer's secret would have been an even value. Because, 2 times v will be an even quantity. And now, from an even quantity, if you subtract some quantity, the overall result will be; or the event depending upon whatever you have subtracted is odd or even, respectively.

That means, if the value of y_2, that means P_2 who is the second shareholder, if he compares y_2 and checks whether it is odd or even; if it is odd, then he can simply tell that, okay, dealer's secret was an odd integer, it cannot be an even integer. And that is a breach of privacy, because, remember, the privacy property demands that, from the viewpoint of the unauthorised subset of parties, it could have, it should be the case that their shares should be equiprobably the shares for any candidate element from the secret space.

And the secret space is the set of all possible integers. But now, what this second party is able to do is, based on whether his share is odd or even, he can narrow down whether the dealer's secret was an odd value or whether it is an even value. And that itself is a breach of privacy, which should not be allowed. That means, if I instantaneous this secret-sharing scheme, where I perform all the plus and dot operations as per integers, then there is a breach of privacy.

(Refer Slide Time: 28:35)



Shamir's Secret-Sharing: Role of a Field

❑ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \ldots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

❑ Randomly pick $a_1, \ldots, a_t \in \mathbb{F}$

$+_n$ $\cdot_n$

❑ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

❑ What will happen if the operations are performed over Integers?

- ❖ Will the privacy hold ?
- ❖ Will the correctness hold ?

❑ Ex: $t = 1$, $s \in \mathbb{Z}$, $x_i = i$ Random poly of deg 1

a_1

$f(X) = s + (v - s)X$, where $v \in_r \mathbb{Z}$

$y_1 = f(1) = v$ $y_2 = f(2) = 2v - s$ $y_3 = f(3) \cdots$

❑ Is the probability distribution of y_2 independent of the underlying s ?

y_2 is even ⇔ s is even

$P_2: \rightarrow 2\boxed{v} - \boxed{s} = y_2$

❑ If the computations are instead performed over $\mathbb{F}$ then the **above does not hold**

❖ $2v - s$ could be any value from $\mathbb{F}$, irrespective of the underlying s

If he thinks $s = a \Rightarrow v = (y_2 + a) \cdot 2^{-1}$

Now, you might be interested to know that why this would not be considered as a breach of privacy, if instead of performing the operations over the integers, I perform operations over a field. How can suddenly changing the algebraic structure make something which is not private with respect to integers make it private suddenly. Why it is not private with integers, but private if I perform the operations over field?

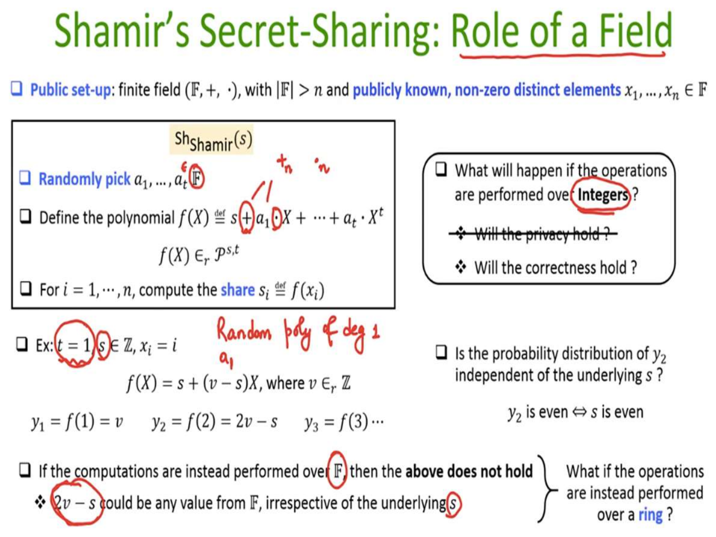
The thing is that, again, if I perform the operations over the field rather performing operations with respect to integers, and if this plus and dot operations are my integer plus and dot operation, then I can no longer say that $2v\ -\ s$ is odd, if and only if s is odd. Because, $2v - s$ could be any element from the field irrespective of whether my s is odd or whether my s is even. It could be any value from the field.

That means, if P_2 is an unauthorised party, he got the value $2v - s$. v is unknown for him, s is unknown for him, but this overall value is known to him. So, I am putting this values in a rectangular box, that means it is unknown for him. And say this overall value is y_2. He has this concrete value of y_2. If he thinks s is equal to say some value a, then that is possible if v would have been y_2 plus a multiplied with 2 inverse.

That means, indeed it could be the case that dealer's secret was the field element a, and the random coefficient v that dealer has chosen is this value, which would have resulted P_2 getting y_2 as the field element, which is $2v - s$. But v, the element v, the coefficient v is a randomly chosen field element. That means, whatever is the value of s that P_2 could guess in his mind, that will give him a candidate v.

And that is why P_2 can no longer narrow down, whether he is seeing a share corresponding to an odd field element or an even field element. But that is not possible if; that is not the case if the instant operations would have been performed over integers, because, in the integer world, we know that $2v\ -\ s$ will be odd, if and only if s is odd, and so on.

(Refer Slide Time: 31:17)



Shamir's Secret-Sharing: Role of a Field

□ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \dots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

□ Randomly pick $a_1, \dots, a_t \overset{\in}{} \mathbb{F}$

$+_n$ $\cdot_n$

□ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

□ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

□ What will happen if the operations are performed over Integers?

~~❖ Will the privacy hold?~~

❖ Will the correctness hold?

□ Ex: $t = 1$, $s \in \mathbb{Z}$, $x_i = i$ Random poly of deg 1

a_1

$f(X) = s + (v - s)X$, where $v \in_r \mathbb{Z}$

$y_1 = f(1) = v \quad y_2 = f(2) = 2v - s \quad y_3 = f(3) \cdots$

□ Is the probability distribution of y_2 independent of the underlying s?

y_2 is even $\Leftrightarrow$ s is even

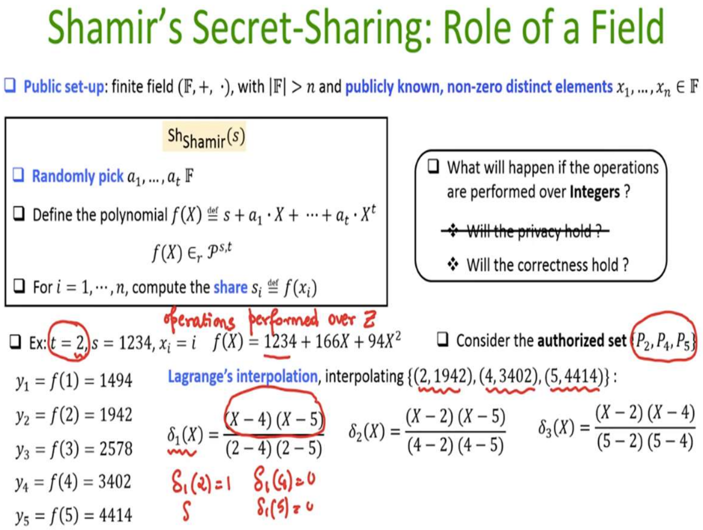
□ If the computations are instead performed over $\mathbb{F}$ then the **above does not hold**

❖ $2v - s$ could be any value from $\mathbb{F}$, irrespective of the underlying s

What if the operations are instead performed over a ring?

So, that is why performing operations over field is important. Now, you might be wondering, why cannot I perform the operations over the ring instead of a field? Again, in the ring, I cannot say that $2v - s$ could be any ring element, if v is randomly chosen ring element. Again, depending upon cases, adversary can narrow down that this could be the candidate s and this cannot be the candidate s. So, the summary is that, in order to achieve the privacy properties, Shamir's secret-sharing scheme has to be instantiated over the field.

(Refer Slide Time: 31:52)



Shamir's Secret-Sharing: Role of a Field

□ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \dots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

□ Randomly pick $a_1, \dots, a_t$ $\mathbb{F}$

□ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

□ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

□ What will happen if the operations are performed over **Integers** ?

~~❖ Will the privacy hold ?~~

❖ Will the correctness hold ?

operations performed over Z

□ Ex: $t = 2$, $s = 1234$, $x_i = i$ $f(X) = 1234 + 166X + 94X^2$

□ Consider the **authorized set** $\{P_2, P_4, P_5\}$

$y_1 = f(1) = 1494$

$y_2 = f(2) = 1942$

$y_3 = f(3) = 2578$

$y_4 = f(4) = 3402$

$y_5 = f(5) = 4414$

Lagrange's interpolation, interpolating $\{(2, 1942), (4, 3402), (5, 4414)\}$:

$\delta_1(X) = \frac{(X-4)\,(X-5)}{(2-4)\,(2-5)}$ $\delta_2(X) = \frac{(X-2)\,(X-5)}{(4-2)\,(4-5)}$ $\delta_3(X) = \frac{(X-2)\,(X-4)}{(5-2)\,(5-4)}$

$\delta_1(2) = 1$ $\delta_1(4) = 0$

S $\delta_1(5) = 0$

Now, what about the correctness property? Will there be a violation if instead of performing the operations over field, I perform operations over integers? So, again, let me demonstrate the problem that might occur. And this is with respect to the correctness. So, for demonstration, assume that my $t = 2$, operations performed over integers. So, suppose $t = 2$, so, that means, dealer has to pick a random polynomial of degree-2 whose coefficients are integer coefficients and whose constant term is the integer 1, 2, 3, 4.

And say, we fix x_1, x_2, x_3, x_4, x_5 to be 1, 2, 3, 4, 5. So, these will be the shares. Now, imagine an authorised subset consisting of the second party, fourth party and the fifth party. That is an authorised set, because my $t = 2$. That means, if 3 shareholders come together, they should be able to get back the secret. So, let us try to interpolate the shares 2, 1942; 4, 3402; 5, 4144. And Lagrange's interpolation should give us back the unique curve.

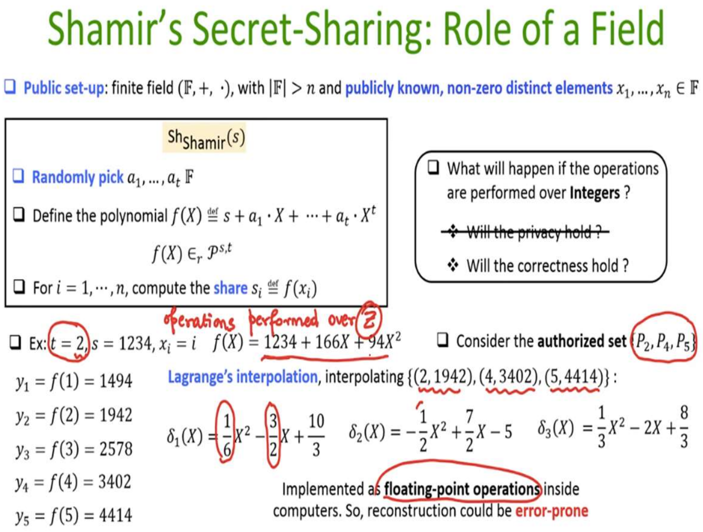
So, remember, in the Lagrange's interpolation, we find out the delta polynomials, and then try to express the unknown polynomial which we want to reconstruct in as a linear combination of the delta polynomials. So, my first delta polynomial should be such that 4 and 5 should

constitute its root. So, that is why, in the numerator, I have written terms $x - 4$ into $x - 5$. And
when evaluated at 2, it should give me the value 1.

That is why in the denominator I have written 2 - 4 and 2 - 5. You can verify the requirements here. So, δ_1 evaluated at 2 will give you 1, and δ_1 evaluated at 4 will be 0; δ_1 evaluated at 5 will be 0. In the same way, the δ_2 polynomial should be such that it should vanish at $x = 2$ and $x = 5$, namely, it should become 0. So, 2 and 5 should be its roots, and it should become 1 at $x = 4$. So, this will be your δ_2 polynomial.

Your δ_3 polynomial should have 2 and 4 as its roots, and it should take the value 1 at $x = 5$. So, these will be your delta polynomials. And now, remember, this division operations are now the integer division operations, because, now I am performing the operations over the integers. And in integer, dividing the numerator by the denominator means, it is an integer division, it is no longer multiplying with the multiplicative inverse here.

(Refer Slide Time: 35:14)



Shamir's Secret-Sharing: Role of a Field

□ Public set-up: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and publicly known, non-zero distinct elements $x_1, \ldots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

□ Randomly pick $a_1, \ldots, a_t$ $\mathbb{F}$

□ Define the polynomial $f(X) \overset{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

□ For $i = 1, \cdots, n$, compute the share $s_i \overset{\text{def}}{=} f(x_i)$

□ What will happen if the operations are performed over Integers ?

~~❖ Will the privacy hold ?~~

❖ Will the correctness hold ?

operations performed over $\mathbb{Z}$

□ Ex: $t = 2$, $s = 1234$, $x_i = i$ $f(X) = 1234 + 166X + 94X^2$

□ Consider the authorized set $\{P_2, P_4, P_5\}$

$y_1 = f(1) = 1494$

$y_2 = f(2) = 1942$

$y_3 = f(3) = 2578$

$y_4 = f(4) = 3402$

$y_5 = f(5) = 4414$

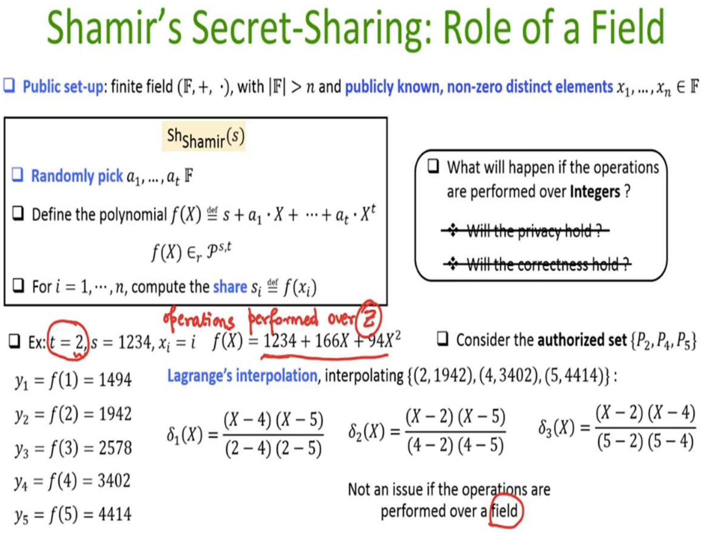
Lagrange's interpolation, interpolating $\{(2, 1942), (4, 3402), (5, 4414)\}$:

$\delta_1(X) = \frac{1}{6}X^2 - \frac{3}{2}X + \frac{10}{3}$ $\delta_2(X) = -\frac{1}{2}X^2 + \frac{7}{2}X - 5$ $\delta_3(X) = \frac{1}{3}X^2 - 2X + \frac{8}{3}$

Implemented as floating-point operations inside computers. So, reconstruction could be error-prone

So, if I now simplify the individual lambda polynomials, it takes this fractional forms here. And now, if I implement this Lagrange's interpolation, this will be implemented as floating point operations. Namely, 1 over 6 will be interpreted as a floating point operation, 3 over 2 will be interpreted as a floating point operations and so on. So, depending upon the precision of your floating point operations as part of your implementation, there might be some error in the recovery process. Namely, there might be some error in reconstructing back the dealer's polynomial $1234 + 166X + 94X^2$, because of this issue.

(Refer Slide Time: 36:05)



Shamir's Secret-Sharing: Role of a Field

❑ **Public set-up**: finite field $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and **publicly known, non-zero distinct elements** $x_1, \ldots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

❑ **Randomly pick** $a_1, \ldots, a_t$ $\mathbb{F}$

❑ Define the polynomial $f(X) \stackrel{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the **share** $s_i \stackrel{\text{def}}{=} f(x_i)$

❑ What will happen if the operations are performed over **Integers** ?

❖ ~~Will the privacy hold ?~~

❖ ~~Will the correctness hold ?~~

operations performed over $\mathbb{Z}$

❑ Ex: $\underset{\sim}{t = 2}$, $s = 1234$, $x_i = i$ $\quad f(X) = \underline{1234 + 166X + 94X^2}$

❑ Consider the **authorized set** $\{P_2, P_4, P_5\}$

$y_1 = f(1) = 1494$

$y_2 = f(2) = 1942$

$y_3 = f(3) = 2578$

$y_4 = f(4) = 3402$

$y_5 = f(5) = 4414$

Lagrange's interpolation, interpolating $\{(2, 1942), (4, 3402), (5, 4414)\}$:

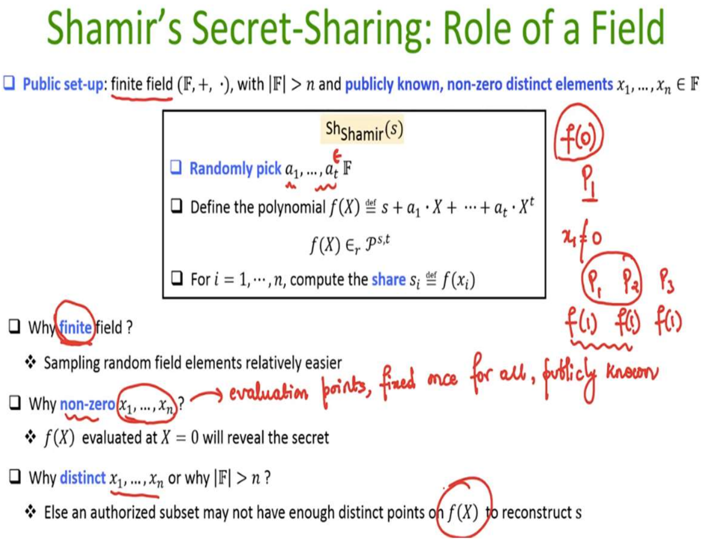
$$\delta_1(X) = \frac{(X-4)\,(X-5)}{(2-4)\,(2-5)} \qquad \delta_2(X) = \frac{(X-2)\,(X-5)}{(4-2)\,(4-5)} \qquad \delta_3(X) = \frac{(X-2)\,(X-4)}{(5-2)\,(5-4)}$$

Not an issue if the operations are performed over a field

Whereas, if you would have performed the operations over the field, then this divisions will not be the integer divisions, rather, the division should be interpreted as multiplying with the multiplicative inverse. So, that is why, if you; and that is possible only when you perform the operations over a field, because multiplicative inverses, they are guaranteed only when you are performing the operations over the field.

If you perform the operations over the ring, not every element is guaranteed to have its multiplicative inverse. So, that is why, for correctness sake also, you need the operations in the Shamir's secret-sharing algorithm to be performed over a field.

(Refer Slide Time: 36:44)



Shamir's Secret-Sharing: Role of a Field

❑ **Public set-up**: $\underset{\sim}{\text{finite field}}$ $(\mathbb{F}, +, \cdot)$, with $|\mathbb{F}| > n$ and **publicly known, non-zero distinct elements** $x_1, \ldots, x_n \in \mathbb{F}$

$\mathsf{Sh}_{\mathsf{Shamir}}(s)$

❑ **Randomly pick** $a_1, \ldots, \overset{t}{a_t}$ $\mathbb{F}$

❑ Define the polynomial $f(X) \stackrel{\text{def}}{=} s + a_1 \cdot X + \cdots + a_t \cdot X^t$

$f(X) \in_r \mathcal{P}^{s,t}$

❑ For $i = 1, \cdots, n$, compute the **share** $s_i \stackrel{\text{def}}{=} f(x_i)$

$f(0)$

$\underline{P_1}$

$x_1 \neq 0$

$P_1 \quad P_2 \quad P_3$

$\underset{\sim}{f(1) \quad f(1)} \quad f(1)$

❑ Why **finite** field ?

❖ Sampling random field elements relatively easier

❑ Why $\underset{\sim}{\textbf{non-zero}}$ $x_1, \ldots, x_n$? → evaluation points, fixed once for all, publicly known

❖ $f(X)$ evaluated at $X = 0$ will reveal the secret

❑ Why **distinct** $\underline{x_1, \ldots, x_n}$ or why $|\mathbb{F}| > n$?

❖ Else an authorized subset may not have enough distinct points on $f(X)$ to reconstruct s

Now, in the public setup, I assume that not only I have a field, but rather we have a finite field. The only restriction was that it should have n + 1 or more number of elements. So, why finite field? The reason we want to operate over a finite field is that, in this step of Shamir's secret-sharing, these coefficients, $a_1, a_2, \dots, a_t$, the coefficients of the sharing polynomial, they are randomly chosen.

And if my field is a finite field, then sampling random field elements is relatively easier. So, that is why I need a finite field. Now, you might be wondering that if I perform the operations over a finite field, but the secret that I want to share is enormously large, it cannot be kind of mapped as an element of field; so, what I can do is, rather I can divide my secret into small pieces; each small piece, I can map to individual field elements, and then run an instance of secret-sharing.

That is not an issue. So, this restriction of using a finite field is not going to lead to any issue if the secret is actually a large element. You always have mechanisms to map that large secret as field elements, the field that you have chosen to instantiate your Shamir's secret-sharing and run the Shamir's secret-sharing algorithm. Now, why you want the $x_1, \dots, x_n$, the evaluation points?

So, this $x_1, \dots, x_n$, they are also called as the evaluation points, they are evaluation points fixed once for all. That means, once it is fixed, then for all the instantiations of Shamir's secret-sharing, the same $x_1, \dots, x_n$ will be used, irrespective of who is the dealer and so on; and publicly known. So, why they have to be non-zero? Why cannot I give say party P_1 the value of the polynomial at $x = 0$?

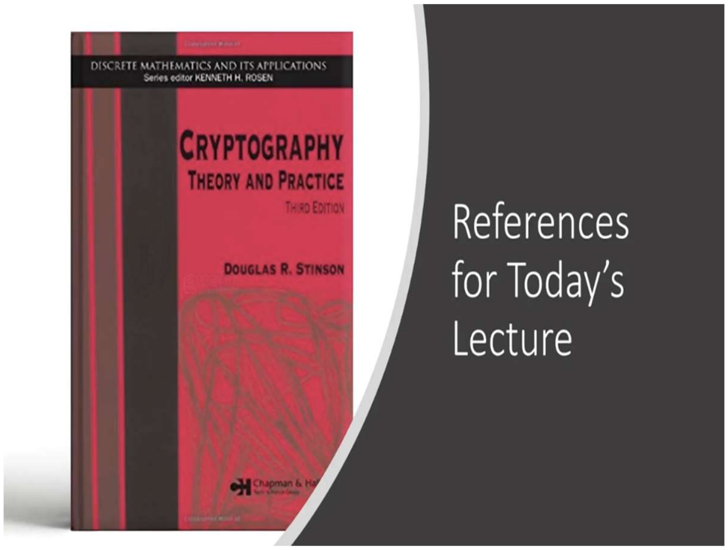
Why cannot I say that x_1; why $x_1 = 0$ not allowed? Well, if I make x_1 to be 0, then basically P_1 is going to learn the value of the secret itself, because $f(0)$ will give the value of the secret which dealer wants to share. And we do not want the P_1 alone just to learn the secret. Only when P_1 is a part of an authorised subset, he should learn the secret. So, that is why, none of these evaluation points we can afford to set as 0.

They have to be the non-zero elements. And now, the final restriction that, why they have to be distinct? Why cannot I give P_1 the value of say $f(1)$ and P_2 also the value of $f(1)$ and P_3

also the value of $f(1)$ and so on? Well, in that case, if say P_1 and P_2, they are authorised set, then they would not have sufficient number of shares or sufficient number of distinct points on the dealer's polynomial to interpolate it back, because both P_1 as well as P_2 have the same point on the dealer's unknown polynomial, even if they constitute authorised subset.

That means, if a subset of t + 1 parties come together, only when we ensure that their evaluation points are distinct, namely the x components are distinct, they basically will have t + 1 distinct points on the dealer's sharing polynomial, and then only the Lagrange's interpolation will work and help you to get back uniquely the dealer's sharing polynomial.

(Refer Slide Time: 40:35)



References for Today's Lecture

So, with that I conclude, with that I end today's lecture. Thank you.